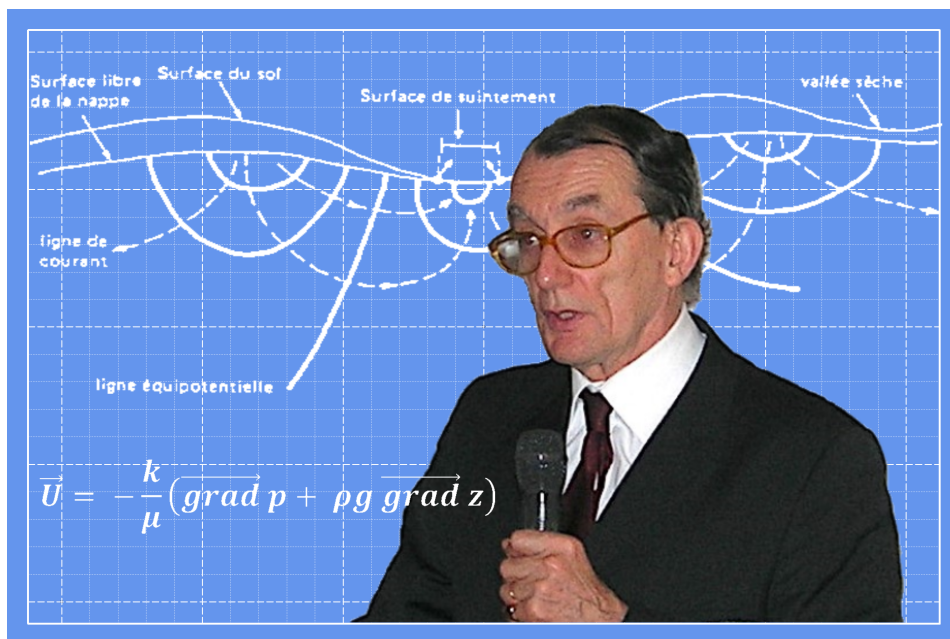


# COMPTES RENDUS DE L'ACADÉMIE DES SCIENCES

1778-7025 (electronic)

## *Géosciences* *Sciences de la Planète*



Volume 355, Special Issue S1, 2023

**Special issue / Numéro thématique**

Geo-hydrological Data & Models / *GDM - Géo-hydrologie, données et modèles*

**Guest editors / Rédacteurs en chef invités**

Vazken Andréassian, Valérie Plagnes,  
Craig Simmons, Pierre Ribstein

# *Comptes Rendus*

---

## *Géoscience*

### **Objective of the journal**

*Comptes Rendus Géoscience* is an internationally peer-reviewed electronic journal covering the full range of earth sciences and sustainable development. It publishes original research articles, review articles, historical perspectives, pedagogical texts, and conference proceedings of unlimited length, in English or French. *Comptes Rendus Géoscience* is published according to a virtuous policy of diamond open access, free of charge for authors (no publication fees) as well as for readers (immediate and permanent open access).

**Editorial director:** Étienne Ghys

**Editors-in-chief:** Éric Calais, Michel Campillo, François Chabaux, Ghislain de Marsily

**Editorial Board:** Jean-Claude André, Pierre Auger, Mustapha Besbes, Sylvie Bourquin, Yves Bréchet, Marie-Lise Chanin, Philippe Davy, Henri Décamps, Sylvie Derenne, Michel Faure, François Forget, Claude Jaupart, Jean Jouzel, Eric Karsenti, Amaëlle Landais, Sandra Lavorel, Yvon Le Maho, Mickaele Le Ravalec, Hervé Le Treut, Benoit Noetinger, Carole Petit, Valérie Plagnes, Pierre Ribstein, Didier Roux, Bruno Scaillet, Marie-Hélène Tusseau-Vuillemin, Élisabeth Vergès

**Editorial secretary:** Adenise Lopes

### **About the journal**

All journal's information, including the text of published articles, which is fully open access, is available from the journal website at <https://comptes-rendus.academie-sciences.fr/geoscience/>.

### **Author enquiries**

For enquiries relating to the submission of articles, please visit this journal's homepage at <https://comptes-rendus.academie-sciences.fr/geoscience/>.

### **Contact**

Académie des sciences  
23, quai de Conti, 75006 Paris, France  
Tel: (+33) (0)1 44 41 43 72  
[CR-Geoscience@academie-sciences.fr](mailto:CR-Geoscience@academie-sciences.fr)



The articles in this journal are published under the license  
Creative Commons Attribution 4.0 International (CC-BY 4.0)  
<https://creativecommons.org/licenses/by/4.0/deed.en>



---

## Contents / Sommaire

<b>Vazken Andréassian, Valérie Plagnes, Craig Simmons, Pierre Ribstein</b> A hydrogeological acoustic: in honour of Ghislain de Marsily .....	1-8
INVERSE METHODS	
<b>Jeremy White, Marsh Lavenue</b> Advances in the pilot point inverse method: Où En Sommes-Nous maintenant? .....	9-17
<b>Przemysław Juda, Julien Straubhaar, Philippe Renard</b> Comparison of three recent discrete stochastic inversion methods and influence of the prior choice .....	19-44
<b>Philippe Ackerer, Jesus Carrera, Frédérick Delay</b> Identification of aquifer heterogeneity through inverse methods .....	45-58
<b>Frederick Delay, Jean-Luc Mari, Gilles Porel, François Chabaux, Philippe Ackerer</b> Is subsurface geophysics as seismic and acoustic investigations a rescue to groundwater flow inversion? .....	59-78
SURFACE HYDROLOGY	
<b>Nicolas Le Moine</b> On the computation of the Basal Envelope Surface of Talwegs using the Analytic Element Method .....	79-97
<b>Ludovic Oudin, Morgane Lalonde</b> Pitfalls of space-time trading when parametrizing a land use dependent hydrological model .....	99-115
<b>Thibault Mathevet, Nicolas Le Moine, Vazken Andréassian, Hoshin Gupta, Ludovic Oudin</b> Multi-objective assessment of hydrological model performances using Nash–Sutcliffe and Kling–Gupta efficiencies on a worldwide large sample of watersheds .....	117-141
<b>Deniz Kilic, Agnès Rivière, Nicolas Gallois, Agnès Ducharne, Shuaitao Wang, Philippe Peylin, Nicolas Flipo</b> Assessing water and energy fluxes in a regional hydrosystem: case study of the Seine basin .....	143-163
<b>Daniel Zimmer, Julien Tournebize, Sami Bouarfa, Cyril Kao, Benoît Lesaffre</b> Land drainage functioning and hydrological impacts in rural catchments: model development and field experiments .....	165-183

### GROUNDWATER HYDROLOGY

<b>François Chabaux, Philippe Négrel, Frederick Gal, Daniel Viville, Coralie Ranchoux, Jérôme Van der Woerd, Frederick Delay, Julien Ackerer, Luc Aquilina, Marc Diraison, Frédéric Masson, Anne-Sophie Mériaux, Thierry Labasque, Yann Lucas, Philippe Adrien Pezard, Thierry Reuschlé, Anne-Désirée Schmitt, Peter Stille, Robert Wyls, Catherine Lerouge, Chrystel Dezayes</b> Hard rock aquifer architecture and water circulation levels in the Strengbach critical zone observatory (France) .....	185-206
<b>Marc Dumont, Valérie Plagnes, Patrick Lachassagne, Roger Guérin, Bayu Nugraha, Febriwan Mohamad, Ludovic Oudin, Arif Fadillah, Danièle Valdès, Gilles Brocard, Jean-Luc Bonjour, Mohamed Saadi, Anne-Sophie Esneu, Aswar Muhammad, Hendar-mawan, Nathalie Dörfliger</b> Water cycle modelling strengthened by probabilistic integration of field data for groundwater management of a quite unknown tropical volcanic hydrosystem .....	207-229
<b>Arnauld Malard, Nathalie Dörfliger, Pierre-Yves Jeannin, Silvan Laube</b> Contribution of the dynamic visualization of a 3D hydrogeological conceptual model to the participatory management of groundwater resources .....	231-244
<b>Anne Jost, Shuaitao Wang, Thomas Verbeke, François Colleoni, Nicolas Flipo</b> Hydrodynamic relationships between gravel pit lakes and aquifers: brief review and insights from numerical investigations .....	245-269
<b>Antoine Pelletier, Vazken Andréassian</b> An underground view of surface hydrology: what can piezometers tell us about river floods and droughts? .....	271-280
<b>Julio Gonçalves, Alexis Nutz, Pierre Séraphin, Amine Chekireb, Lahcen Kabiri, Pierre Deschamps</b> Dealing with hydrologic data scarcity: the case of the Tindouf basin .....	281-300

#### WATER IN THE ANTHROPOCENE

<b>Gilles Billen, Josette Garnier</b> The water-agro-food system: upscaling from the Seine river basin to the global scale .	301-315
<b>Michel Meybeck, Gabrielle Bouleau, Catherine Carré, Josette Garnier, Laurence Lestel</b> Rivers help us to quantify the socio-ecological functioning of their basin at the Anthropocene: the Seine example (1850–2020) .....	317-335

#### MODELLING PHILOSOPHY & LAW

<b>Vazken Andréassian</b> On the (im)possible validation of hydrogeological models .....	337-345
<b>Maria-Theresia Schafmeister</b> High-level radioactive waste repository: How Geology combined with societal principles can lead to public acceptance—the German experiment .....	347-361
<b>Anne Coudrain, Matthieu Le Duff, Danielle Mitja</b> The Anthropocene is shifting the paradigm of geosciences and science .....	363-380

#### CLIMATE CHANGE

<b>Philémon Autin, Jean Emmanuel Sicart, Antoine Rabatel, Regine Hock, Vincent Jomelli</b>	
Climate reconstruction of the Little Ice Age maximum extent of the tropical Zongo Glacier using a distributed energy balance model .....	381-398
<b>Chantal Gascuel-Oudou, Ophélie Fovet, Mikaël Fauchoux, Jordy Salmon-Monviola, Laurent Strohmenger</b>	
How to assess water quality change in temperate headwater catchments of western Europe under climate change: examples and perspectives .....	399-409
<b>Roland Yonaba, Lawani Adjadi Mounirou, Fowé Tazen, Mahamadou Koïta, Angelbert Chabi Biaou, Cheick Oumar Zouré, Pierre Queloz, Harouna Karambiri, Hama Yacouba</b>	
Future climate or land use? Attribution of changes in surface runoff in a typical Sahelian landscape .....	411-438
<b>Justine Tirogo, Anne Jost, Angelbert Biaou, Youssouf Koussoubé, Pierre Ribstein, Denis Dakouré</b>	
Impacts of climate change and pumping on groundwater resources in the Kou River basin, Burkina Faso .....	439-464
<b>Mustapha Besbes, Jamel Chahed</b>	
Predictability of water resources with global climate models. Case of Northern Tunisia	465-486

#### KARSTS & SUBMARINE SPRINGS

<b>Perrine Fleury, Séverin Pistre, Michel Bakalowicz</b>	
Coastal karst aquifers and submarine springs: what future for their water resources? .	487-500
<b>Esteban Sanz, Jesús Carrera, Carlos Ayora, Alfredo Barón, Concha González</b>	
Assessing the salinization mechanisms of coastal brackish springs .....	501-522
<b>David Labat, Thomas Guilloteau, Sébastien Lamagnère, Thaïs Dumareau, Gabin Haryouli, Sandra Béranger, Franck Granouillac, Vianney Sivel</b>	
New insights from an enhancement of hydrogeological archives of a French karst watershed over the last 50 years .....	523-531
<b>Guy Vasseur</b>	
Geothermal heat advected by the recharge of underground conduit. Case study of the karstic spring of Lez (Hérault, France) .....	533-557

#### GEOSTATISTICS

<b>Benoît Noetinger</b>	
Random fields and up scaling, towards a more predictive probabilistic quantitative hydrogeology .....	559-572
<b>Hayet Chihi, Mohamed Amin Hammami, Imen Mezni, Habib Belayouni, Abdallah Ben Mammou</b>	
Multiscale modeling of reservoir systems using geostatistical methods .....	573-603
<b>Chantal de Fouquet, Jérémy Nos</b>	
Cokriging of transmissivity from head measurements revisited in the case of two quasi steady state flows .....	605-615

<b>Vanessa Teles, Youri Hamon, Rémy Deschamps, Sébastien Rohais, Fadi H. Nader, Elodie Heckenmeyer, Marta Gasparrini, Mickael Barbier, Olivier Lerat, Philippe Joseph, Brigitte Doligez</b>	
Modelling the coupled heterogeneities of the lacustrine microbialite-bearing carbonate reservoir of the Yacoraite Formation (Salta, Argentina) .....	617-636

GEOLOGY

<b>Emmanuel Ledoux, Emmanuel Hertz, Jean-Charles Robinet, Pierre Combes</b>	
Reflections on the role of chemical osmosis mechanisms on the long-term behavior of a collapsed salt cavity .....	637-665
<b>Philippe Davy, Romain Le Goc, Caroline Darcel, Jan-Olof Selroos</b>	
Scaling of fractured rock flow. Proposition of indicators for selection of DFN based flow models .....	667-690



Research article

Geo-hydrological Data & Models

# A hydrogeological acrostic: in honour of Ghislain de Marsily

Vazken Andréassian<sup>Ⓢ,\*<sup>a</sup></sup>, Valérie Plagnes<sup>Ⓢ,<sup>b</sup></sup>, Craig Simmons<sup>Ⓢ,<sup>c</sup></sup> and Pierre Ribstein<sup>Ⓢ,<sup>b</sup></sup>

<sup>a</sup> INRAE, UR HYCAR, Antony, France

<sup>b</sup> Sorbonne Université, UMR METIS, Paris, France

<sup>c</sup> University of Newcastle, Newcastle, Australia

E-mails: vazken.andreassian@inrae.fr (V. Andréassian),

valerie.plagnes@sorbonne-universite.fr (V. Plagnes),

Craig.Simmons@newcastle.edu.au (C. Simmons),

pierre.ribstein@sorbonne-universite.fr (P. Ribstein)

*Manuscript received 28 June 2023, accepted 19 July 2023.*

Great scientists are sometimes excessively modest, and ostentatious marks of appreciation may embarrass them. This special issue of *Comptes Rendus Géosciences* was initially thought as a hydrogeological firework launched by former students and current friends and colleagues of Professor Ghislain de Marsily to honour his multifaceted achievements and talents.

However, because we do not want to embarrass him more than strictly needed with our expressions of gratitude and admiration, we tried in this editorial to do so with as much discretion as we could. A recent biographical note by Simmons [2021] has already discussed the career, scientific contributions and international recognition of Ghislain de Marsily. These are not repeated here.

Inverse methods are the focus of several papers of this special issue: White and Lavenue [2023] review the *Pilot Point Method* for model inversion and uncertainty quantification, which has become a reference in the domain of inversion for groundwater

modelling. Juda et al. [2023] compare several stochastic inversion methods, using a synthetic example (a pumping test in a fluvial channelized aquifer). Ackerer et al. [2023] discuss the difficulties that occur during the identification of aquifer heterogeneity using inverse methods and underline how seminal ideas introduced by Ghislain de Marsily helped address key inversion issues (stabilization of the inverse problem and parsimonious reproduction of the natural heterogeneity of the subsurface). Delay et al. [2023] also investigated the flow inversion issue, adopting a geophysical point of view with information produced by a seismic survey: they compare inverted hydraulic conductivity maps with the distribution of porous bodies identified by seismic data, and discuss the disappointing lack of a match. They end up questioning the idea of guiding conventional inversions with a geophysically-based prior guess of the subsurface structure.

Several papers deal with catchment-scale hydrology: Oudin and Lalonde [2023] present an attempt to parameterize the effect of land use within a “Budyko-type” hydrological formula. Using a worldwide set of more than 5000 catchments, they show a clear modulation of land use on evapotranspiration, suggesting larger and lower evaporation rates over

\* Corresponding author.

croplands and urban areas respectively. Mathevet *et al.* [2023] present a diagnosis tool (a modification of the classical Taylor diagram) to analyze simulation results over large catchment sets to simultaneously visualize several error components. Kilic *et al.* [2023] presents an attempt to jointly model water and energy fluxes for the entire Seine catchment, both for surface- and groundwater.

Long-term observatories are also dealt with in this special issue: Chabaux *et al.* [2023] use the Strengbach Critical zone observatory (France), integrating petrological and hydrogeological information from boreholes and piezometers, to characterize water circulation channels at the catchment scale.

Aquifer properties are addressed in several of the papers: Dumont *et al.* [2023] present work on the Andesitic volcanic hydrosystems of Indonesia, using electrical resistivity tomography to characterize the aquifer and further parameterize the hydrological model structure. Malard *et al.* [2023] present a flow animation module (FAM) designed to increase the understanding of dynamic hydrogeological processes within a web-based modelling tool called Visual KARSYS. Jost *et al.* [2023] present their investigations relative to the link between gravel pit lakes and aquifers: they detail a gravel pit lake module developed for the CaWaQS modelling platform, which they test against the LAK package of Modflow.

In order to better understand how aquifers and rivers interact during extreme events, Pelletier and Andréassian [2023] study how yearly extremes of groundwater level and streamflow time series are related, using a large set of catchments and piezometers: the result is an aquifer typology, which distinguishes between aquifers that demonstrate correlations between groundwater level and streamflows only with high-flow events, only with low flows, with both types of events or with none of them.

New numerical approaches to characterize aquifers are investigated by Le Moine [2023], who presents a method for interpolating the elevation between talwegs using the Analytic Element Method and discusses a new index named *Height Above the Basal Envelope Surface of Talwegs (HABEST)* to characterize groundwater levels at catchment scale. Zimmer *et al.* [2023] present a retrospective analysis of the development of the theory of subsurface drainage, showing how a parsimonious parameterization of the physical system was sufficient for pre-

dicting the dynamics of subsurface drain flow rates. Gonçalves *et al.* [2023] propose a conceptual model for the hydrogeological functioning of the Tindouf Aquifer, which spreads over Algeria, Morocco, Mauritania and Western Sahara (a region of lifelong interest for Ghislain de Marsily), building upon geomorphological, geological and hydrological remotely sensed and ground observations. Last, Schafmeister [2023] looks at aquifers from the radioactive waste repository point of view (which was also a topic of major interest for Ghislain de Marsily) and describes the current strategy in Germany, as well as the mistakes made in the seventies.

During the last decade, the issue of Water Security has been a topic of major interest for Ghislain de Marsily, and it is thus natural to see several related papers in this special issue: Billen and Garnier [2023] discuss the concept of water-agro-food system, first at the scale of the Seine basin, then enlarging it to the national, European and global scales. They couple it to the Riverstrahler model, in order to simultaneously address issues of food production, river water quality and nutrient delivery to the coastal zone. Meybeck *et al.* [2023] describe the work undertaken since 1989 in the framework of the interdisciplinary PIREN-Seine research program, combining historical river fluxes, material flows, river ecology, environmental history and political ecology. Last, Coudrain *et al.* [2023] discuss the Anthropocene concept, which tends to move science towards action, in order to actively promote solutions that adapt to changes.

Evaluation of hydrogeological models is a domain where Ghislain de Marsily had a remarkable contribution, being himself involved 30 years ago in the “impossible validation” controversy, which Andréassian [2023] revisits insisting on the necessity to look at the issue from two complementary points of view: model’s explanatory power (theoretical content) and the point of view of its predictive power.

Modelling the impact of climatic changes is the topic of several of the contributions of this special issue. Autin *et al.* [2023] use a distributed energy-balance model to reconstruct the maximum extent of a tropical glacier during the Little Ice Age in the late 17th century, using physically coherent climate scenarios constrained by information taken from paleoclimate proxies and sensitivity studies of past glacier advances. Their results highlight the impor-



tance of the seasonal pattern of precipitation to permit the existence of these glaciers.

An analysis of the complex interaction of water quality and climatic changes is proposed by Gascuel-Oudoux *et al.* [2023], who propose a retrospective analysis based on data from European observatories. They argue for the need of a coupled approach because water quality integrates current and past (legacy) conditions, flow pathways, and biogeochemical reactivity. Besbes and Chahed [2023] discuss the predictability of water resources (blue and green water) made directly with Global Climate Models (GCMs), in the case of Northern Tunisia. The authors argue that the use of the raw GCMs predictions on large basins is possible with precisions comparable to what can be obtained with Regional Climate Models in medium size basins.

Runoff and the specificities of its generation in Sahelian landscapes is addressed by two papers dealing with the possible impact of climate change in Burkina Faso: Yonaba *et al.* [2023] use the Soil and Water Assessment Tool (SWAT) model to evaluate possible future changes in surface runoff in the Tougou catchment, paying special attention to the possible interactions between future climate changes and land use changes. Tirogo *et al.* [2023] focus on the impacts of climate change and pumping on groundwater in the Kou catchment: they show that the observed decrease in groundwater levels is due to the combined effect of precipitation deficits and increased pumping.

Submarine springs and karsts are discussed by Fleury *et al.* [2023], who analyze issues for the conservation of the water resources in coastal karst aquifers: they show that the possibility of natural seawater intrusion makes these aquifers particularly fragile, with the combined effect of the increase of sea level, that of more frequent and severe droughts, and that of an increase in withdrawals from aquifers for urban development. Sanz *et al.* [2023] analyze the controversy on the possible salinization mechanisms for coastal brackish springs (which present the puzzling characteristic to be able to discharge high flow rates with significant salinities at elevations of several meters above sea level). They show that spring salinity is mostly controlled by the weight of the water column flowing towards the spring mouth (for low flow rates) and by energy dissipation (for high flow rates). Labat *et al.* [2023] analyse a long-discharge time series

available at an hourly time step for the Baget karstic spring (French Pyrenees). The analyses confirm the multiscale nature and the non-stationarity of the hydrological response. Last, Vasseur [2023] present a geothermal analysis of the Lez karstic spring (located near Montpellier, France): temperature measurements made above and below the underground karst network allow building a model to estimate the amount of geothermal energy captured by the flow system.

In the field of Geostatistics, four contributions are presented: first, Noetinger [2023] presents a study dealing with the upscaling of random fields, for improving quantitative hydrogeological predictions. He discusses how upscaling techniques allow lowering the dimension of the parameter space and, in the stochastic case, reduce the interaction between the conductivity spatial distribution and the flow pattern. Chihi *et al.* [2023] present a geostatistical analysis of a multiscale groundwater reservoir system, combining geological–geophysical investigations, geostatistics, and 3D geological modelling in the Jeffara basin of southeast Tunisia. The behaviour of a complex heterogeneous lacustrine carbonate formation in Argentina is modelled by Teles *et al.* [2023], using a bivariate plurigaussian geostatistical method to constrain the porosity simulation. Last, de Fouquet and Nos [2023] discuss the cokriging of transmissivity using head measurements and show that the estimation is improved in comparison with the case of head data from a single flow.

Little would have been achieved without a deep understanding of geology: First, Ledoux *et al.* [2023] examine chemical osmosis mechanisms and their impact on a collapsed salt cavity: the goal is to understand the mechanisms likely to generate salt fluxes towards surrounding aquifers and to allow the establishment of brine and dissolved salt balance. Davy *et al.* [2023] focus on fractured rock aquifers and propose three scale-independent indicators describing the fundamental characteristics of the flow/permeability relationship.

Yearning for the end of this already-too-long editorial, it seems natural to allocate these last lines to thank all the participants of this special issue for their enthusiasm, as well as the editorial team of the *Compte Rendus Géosciences*, who helped us achieve what we believe is a well-deserved tribute to our great colleague, friend, researcher and teacher, who has

had a lasting impact on Geosciences over at least half a century.

## Conflicts of interest

Authors have no conflict of interest to declare.

## References

- Ackerer, P., Carrera, J., and Delay, F. (2023). Identification of aquifer heterogeneity through inverse methods. *C. R. Géosci.*, 355(S1), 1–14. Online first.
- Andréassian, V. (2023). On the (im)possible validation of hydrogeological models. *C. R. Géosci.*, 355(S1), 1–9. Online first.
- Autin, P., Sicart, J. E., Rabatel, A., Hock, R., and Jomelli, V. (2023). Climate reconstruction of the Little Ice Age maximum extent of the tropical Zongo Glacier using a distributed energy balance model. *C. R. Géosci.*, 355(S1), 1–18. Online first.
- Besbes, M. and Chahed, J. (2023). Predictability of water resources with global climate models. Case of Northern Tunisia. *C. R. Géosci.*, 355(S1), 1–22. Online first.
- Billen, G. and Garnier, J. (2023). The water-agro-food system: upscaling from the Seine river basin to the global scale. *C. R. Géosci.*, 355(S1), 1–15. Online first.
- Chabaux, F., Lerouge, C., Gal, F., Ranchoux, C., Viville, D., Négrel, P., and Dezayes, C. (2023). Water pathway geometry in the critical zone (Strengbach Critical Zone Observatory—France). *C. R. Géosci.* (in press). Online first.
- Chihi, H., Hammami, M. A., Mezni, I., Belayouni, H., and Mammou, A. B. (2023). Multiscale modeling of reservoir systems using geostatistical methods. *C. R. Géosci.*, 355(S1), 1–31. Online first.
- Coudrain, A., Le Duff, M., and Mitja, D. (2023). The Anthropocene is shifting the paradigm of geosciences and science. *C. R. Géosci.*, 355(S1), 1–18. Online first.
- Davy, P., Le Goc, R., Darcel, C., and Selroos, J.-O. (2023). Scaling of fractured rock flow. Proposition of indicators for selection of DFN based flow models. *C. R. Géosci.*, 355(S1), 1–23. Online first.
- de Fouquet, C. and Nos, J. (2023). Cokriging of transmissivity from head measurements revisited in the case of two quasi steady state flows. *C. R. Géosci.* (in press). Online first.
- Delay, F., Mari, J.-L., Porel, G., Chabaux, F., and Ackerer, P. (2023). Is subsurface geophysics as seismic and acoustic investigations a rescue to groundwater flow inversion? *C. R. Géosci.*, 355(S1), 1–20. Online first.
- Dumont, M., Plagnes, V., Lachassagne, P., Guérin, R., Nugraha, B., Mohamad, F., Oudin, L., Fadillah, A., Valdès, D., Brocard, G., Bonjour, J.-L., Saadi, M., Esneu, A.-S., Muhammad, A., Hendarmawan, and Dörfliger, N. (2023). Water cycle modelling strengthened by probabilistic integration of field data for groundwater management of a quite unknown tropical volcanic hydrosystem. *C. R. Géosci.*, 355(S1), 1–23. Online first.
- Fleury, P., Pistre, S., and Bakalowicz, M. (2023). Coastal karst aquifers and submarine springs: what future for their water resources? *C. R. Géosci.*, 355(S1), 1–14. Online first.
- Gascuel-Oudou, C., Fovet, O., Faucheux, M., Salmon-Monviola, J., and Strohmer, L. (2023). How to assess water quality change in temperate headwater catchments of western Europe under climate change: examples and perspectives. *C. R. Géosci.*, 355(S1), 1–11. Online first.
- Gonçalvès, J., Nutz, A., Séraphin, P., Chekireb, A., Kabiri, L., and Deschamps, P. (2023). Dealing with hydrologic data scarcity: the case of the Tindouf basin. *C. R. Géosci.*, 355(S1), 1–19. Online first.
- Jost, A., Wang, S., Verbeke, T., Colleoni, E., and Flipo, N. (2023). Hydrodynamic relationships between gravel pit lakes and aquifers: brief review and insights from numerical investigations. *C. R. Géosci.*, 355(S1), 1–25. Online first.
- Juda, P., Straubhaar, J., and Renard, P. (2023). Comparison of three recent discrete stochastic inversion methods and influence of the prior choice. *C. R. Géosci.*, 355(S1), 1–26. Online first.
- Kilic, D., Rivière, A., Gallois, N., Ducharme, A., Wang, S., Peylin, P., and Flipo, N. (2023). Assessing water and energy fluxes in a regional hydrosystem: case study of the Seine basin. *C. R. Géosci.*, 355(S1), 1–21. Online first.
- Labat, D., Guilloteau, T., Lamagnère, S., Dumareau, T., Haryouli, G., Béranger, S., Granouillac, E., and Sivelles, V. (2023). New insights from an enhancement of hydrogeological archives of a French karst watershed over the last 50 years. *C. R. Géosci.*,

- 355(S1), 1–9. Online first.
- Le Moine, N. (2023). On the computation of the basal envelope surface of talwegs using the analytic element method. *C. R. Géosci.*, 355(S1), 1–19. Online first.
- Ledoux, E., Hertz, E., Robinet, J.-C., and Combes, P. (2023). Reflections on the role of chemical osmosis mechanisms on the long-term behavior of a collapsed salt cavity. *C. R. Géosci.*, 355(S1), 1–29. Online first.
- Malard, A., Dörfli, N., Jeannin, P.-Y., and Laube, S. (2023). Contribution of the dynamic visualization of a 3D hydrogeological conceptual model to the participatory management of groundwater resources. *C. R. Géosci.*, 355(S1), 1–14. Online first.
- Mathevet, T., Le Moine, N., Andréassian, V., Gupta, H., and Oudin, L. (2023). Multi-objective assessment of hydrological model performances using Nash–Sutcliffe and Kling–Gupta efficiencies on a worldwide large sample of watersheds. *C. R. Géosci.*, 355(S1), 1–25. Online first.
- Meybeck, M., Bouleau, G., Carré, C., Garnier, J., and Lestel, L. (2023). Rivers help us to quantify the socio-ecological functioning of their basin at the Anthropocene: the Seine example (1850–2020). *C. R. Géosci.*, 355(S1), 1–19. Online first.
- Noetinger, B. (2023). Random fields and up scaling, towards a more predictive probabilistic quantitative hydrogeology. *C. R. Géosci.*, 355(S1), 1–14. Online first.
- Oudin, L. and Lalonde, M. (2023). Pitfalls of space-time trading when parametrizing a land use dependent hydrological model. *C. R. Géosci.*, 355(S1), 1–17. Online first.
- Pelletier, A. and Andréassian, V. (2023). An underground view of surface hydrology: what can piezometers tell us about river floods and droughts? *C. R. Géosci.*, 355(S1), 1–11. Online first.
- Sanz, E., Carrera, J., Ayora, C., Barón, A., and González, C. (2023). Assessing the salinization mechanisms of coastal brackish springs. *C. R. Géosci.*, 355(S1), 1–22. Online first.
- Schafmeister, M.-T. (2023). High-level radioactive waste repository: How geology combined with societal principles can lead to public acceptance—the German experiment. *C. R. Géosci.*, 355(S1), 1–15. Online first.
- Simmons, C. T. (2021). Ghislain de Marsily. In Daya Sagar, B. S. et al., editors, *Encyclopedia of Mathematical Geosciences*, Encyclopedia of Earth Sciences Series. Springer Nature, Switzerland AG.
- Teles, V., Hamon, Y., Deschamps, R., Rohais, S., Nader, F. H., Heckenmeyer, E., Gasparrini, M., Barbier, M., Lerat, O., Joseph, P., and Doligez, B. (2023). Modelling the coupled heterogeneities of the lacustrine microbialite-bearing carbonate reservoir of the Yacoraite Formation (Salta, Argentina). *C. R. Géosci.*, 355(S1), 1–20. Online first.
- Tirogo, J., Jost, A., Biaou, A., Koussoubé, Y., Ribstein, P., and Dakouré, D. (2023). Impacts of climate change and pumping on groundwater resources in the Kou River basin, Burkina Faso. *C. R. Géosci.*, 355(S1), 1–25. Online first.
- Vasseur, G. (2023). Geothermal heat advected by the recharge of underground conduit. Case study of the karstic spring of Lez (Hérault, France). *C. R. Géosci.*, 355(S1), 1–25. Online first.
- White, J. and Lavenue, M. (2023). Advances in the pilot point inverse method: Où En Sommes-Nous maintenant? *C. R. Géosci.*, 355(S1), 1–9. Online first.
- Yonaba, R., Mounirou, L. A., Tazen, F., Koïta, M., Biaou, A. C., Zouré, C. O., Queloz, P., Karambiri, H., and Yacouba, H. (2023). Future climate or land use? Attribution of changes in surface runoff in a typical Sahelian landscape. *C. R. Géosci.*, 355(S1), 1–28. Online first.
- Zimmer, D., Tournebize, J., Bouarfa, S., Kao, C., and Lesaffre, B. (2023). Land drainage functioning and hydrological impacts in rural catchments: model development and field experiments. *C. R. Géosci.*, 355(S1), 1–19. Online first.

## Appendix

This appendix lists (in alphabetic order) the original works of Ghislain de Marsily which were cited by the papers in this special issue. This list is obviously not an exhaustive one, and is only aimed at providing a rapidly accessible account of Ghislain de Marsily's scientific production.

- Ahmed, S., **de Marsily**, G., and Talbot, A. (1988). Combined use of hydraulic and electrical properties of an aquifer in a geostatistical estimation of transmissivity. *Groundwater*, 26(1), 78–86.
- Arfib, B. and **de Marsily**, G. (2004). Modeling the salinity of an inland coastal brackish karstic spring

- with a conduit-matrix model. *Water Resour. Res.*, 40(11), article no. W11506, 1–10.
- Arfib, B., **de Marsily**, G., and Ganoulis, J. (2002). Les sources karstiques côtières en Méditerranée: étude des mécanismes de pollution saline de l'Almyros d'Héraklion (Crète), observations et modélisation. *Bull. Soc. Géol. France*, 173(3), 245–253.
- Cacas, M. C., Ledoux, E., **de Marsily**, G., Barbreau, A., Calmels, P., Gaillard, B., and Magritta, R. (1990). Modeling fracture flow with a stochastic discrete fracture network: calibration and validation. 2. The transport model. *Water Resour. Res.*, 26(3), 491–500.
- Cacas, M. C., Ledoux, E., **de Marsily**, G., Tillie, B., Barbreau, A., Durand, E., Feuga, G., and Peaudecerf, P. (1990). Modeling fracture flow with a stochastic discrete fracture network: calibration and validation. 1. The flow model. *Water Resour. Res.*, 26(3), 479–489.
- Castro, M. C., Goblet, P., Ledoux, E., Violette, S., and **de Marsily**, G. (1998). Noble gases as natural tracers of water circulation in the Paris Basin: 2. Calibration of a groundwater flow model using noble gas isotope data, *Water Resour. Res.*, 34(10), 2467–2483.
- Chihi, H., **de Marsily**, G., Belayouni, H., and Yahyaoui, H. (2015). Relationship between tectonic structures and hydrogeochemical compartmentalization in aquifers: Example of the “Jef-fara of Medenine” system, south-east Tunisia. *J. Hydrol. Reg. Stud.*, 4(part B), 410–430.
- Chihi, H., **de Marsily**, G., Bourges, M., and Sbeaa, M. (2016). A constrained geostatistical approach for efficient multilevel aquifer system characterization. *J. Water Resour. Hydraul. Eng.*, 5(3), 80–95.
- de Marsily**, G. (1978). De l'identification des systèmes hydro-géologiques. *Thèse de doctorat d'Etat de l'Université Pierre et Marie Curie*, Paris.
- de Marsily**, G. (1980). Introduction. In: **de Marsily**, G. and Merriam, D. F., editors, (1982). *Predictive Geology with Emphasis on Nuclear-Waste Disposal*, pages xi–xv. Pergamon Press, England.
- de Marsily**, G. (1982). Predictive Geology—with emphasis to nuclear waste disposal. In: **de Marsily**, G. and Merriam, D. F., editors, *Proceedings IAMG at 26th IGC Paris 1980*, Pergamon Press, Oxford, UK.
- de Marsily**, G. (1986). *Quantitative Hydrogeology*, page 440. Academic Press, New York.
- de Marsily**, G. (1994). Quelques réflexions sur l'utilisation des modèles en hydrologie. [Tribune libre]. *Rev. Sci. Eau/J. Water Sci.*, 7(3), 219–234.
- de Marsily**, G. (2003). Perspective importance of the maintenance of temporary ponds in arid climates for the recharge of groundwater. *C. R. Geosci.*, 335, 933–934.
- de Marsily**, G. (2007). An overview of the world's water resources problems in 2050. *Ecohydrol. Hydrobiol.*, 7, 147–155.
- de Marsily**, G. (2008). Eau, changements climatiques, alimentation et évolution démographique. *Rev. Sci. Eau*, 21(2), 111–128.
- de Marsily**, G. (2009). *L'eau, un trésor en partage*. Dunod, Paris.
- de Marsily**, G. (2020). Will we soon run out of water? *Ann. Nutr. Metab.*, 76(1), 10–16.
- de Marsily**, G. and Tardieu, B. (2018). *Stratégie d'utilisation des ressources du sous-sol pour la transition énergétique française : les métaux rares*, Académie des Sciences, Académie des Technologies, Paris, France, <https://www.academie-sciences.fr/fr/Rapports-ouvrages-avis-et-recommandations-de-l-Academie/utilisation-des-ressources-du-sous-sol-pour-la-transition-energetique.html>.
- de Marsily**, G. and Abarca-del-Rio, R. (2016). Water and food in the twenty-first century. In: *Remote Sensing and Water Resources*, pages 313–337. Springer, Cham.
- de Marsily**, G. and Besbes, M. (2017). Les Eaux souterraines. In: *Annales Des Mines - Responsabilité et Environnement*, 2017/2 B°86, pages 25–30.
- de Marsily**, G. and Chabaux, F. (2020). Des géosciences aux sciences de la Planète / From Geoscience to Sciences of the Planet. *C. R. Géosci.*, 352, 1–6.
- de Marsily**, G., Combes, P., and Goblet, P. (1992). Comment on “Ground-water models cannot be validated”, by L. F. Konikow and J. D. Bredehoeft. *Adv. Water Resour.*, 15, 367–369.
- de Marsily**, G., Delay, F., Gonçalves, J., Renard, Ph., Teles, V., and Violette, S. (2005). Dealing with spatial heterogeneity. *Hydrogeol. J.*, 13, 161–183.
- de Marsily**, G., Delay, F., Teles, V., and Schafmeister, M. T. (1998). Some current methods to represent the heterogeneity of natural media in hydrogeology. *Hydrogeol. J.*, 6, 115–130.
- de Marsily**, G., Delhomme, J. P., Coudrain-Ribstein,

- A., and Lavenue, A. M. (2000). Four decades of inverse problems in hydrogeology. In: Zhang, D. and Winter, C. L., editors, *Theory, Modeling, and Field Investigation in Hydrogeology: A Special Volume in Honor of Shlomo P. Neuman's 60th Birthday*, Geological Society of America Special Paper 348, pages 1–17. Boulder, Colorado.
- de Marsily**, G., Delhomme, J.-P., Delay, F., and Buoro, A. (1999). Regards sur 40 ans de problèmes inverses en hydrogéologie. *C. R. Acad. Sci. Ser. IIA Earth Planet. Sci.*, 329(2), 73–87.
- de Marsily**, G. and Lallier-Vergès, E. (2015). Les recherches sur l'eau: de nombreux défis pour l'avenir. In: *L'eau à découvert*, éditions du CNRS, pages 320–321.
- de Marsily**, G., Lavedan, G., Boucher, M., and Fasamino, G. (1984). Interpretation of interference tests in a well field using geostatistical techniques to fit the permeability distribution in a reservoir model. In: *Geostatistics for Natural Resources Characterization*, pages 831–849. NATO advanced Study Institute.
- de Marsily**, G., Ledoux, E., Levassor, A., Poitral, D., and Salem, A. (1978). Modelling of large multi-layered aquifer systems: theory and applications. *J. Hydrol.*, 36, 1–34.
- de Marsily**, G. and Besbes, M. (2017). Les eaux souterraines. In: *Responsabilités et Environnement, 2017/2 (N°86)*, pages 25–30.
- Dieng, B., Ledoux, E., and **de Marsily**, G. (1990). Palaeohydrogeology of the Senegal sedimentary basin: a tentative explanation of the piezometric depressions. *J. Hydrol.*, 118, 357–371.
- Emselem, Y. and **de Marsily**, G. (1971). An automatic solution for the inverse problem. *Water Resour. Res.*, 7(5), 1264–1283.
- Fleury, P., Bakalowicz, M., and **de Marsily**, G. (2007). Submarine springs and coastal karst aquifers: A review. *J. Hydrol.*, 339, 79–92.
- Fleury, P., Bakalowicz, M., **de Marsily**, G., and Cortes, J. M. (2007). Functioning of a coastal karstic system with a submarine outlet, in southern Spain. *Hydrogeol. J.*, 16(1), 75–85.
- Garnier, J., Meybeck, M., Ayrault, S., Billen, G., Blanchoud, H., Carré, C., Flipo, N., Gasperi, J., Lestel, L., **de Marsily**, G., Mouchel, J. M., Servais, P., and Tales, E. (2022). Continental Atlantic rivers: the Seine basin. In: Tockner, K., Zarfl, C., and Robinson, C. T., editors, *Rivers of Europe*, pages 291–330. Elsevier, 2nd edition.
- Lavenue, A. M., RamaRao, B. S., **de Marsily**, G., and Marietta, M. G. (1995). Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields: 2. Application. *Water Resour. Res.*, 31(3), 495–516.
- Lavenue, M. and **de Marsily**, G. (2001). Three-dimensional interference test interpretation in a fractured aquifer using the pilot point inverse method. *Water Resour. Res.*, 37(11), 2659–2675.
- Ledoux, E., Girard, G., **de Marsily**, G., Villeneuve, J. P., and Deschenes, J. (1989). *Spatially Distributed Modeling: Conceptual Approach, Coupling Surface Water and Groundwater*, pages 435–454. Springer Netherlands, Dordrecht.
- Matheron, G. and **de Marsily**, G. (1980). Is transport in porous media always diffusive? a counterexample. *Water Resour. Res.*, 16(5), 901–917.
- Meybeck, M., Fustec, E., and **de Marsily**, G., 1998. *La Seine en son bassin: fonctionnement écologique d'un système fluvial anthropisé*, page 736. Elsevier, Masson.
- RamaRao, B. S., Lavenue, A. M., **de Marsily**, G., and Marietta, M. G. (1995). Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields: 1. Theory and computational experiments. *Water Resour. Res.*, 31(3), 475–493.
- Renard, P. and **de Marsily**, G. (1997). Calculating equivalent permeability: a review. *Adv. Water Resour.*, 20(5), 253–278.
- Rousseau-Gueutin, P., Love, A. J., Vasseur, G., Robinson, N. I., Simmons, C. T., and **de Marsily**, G. (2013). Time to reach near-steady state in large aquifers. *Water Resour. Res.*, 49, 6893–6908.
- Schafmeister, M.-Th. and **de Marsily**, G. (1994). The influence of correlation length of highly conductive zones in alluvial media on the transport behavior. In: Dracos, T. and Stauffer, S., editors, *Transport and Reactive Processes in Aquifers*, pages 171–176. Balkema, Rotterdam.
- Teles, V., Bravard, J.-P., **de Marsily**, G., and Perrier, E. (2001). Modeling of the construction of the Rhône alluvial plain since 15,000 years BP. *Sedimentology*, 48, 1209–1224.
- Teles, V., Delay, F., **de Marsily**, G. (2004). Comparison of genesis and geostatistical methods for characterizing the heterogeneity of alluvial media: Groundwater flow and simulations. *J. Hydrol.*, 294,

103–121.

- Zimmerman, D., **de Marsily**, G., Gotway, C. A., Marietta, M. G., Axness, C. L., Beauheim, R. L., Bras, R. L., Carrera, J., Dagan, G., Davies, P. B., Gallegos, D. P., Galli, A., Gómez-Hernández, J., Grindrod, P., Gutjahr, A. L., Kitanidis, P. K., Lavenue, A. M., McLaughlin, D., Neuman, S. P., RamaRao, B. S., Ravenne, C., and Rubin, Y. (1998). A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow. *Water Resour. Res.*, 34(6), 1373–1413.

# **Inverse Methods**







---

Research article

Geo-hydrological Data & Models

# Advances in the pilot point inverse method: OÙ En Sommes-Nous maintenant?

Jeremy White<sup>Ⓢ,\*</sup>,<sup>a</sup> and Marsh Lavenue<sup>Ⓢ</sup>,<sup>b</sup>

<sup>a</sup> INTERA, Inc., Fort Collins, Colorado, 80525, USA

<sup>b</sup> INTERA, Inc., Austin, Texas, 78759, USA

E-mails: [jwhite@intera.com](mailto:jwhite@intera.com) (J. White), [mlavenue@intera.com](mailto:mlavenue@intera.com) (M. Lavenue)

**Abstract.** At a conference some years ago, one of the attendees came up to Ghislain de Marsily and asked, “Excuse me, but aren’t you the de Marsily who developed the Pilot Point Method? It is an honor to meet you.” His response highlights one of Ghislain’s greatest qualities, his humility: “Yes I am, thank you, you are very kind, but that was a very long time ago.” Since Ghislain de Marsily first developed the Pilot Point Method (PPM) in 1978, its development and use has grown significantly in applied decision-support modeling settings including hydrogeology, as well as in other industries, e.g., petroleum reservoir engineering. A technique that was once confined to academic realms, the PPM is now widely accepted as one of the industry pillars of inversion and uncertainty quantification for predictive groundwater modeling. Herein, we provide an update to de Marsily’s paper entitled “Four Decades of Inverse Problems in Hydrogeology” [De Marsily et al., 2000], but with a particular focus on the incredible adoption and advancement of de Marsily’s PPM and related inverse techniques over the last twenty years in the field of predictive groundwater modeling.

Much has been written about the vast array of inverse techniques developed by researchers and practitioners since the 1960s. de Marsily’s PPM, like many methods developed in the late 70s and early 80s, structured its approach to parameterization to overcome many of the challenges of applying inverse methods to real world problems, namely, limited head and transmissivity data relative to the number of unknowns to be estimated, measurement errors, inferred covariance structures of the state variables, and limited computational resources. While inversion research continues, only de Marsily’s PPM has achieved wide-spread adoption within the groundwater modeling community.

The reasons for the popularity of the PPM are many but none more important than its adoption by John Doherty’s PEST inversion software in the early 2000s [Doherty, 2003]. This paper will review the advancements and adaptations of the PPM over the last two decades which have led to its ubiquity. Significant reductions in inversion time are transforming the way in which practitioners are deploying the PPM to improve their understanding of hydrogeologic systems, and, ultimately, to provide decision support for water resource management. The paper ends with newly developed applications of the PPM, given modern machine learning capabilities, and some foreshadowing as to where the PPM might evolve.

**Keywords.** Geostatistics, Groundwater modeling, Pilot points, Inversion, Uncertainty analysis.

*Manuscript received 1 April 2022, revised 8 June 2022, accepted 12 September 2022.*

---

\* Corresponding author.

## 1. Introduction

Since the 1970's, the advancements in computer models and associated tools and techniques used to simulate groundwater systems have enhanced hydrogeologists' understanding of the uncertainty in important simulated outcomes, and ultimately, the risks of decisions related to water supply planning, groundwater remediation and other commercial or academic pursuits. These models, tools and techniques collectively make up an important facet of water resource management decision support which aspire to simulate groundwater flow over space and time and the associated uncertainty in the simulation results. Appropriately representing the salient aspects of complex natural groundwater system with a necessarily simple numerical representation and then combining that simplified representation with formal inverse and data assimilation techniques to extract meaningful information from available state observations is one of the drivers behind the advancements in predictive groundwater modeling. Transformation from data to decisions has therefore been the underlying motivation for research in the area of decision support systems. In this regard, Ghislain de Marsily's contribution to the advancement of the inverse methods, a key component of decision support systems, cannot be understated.

Prior to de Marsily's doctoral research published in French in 1978, modelers had resorted to very simplistic representations of the hydrogeologic data obtained from field investigations due to the lack of techniques to appropriately leverage information embedded in the data. These simplistic techniques typically included the use of zones of pre-determined shape and of large spatial extent, which were assigned constant values for hydrogeologic properties such as transmissivity. The zones could not represent the expected (and necessarily stochastic) patterns of spatial heterogeneity often observed in hydrogeologic system properties at many scales. Therefore, the solutions from these early models were plagued with large and unknown errors initiating from the blocky, unrealistic representation of hydrogeologic properties used in the models and propagated to the decisions these models were meant to inform. The earliest attempts to use formal optimization methods to invert the solution of the groundwater flow equation were thwarted when it was discovered that errors

in the hydraulic head measurements, and moreover in the hydraulic gradient, could lead to instabilities in the inverse solution [Emsellem and De Marsily, 1971, Kleinecke, 1971, De Marsily et al., 2000].

de Marsily's doctoral research presented an alternative inverse method which overcame these challenges. He posed the coupling of a new field of research referred to as geostatistics, or spatial statistics, to represent the groundwater system property fields as continuously varying surfaces, thereby overcoming the known limitations of zonation-based schemes. He parameterized the inversion using an indirect inversion technique which enabled him to overcome the numerical instabilities from errors in the measured head data. Lastly, he also adopted the recently published adjoint technique [Chavent, 1974] to efficiently compute sensitivities of the computed head field to changes in the parameter values enabling him to quickly obtain an inverse solution in high dimensions. de Marsily called his technique the "Pilot Point Method" (PPM) because of the way in which he used specific locations to impose changes in parameter field, locations he called "pilot points". The pilot points would be added to the geostatistical kriging equations and their values assigned optimally to change or "warp" the modeled transmissivity field to reduce the error between the simulated groundwater levels and their measured counterparts. The PPM method fundamentally changed the way groundwater modelers approached the inverse method and is still ubiquitous in groundwater modeling today.

Herein, we provide a brief overview of the theory and history of the PPM, focusing on the important contributions of de Marsily. We finish with some brief remarks related to some of the more recent advances in the PPM and how it continues to serve the decision-support groundwater modeling community.

## 2. Concepts and theory of the PPM

In the PPM, the Bayesian viewpoint of the groundwater model inverse problem is adopted. Originally, the PPM focused on the estimation of the modeled transmissivity field. The log of measured transmissivity is used to transform the statistical distribution of transmissivity (considered to be logarithmic; [Law, 1944, Sánchez-Vila et al., 1996]) to be Gaussian. The unknown log transmissivities in between measurements are viewed as variables of a

Multi-Gaussian random function (RF). Log transmissivity at a point in space is considered a Gaussian random variable (RV) which may or may not have a distribution which is statistically dependent on, or correlated to, its neighboring transmissivities. A transmissivity measurement at location  $(x, y, z)$  is conceptually considered as a sample from the RV distribution at point  $(x, y, z)$ . As presented by Matheron and Blondel [1962], Matheron [1963], kriging provides the best unbiased estimate of the RV at any point in space given observed data and the correlation between the data.

In de Marsily's work, the RV was log transmissivity and the estimation points were the grid block centroids of his finite-difference groundwater flow model. He employed ordinary kriging (OK) to estimate the initial log transmissivity field. OK provides estimates of a RV at any point (or area) in space through the use of measurements of the RV already available and their covariance [see Chilès and Desassis, 2018]. The OK estimator is given by:

$$Z_m^*(u) = \sum_{\beta=1}^n v_{\beta,m} Z(u_{\beta}) \quad (1)$$

where  $Z_m^*(u)$  is the log of the transmissivity estimate at the centroid of grid block  $m$  (located at  $(u) = (x_1, x_2, x_3)$ ),  $Z(u_{\beta})$  is the log of the measured transmissivity data at point  $u_{\beta}$ , and  $v_{\beta,m}$  is the kriging weight assigned to measurement location  $\beta$  given the estimation point  $(u)$ . The covariance of the RV affects the magnitude of the kriging weights through the solution of the following system of equations used to determine  $v_{\beta,m}$ .

$$\begin{cases} \sum_{\beta=1}^n v_{\beta,m} C(u_{\beta} - u_{\alpha}) + \mu(u) = C(u - u_{\alpha}), \alpha = 1, \dots, n \\ \sum_{\beta=1}^n v_{\beta,m} = 1 \end{cases} \quad (2)$$

where the  $v_{\beta,m}$  are the OK weights and  $C(u_{\beta} - u_{\alpha})$  is the covariance of the RV between measurements points  $\beta$  and  $\alpha$ .  $\mu(u)$  is the LaGrange parameter associated with the constraint in the second expression in Equation (2) designed to "filter" out the local mean value to preserve stationarity.

The solution of the inverse problem using the PPM relies upon an expanded form of (1):

$$Z_m^*(u) = \sum_{\beta=1}^n v_{\beta,m} Z(u_{\beta}) + \sum_{p=1}^N v_{p,m} Z(u_p) \quad (3)$$

where  $Z(u_p)$  is the log transmissivity at a pilot point,  $v_{p,m}$  is its associated kriging weight and  $N$  is the total number of pilot points. The location and number of the pilot points were subjectively chosen by de Marsily. He suggested the number of pilot points to be less than the observed transmissivity values and that the pilot points should be placed in areas with high hydraulic-head gradient. As illustrated in Equation (2), the kriging weights  $v_{\beta,m}$  (and thus  $v_{p,m}$ ) do not depend on actual values of the RF,  $Z(u)$ . Thus, they may be solved for, prior to assigning transmissivities estimates/values to the pilot points. de Marsily's PPM thus consists of estimating  $Z(u_p)$  so as to minimize the weighted least-squares objective function of the measured-to-simulated residual vector,  $\phi_{\hat{z}}$ :

$$\phi_{\hat{z}} = (\mathbf{h}^* - \hat{\mathbf{h}})^T \mathbf{V}_h^{-1} (\mathbf{h}^* - \hat{\mathbf{h}}) \quad (4)$$

where  $\mathbf{h}^*$  is the vector of simulated groundwater level,  $\hat{\mathbf{h}}$  is the corresponding measured values of groundwater level, and  $\mathbf{V}_h$  is the covariance matrix of measurement noise.

The sensitivity derivatives required to assign the optimal value of pilot-point transmissivities are computed with the adjoint technique [Chavent, 1974] and the kriging equations (Equation (2)). The overall equation solved by the PPM may be presented as follows.

Let  $Z_m$  represent the log transmissivity value assigned to grid block  $m$ . Using the chain rule, the sensitivity of the weighted least-squares objective function,  $\phi$ , to the log transmissivity value assigned to each pilot point may be expressed by:

$$\frac{d\phi}{dZ_p} = \sum_{m=1}^M \frac{\partial\phi}{\partial Z_m} \frac{\partial Z_m}{\partial Z_p} \quad (5)$$

where  $M$  is the total number of grid blocks in the flow model. The second term on the right-hand side of Equation (5) may be reduced by taking the derivative of Equation (3) with respect to pilot-point transmissivity. Therefore,

$$\frac{\partial Z_m^*(u_m)}{\partial Z_p(u_p)} = v_{p,m} \quad (6)$$

Substituting the right-hand side of Equation (6) into Equation (5) and simplifying yields,

$$\frac{d\phi}{dZ_p} = 2.3 \sum_{m=1}^M v_{p,m} k_m \frac{dk_m}{dk_m} \quad (7)$$

where  $k$  is the permeability of grid block  $m$ . The sensitivity coefficient,  $d\phi/dk_m$ , can be obtained by

adjoint sensitivity analysis or by perturbation methods and used to determine the optimal values of the pilot-point's log transmissivity by modifying the initial guess of the pilot point transmissivities (taken as the kriged estimate at  $\nu_p$  obtained from the measured data). de Marsily states that constraints to the pilot-point transmissivities could be applied to ensure that the pilot-point transmissivities lie within  $\pm 2$  standard deviations of the kriged estimate.

As mentioned earlier, de Marsily's PPM was the first to solve many of the problems inherent in the earlier inverse techniques. By incorporating geostatistics and logarithms of transmissivity, he ensured a smooth but spatially varying transmissivity field even in the presence of significant errors in the head field. In addition, he stabilized the inverse solution by:

- adding prior information, e.g. regularization, in the form of initial kriged estimates at the pilot point locations,
- properly posing the inverse problem by reducing the number of unknowns to a small number of pilot points,
- implementing the transmissivity changes in the neighborhood of a pilot point in accord with the variogram, and
- utilizing efficient adjoint sensitivity techniques to obtain the necessary derivatives for the optimization routine.

### 3. Early PPM advances

The first mention of the PPM was in de Marsily's PhD dissertation [De Marsily, 1978], which describes using "pilot points" at discrete locations. One of the first applications of the PPM published in English was presented in De Marsily et al. [1984]. Here the PPM was used to obtain the permeability field which reproduced interference tests in a well field.

In the late 1980s and early 1990s, research of geostatistical methods by Ahmed and De Marsily [1987] and research of the inverse problem by Certes and De Marsily [1991] resulted in improvements to the PPM. Lavenue and Pickens [1992] coupled ordinary kriging and the PPM to history match a regional groundwater model at the Waste Isolation Pilot Plan (WIPP) site in Carlsbad, New Mexico. They also adopted an adjoint sensitivity method to optimally locate pilot points, which were sequentially

added to the model during an iterative inversion process. One of the significant features added to the model transmissivities during inversion was a high-transmissivity fracture zone which provided a fast offsite pathway to the WIPP-site boundary potentially threatening the WIPP site safety assessment.

A technical review panel comprising the leading groundwater inversion researchers of the 1990s was convened by Sandia National Laboratories to review the WIPP model construction, inversion and travel time calculations. This panel, called the Geostatistics Expert Group (GXG), was headed by Ghislain de Marsily. The GXG recommended that the PPM approach be expanded to calibrate an ensemble of conditionally simulated transmissivity fields in order to investigate groundwater travel time uncertainty to the WIPP-site boundary. de Marsily, Lavenue, and RamaRao did so by linking a conditional simulation front-end routine (using turning bands), a flow model (SWIFT II), an adjoint sensitivity routine to optimally located pilot points and a conjugate gradient optimization routine to assign the pilot-point transmissivity values. In 1995, the resulting new PPM method referred to as GRASP-INV and its application to the WIPP site produced an ensemble of 100 conditionally simulated transmissivity fields [Rama-rao et al., 1995, Lavenue et al., 1995].

This extension of the PPM method to produce an ensemble of solutions, all equally calibrated or "history matched" to the observed data, was a significant step forward in the quantification of posterior uncertainty, which ultimately allowed for a more complete risk assessment to support decision making. It also laid a foundation for the use of inversion and sensitivity techniques to assess data worth which guided future data acquisition campaigns at the WIPP site to further reduce uncertainties in our understanding of the hydrogeologic system.

Shortly afterward, Gomez-Hernandez and his fellow researchers at the University of Valencia, Spain, developed a non-linear pilot-point technique (or a "pilot-point-like" method as one reviewer pointed out) that differs from the original approach of De Marsily [1978] and Lavenue et al. [1995]. These works include Sahuquillo et al. [1992], Gómez-Hernández et al. [1997], and Capilla et al. [1997]. The solution of the inverse problem using Gomez-Hernandez' approach is significantly faster than that of Lavenue and Pickens [1992] or Certes and

De Marsily [1991] for several possible reasons. First, Gomez-Hernandez did not update the sensitivity derivatives as frequently as the other pilot-point approaches. The second major difference in Gomez-Hernandez's method is that he assumed the conductance between two finite-difference grid blocks is expressed by the geometric mean (as opposed to the harmonic mean) of the associated grid-block transmissivities, which significantly reduced the time required to compute sensitivities. The combined effect of these differences significantly sped up the inversion process using Gomez-Hernandez's technique which he called the sequential self-calibration (SSC) pilot point method.

Sandia National Laboratory decided to benchmark GRASP-INV against other inversion techniques represented by the GXG members and invited Gomez-Hernandez to participate. This unique bench-marking exercise, described in Zimmerman et al. [1998], evaluated the performance of virtually every inversion technique used by industry and academics at that time. The ultimate objective of the study was to determine which of several geostatistical inverse techniques is better suited for making probabilistic forecasts of the potential transport of solutes in an aquifer where spatial variability and uncertainty in hydrogeologic properties are significant. Seven geostatistical methods were compared on four synthetic data sets. The inverse methods tested were categorized as being either linearized or non-linear [Carrera and Glorioso, 1991]. The linearized approaches are generally based upon simplifying assumptions about the flow field (e.g., a uniform hydraulic head gradient, a small  $\log(T)$  variance, etc.), that lead to a linearized relation between  $T$  and head using a perturbation expansion of the head and transmissivity fields. This equation can then be solved analytically or numerically. The nonlinear approaches have no such restrictions placed on them and can, in principle, handle more complex flow fields or larger  $\log(T)$  variances. Linearized methods included the Fast Fourier Transform inverse method by Gutjahr et al. [1994], the Linearized Cokriging Method by Kitanidis and Lane [1985], and the Linearized Semianalytical method by Rubin et al. [1992], while the non-linear methods included the Fractal Simulation inverse method by Grindrod and Impey [1991], the Maximum Likelihood inverse method by Carrera [1994], the Pilot Point Method [Rama-

rao et al., 1995] and the SSC pilot point method by Gomez-Hernandez [Gómez-Hernández et al., 1997]. While the non-linear inverse methods were superior to the linearized inverse methods, the most robust results came from the inverse methods that were able to incorporate rigorous geostatistics into the inversion process, that is, inversion suffered when a technique could not adequately represent nor modify the expected spatial (e.g. Prior) patterns of variability of the aquifer properties.

Lessons learned in Zimmerman et al. [1998] regarding the importance of robust geostatistics led de Marsily, Lavenue and RamaRao to enhance their PPM by adopting the GSLIB [Deutsch and Journel, 1992] routines for conditional simulation. They collaborated with Gomez-Hernandez in this effort which resulted in a new conditional simulator for the PPM which could produce an ensemble of multi-facies geostatistical fields with unique covariance structures using categorical indicator simulation followed by sequential gaussian simulation to overlay spatial variability within each facies. de Marsily, Lavenue and RamaRao linked this new capability to their PPM method and later applied it to the WIPP site regional model [Lavenue, 1998]. Unique to this enhancement was the ability of the PPM to optimize the spatial variability within each facies (fractured and unfractured zones) separately to match transient water levels. It also enabled the delineation of a sharp fractured/unfractured zone interface within the regional aquifer which reduced travel time uncertainty. Lavenue and de Marsily then extended this version of the PPM to three dimensions and demonstrated its utility by matching a series of three-dimensional interference tests [Lavenue and De Marsily, 2001].

Other research motivated by Zimmerman et al. [1998] included methods to enforce prior plausibility in a Bayesian sense, the work of Alcolea et al. [2006], Riva et al. [2010] and later Jiménez et al. [2016] worked to include expert knowledge into the PPM solution process so that the resulting property fields were adjusted in harmony with the prior estimates.

#### 4. The tipping point

Notwithstanding the developments and advancements of the PPM cited above, its use was not widespread since the PPM was largely coupled to

groundwater model software not widely used by industry. This changed in the late 1990s, when John Doherty, adopted the PPM for his PEST software suite and produced a full suite of advanced software tools that allowed the PPM to be applied at scale in an unlimited number of settings. Unlike the existing adjoint-based PPM approaches, the PEST software suite relies on non-intrusive/model-independent approaches, such as finite-difference derivatives to estimate the first-order relation between parameters and simulation results, so that it can be applied to practically any forward numerical groundwater model. Doherty realized that like PEST, the PPM can be implemented in a non-intrusive/model-independent approach, so that the PPM could also be applied to practically any forward numerical groundwater model. This shifted the use of the PPM from the bespoke research and development projects cited above to the broader groundwater modeling community, a tipping point of scaling the PPM to industry at large.

Within the PEST framework, the PPM serves as an indispensable parameterization scheme, one that balances the need to express uncertain spatial heterogeneity in model inputs with the requirement to maintain a computationally tractable inverse problem and represent broad-scale heterogeneity. We generalize from transmissivity to “model inputs” here because with the non-intrusive PEST framework, pilot points are commonly used to represent realistic spatial heterogeneity in a wide-range of uncertain properties and boundary condition elements that numerical models require. For example, Knowling and Werner [2016] and later Knowling and Werner [2017] used pilot points to represent uncertain spatial patterns of recharge, while McKenna et al. [2020] and later White et al. [2020a] used pilot points in a multi-scale parameterization scheme to explicitly represent different scales of heterogeneity and associated uncertainty.

The issue of inverse problem stability and posedness with the PPM were both overcome through the adoption of formal Tikhonov regularization [Tikhonov, 1963, Tikhonov and Arsenin, 1977] [see Tonkin and Doherty, 2005, Fienen et al., 2009, Doherty et al., 2010, Doherty, 2015], in combination with truncated subspace solution techniques [see Oliver et al., 2008, Aster et al., 2018, Doherty, 2015]. These two advancements, which were implemented

at scale to function in concert with each other for the first time in PEST software suite, allowed practitioners to forego the need to *a priori* reduce the number of pilot points to stabilize the inverse problem—Tikhonov regularization and truncated subspace techniques optimally and unconditionally stabilize the inverse problem. Now practitioners were free to express expected property heterogeneity (and associated uncertainty) at scales that are relevant to the predictive outcomes of the model.

## 5. OÙ EN SOMMES-NOUS MAINTENANT?

With the recent advancements in the PPM mentioned above, the combined use of the PPM and PEST has led to wide-spread investigation of predictive uncertainties in applied groundwater modeling, which is facilitated by the PPM allowance for a more complete and robust expression of expected patterns of spatial heterogeneity in subsurface properties [see for example Moore and Doherty, 2005, James et al., 2009, Dausman et al., 2010, Herckenrath et al., 2011, Tonkin and Doherty, 2009, Christensen and Doherty, 2008, Keller et al., 2021].

The PPM, as embodied within Doherty’s PEST software suite, continues to serve an important role in decision-support environmental simulation around the world and new developments continue. Several recent advancements have focused on expanding the use of pilot points from 2-point geostatistics to multipoint geostatistics [see Li et al., 2013, Mariethoz and Caers, 2014, Ma and Jafarpour, 2018, Khambhammettu et al., 2020, Liu et al., 2021]. The capability of these pilot point methods to facilitate the use of multi-point geostatistical property representations within formal data assimilation is a major advancement in settings where high-order multi-point geostatistical property representations are important, such as settings where connected permeability are important for representing the fate and transport of dissolved-phase constituents.

So “Where are We Now?” Groundwater modelers have more computational power than ever before to build highly complex models and invert them with PEST, and more recently, the PEST++ software suite [White et al., 2020b]. With the related advancements in history matching using the iterative ensemble smoother approaches [e.g. Chen and Oliver, 2013, White, 2018], practitioners are able to more

efficiently employ high-dimensional inversion and uncertainty quantification in non-intrusive ways and at scales not previously possible except for those with access to high-performance computing. This “gilded age” of inversion however comes with the responsibility of parsimony. It cannot be overstated that sound groundwater model analyses are built on the foundation of a solid hydrogeologic conceptual model that is focused on the predictive purpose of the modeling, and a firm grasp of the sources and magnitudes of model input (i.e. parameter) and conceptual model uncertainty. Decisions based on the associated numerical model can only then take full advantage of the decision-support simulation tools discussed herein.

A final comment: When asked about his upcoming conference presentation in the early 1980s, de Marsily mentioned that his talk was on applying geostatistics in hydrogeology, which at the time was virtually unknown in the US. He went on to say that he felt it was important to explain to the audience how powerful geostatistical tools are to hydrogeologists and that it “was never too late to learn new things”. He was right on both points. Countless hydrogeologists have been positively impacted by the “new things” Ghislain de Marsily taught us and by his example of leading with humility, style and grace. Thank you, Ghislain.

## Conflicts of interest

Authors have no conflict of interest to declare.

## References

- Ahmed, S. and De Marsily, G. (1987). Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity. *Water Resour. Res.*, 23(9), 1717–1737.
- Alcolea, A., Carrera, J., and Medina, A. (2006). Pilot points method incorporating prior information for solving the groundwater flow inverse problem. *Adv. water Resour.*, 29(11), 1678–1689.
- Aster, R. C., Borchers, B., and Thurber, C. H. (2018). *Parameter Estimation and Inverse Problems*. Elsevier, Amsterdam.
- Capilla, J., Gómez-Hernández, J. J., and Sahuquillo, A. (1997). Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data 2. Demonstration on a synthetic aquifer. *J. Hydrol.*, 203(1–4), 175–188.
- Carrera, J. (1994). Invert-4. a fortran program for solving the groundwater flow inverse problem. User's guide. Technical report, CIMNE, Barcelona, Spain. 160 pp. þ appendices.
- Carrera, J. and Glorioso, L. (1991). On geostatistical formulations of the groundwater flow inverse problem. *Adv. Water Resour.*, 14(5), 273–283.
- Certes, C. and De Marsily, G. (1991). Application of the pilot point method to the identification of aquifer transmissivities. *Adv. Water Resour.*, 14(5), 284–300.
- Chavent, G. (1974). Estimation de parametres distribués dans des equations aux derivees partielles. In *Computing Methods in Applied Sciences and Engineering Part 2*, pages 361–390. Springer, Berlin.
- Chen, Y. and Oliver, D. S. (2013). Levenberg–marquardt forms of the iterative ensemble smoother for efficient history matching and uncertainty quantification. *Comput. Geosci.*, 17(4), 689–703.
- Chilès, J.-P. and Desassis, N. (2018). Fifty years of kriging. In *Handbook of Mathematical Geosciences*, pages 589–612. Springer, Cham.
- Christensen, S. and Doherty, J. (2008). Predictive error dependencies when using pilot points and singular value decomposition in groundwater model calibration. *Adv. Water Resour.*, 31(4), 674–700.
- Dausman, A. M., Doherty, J., Langevin, C. D., and Sukop, M. C. (2010). Quantifying data worth toward reducing predictive uncertainty. *Groundwater*, 48(5), 729–740.
- De Marsily, G. (1978). *De l'identification des systèmes hydrogéologiques*. PhD thesis, Pierre and Marie Curie Univ., Paris.
- De Marsily, G., Delhomme, J., Coudrain-Ribstein, A., and Lavenue, A. (2000). Four decades of inverse problems in hydrogeology. In Zhang, D. and Winter, C. L., editors, *Theory, Modeling and Field Investigation in Hydrogeology: A Special Volume in Honor of Shlomo P. Neuman's 60th Birthday*, Special Paper 348, pages 1–17. Geological Society of America, Boulder, Colorado.
- De Marsily, G., Lavedan, G., Boucher, M., and Fasanino, G. (1984). Interpretation of interference tests in a well field using geostatistical techniques to fit the permeability distribution in a reservoir model. In Verly, G., David, M., Journel, A. G.,

- and Marecha, A., editors, *Geostatistics for Natural Resources Characterization 2nd NATO Advanced Study Institute, South Lake Tahoe, CA, September 6–17, 1987*, pages 831–849. D. Reidel, Hingham, MA. Pt. 2.
- Deutsch, C. and Journel, A. (1992). *GSLIB: Geostatistical Software Library and User's Guide*. Oxford University Press, New York, NY.
- Doherty, J. (2003). Ground water model calibration using pilot points and regularization. *Groundwater*, 41(2), 170–177.
- Doherty, J. E. (2015). *PEST and its Utility Support Software, Theory*. Watermark Numerical Publishing, Brisbane.
- Doherty, J. E., Fienen, M. N., and Tonkin, M. J. (2010). Approaches to highly parameterized inversion: Pilot-point theory, guidelines, and research directions. U.S. Geological Survey Scientific Investigations Report 2010-5168, <https://doi.org/10.3133/sir20105168>.
- Emsellem, Y. and De Marsily, G. (1971). An automatic solution for the inverse problem. *Water Resour. Res.*, 7(5), 1264–1283.
- Fienen, M. N., Muffels, C. T., and Hunt, R. J. (2009). On constraining pilot point calibration with regularization in pest. *Groundwater*, 47(6), 835–844.
- Gómez-Hernández, J. J., Sahuquillo, A., and Capilla, J. (1997). Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data—i. Theory. *J. Hydrol.*, 203(1–4), 162–174.
- Grindrod, P. and Impey, M. (1991). Fractal field simulations of tracer migration within the wipp culebra dolomite. *Rep. IM2856-1, Intera Information Technologies*, 2, 62.
- Gutjahr, A., Bullard, B., Hatch, S., and Hughson, L. (1994). Joint conditional simulations and the spectral approach for flow modeling. *Stoch. Hydrol. Hydraul.*, 8(1), 79–108.
- Herckenrath, D., Langevin, C. D., and Doherty, J. (2011). Predictive uncertainty analysis of a saltwater intrusion model using null-space Monte Carlo. *Water Resour. Res.*, 47(5), 1–16.
- James, S. C., Doherty, J. E., and Eddebbarh, A.-A. (2009). Practical postcalibration uncertainty analysis: Yucca mountain, Nevada. *Groundwater*, 47(6), 851–869.
- Jiménez, S., Mariethoz, G., Brauchler, R., and Bayer, P. (2016). Smart pilot points using reversible-jump markov-chain monte carlo. *Water Resour. Res.*, 52(5), 3966–3983.
- Keller, J., Franssen, H.-J. H., and Nowak, W. (2021). Investigating the pilot point ensemble kalman filter for geostatistical inversion and data assimilation. *Adv. Water Resour.*, 155, article no. 104010.
- Khambhammettu, P., Renard, P., and Doherty, J. (2020). The traveling pilot point method. a novel approach to parameterize the inverse problem for categorical fields. *Adv. Water Resour.*, 138, article no. 103556.
- Kitanidis, P. K. and Lane, R. W. (1985). Maximum likelihood parameter estimation of hydrologic spatial processes by the gauss-newton method. *J. Hydrol.*, 79(1–2), 53–71.
- Kleinecke, D. (1971). Use of linear programming for estimating geohydrologic parameters of groundwater basins. *Water Resour. Res.*, 7(2), 367–374.
- Knowing, M. J. and Werner, A. D. (2016). Estimability of recharge through groundwater model calibration: Insights from a field-scale steady-state example. *J. Hydrol.*, 540, 973–987.
- Knowing, M. J. and Werner, A. D. (2017). Transient recharge estimability through field-scale groundwater model calibration. *Groundwater*, 55(6), 827–840.
- Lavenue, A. M. (1998). *A new pilot point inverse method in hydrogeology: Generating an ensemble of conditionally-simulated transmissivity fields*. PhD thesis, School of Mines, Paris.
- Lavenue, A. M. and Pickens, J. F. (1992). Application of a coupled adjoint sensitivity and kriging approach to calibrate a groundwater flow model. *Water Resour. Res.*, 28(6), 1543–1569.
- Lavenue, A. M., Ramarao, B. S., De Marsily, G., and Marietta, M. G. (1995). Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields: 2. Application. *Water Resour. Res.*, 31(3), 495–516.
- Lavenue, M. and De Marsily, G. (2001). Three-dimensional interference test interpretation in a fractured aquifer using the pilot point inverse method. *Water Resour. Res.*, 37(11), 2659–2675.
- Law, J. (1944). A statistical approach to the interstitial heterogeneity of sand reservoirs. *Trans. AIME*, 155(01), 202–222.
- Li, L., Srinivasan, S., Zhou, H., and Gómez-Hernández, J. J. (2013). A pilot point guided pattern matching approach to integrate dynamic data



- into geological modeling. *Adv. Water Resour.*, 62, 125–138.
- Liu, J.-b., Jiang, S.-m., Zhou, N.-q., Cai, Y., Cheng, L., and Wang, Z.-y. (2021). Groundwater contaminant source identification based on QS-ilues. *J. Groundw. Sci. Eng.*, 9(1), 73–82.
- Ma, W. and Jafarpour, B. (2018). Pilot points method for conditioning multiple-point statistical facies simulation on flow data. *Adv. Water Resour.*, 115, 219–233.
- Mariethoz, G. and Caers, J. (2014). *Multiple-Point Geostatistics: Stochastic Modeling with Training Images*. John Wiley & Sons, Hoboken, NJ.
- Matheron, G. (1963). *Traité de géostatistique appliquée tome ii: Le krigeage, Memoires, Bureau du Recherches Geologiques et Minières*, volume 24, issue 14. Éditions Technip, Paris.
- Matheron, G. and Blondel, F. (1962). *Traité de géostatistique appliquée, tome i*. Editions Technip, Paris. Á. Martínez-del-Pozo et al.
- McKenna, S. A., Akhriev, A., Ciaurri, D. E., and Zhuk, S. (2020). Efficient uncertainty quantification of reservoir properties for parameter estimation and production forecasting. *Math. Geosci.*, 52, 233–251.
- Moore, C. and Doherty, J. (2005). Role of the calibration process in reducing model predictive error. *Water Resour. Res.*, 41(5), 1–14.
- Oliver, D. S., Reynolds, A. C., and Liu, N. (2008). *Inverse Theory for Petroleum Reservoir Characterization and History Matching*. Cambridge University Press, Cambridge.
- Ramarao, B. S., Lavenue, A. M., De Marsily, G., and Marietta, M. G. (1995). Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields: 1. Theory and computational experiments. *Water Resour. Res.*, 31(3), 475–493.
- Riva, M., Guadagnini, A., De Gaspari, F., and Alcolea, A. (2010). Exact sensitivity matrix and influence of the number of pilot points in the geostatistical inversion of moment equations of groundwater flow. *Water Resour. Res.*, 46, 1–15.
- Rubin, Y., Mavko, G., and Harris, J. (1992). Mapping permeability in heterogeneous aquifers using hydrologic and seismic data. *Water Resour. Res.*, 28(7), 1809–1816.
- Sahuquillo, A., Capilla, J., Gómez-Hernández, J., and Andreu, J. (1992). Conditional simulation of transmissivity fields honoring piezometric data. *Hydraul. Eng. Softw. IV, Fluid Flow Model.*, 2, 201–214.
- Sánchez-Vila, X., Carrera, J., and Girardi, J. P. (1996). Scale effects in transmissivity. *J. Hydrol.*, 183(1–2), 1–22.
- Tikhonov, A. N. (1963). Regularization of incorrectly posed problems. *Sov. Math. Dokl.*, 4, 1624–1627.
- Tikhonov, A. N. and Arsenin, V. Y. (1977). *Solutions of Ill-Posed Problems*. V.H. Winston & Sons, Washington, DC: John Wiley & Sons, New York, NY.
- Tonkin, M. and Doherty, J. (2009). Calibration-constrained monte carlo analysis of highly parameterized models using subspace techniques. *Water Resour. Res.*, 45(12), 1–17.
- Tonkin, M. J. and Doherty, J. (2005). A hybrid regularized inversion methodology for highly parameterized environmental models. *Water Resour. Res.*, 41(10), 1–16.
- White, J. T. (2018). A model-independent iterative ensemble smoother for efficient history-matching and uncertainty quantification in very high dimensions. *Environ. Model. Softw.*, 109, 191–201.
- White, J. T., Foster, L. K., Fienen, M. N., Knowling, M. J., Hemmings, B., and Winterle, J. R. (2020a). Toward reproducible environmental modeling for decision support: A worked example. *Front. Earth Sci.*, 8, article no. 50.
- White, J. T., Hunt, R. J., Fienen, M. N., and Doherty, J. E. (2020b). Approaches to Highly Parameterized Inversion: PEST++ Version 5, a Software Suite for Parameter Estimation, Uncertainty Analysis, Management Optimization and Sensitivity Analysis. U.S. Geological Survey Techniques and Methods 7C26, 52 p.
- Zimmerman, D., De Marsily, G., Gotway, C. A., Marietta, M. G., Axness, C. L., Beauheim, R. L., Bras, R. L., Carrera, J., Dagan, G., Davies, P. B., et al. (1998). A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow. *Water Resour. Res.*, 34(6), 1373–1413.





Research article

Geo-hydrological Data & Models

# Comparison of three recent discrete stochastic inversion methods and influence of the prior choice

Przemysław Juda<sup>✉,a</sup>, Julien Straubhaar<sup>✉,a</sup> and Philippe Renard<sup>✉,\*,b,c</sup>

<sup>a</sup> Stochastic Hydrogeology and Geostatistics Group, Centre for Hydrogeology and Geothermics, University of Neuchâtel, Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland

URL: <http://www.unine.ch/philippe.renard>

<sup>b</sup> Department of Geosciences, University of Oslo, Oslo, Norway

E-mails: [przemyslaw.juda@unine.ch](mailto:przemyslaw.juda@unine.ch) (P. Juda), [julien.straubhaar@unine.ch](mailto:julien.straubhaar@unine.ch) (J. Straubhaar), [philippe.renard@unine.ch](mailto:philippe.renard@unine.ch) (P. Renard)

**Abstract.** Groundwater flow depends on subsurface heterogeneity, which often calls for categorical fields to represent different geological facies. The knowledge about subsurface is however limited and often provided indirectly by state variables, such as hydraulic heads of contaminant concentrations. In such cases, solving a categorical inverse problem is an important step in subsurface modeling. In this work, we present and compare three recent inverse frameworks: Posterior Population Expansion (PoPEx), Ensemble Smoother with Multiple Data Assimilation (ESMDA), and DREAM-ZS (a Markov chain Monte Carlo sampler). PoPEx and ESMDA are used with Multiple-point statistics (MPS) as geostatistical engines, and DREAM-ZS is used with a Wasserstein generative adversarial network (WGAN). The three inversion methods are tested on a synthetic example of a pumping test in a fluvial channelized aquifer. Moreover, the inverse problem is solved three times with each method, each time using a different training image to check the performance of the methods with different geological priors. To assess the quality of the results, we propose a framework based on continuous ranked probability score (CRPS), which compares single true values with predictive distributions. All methods performed well when using the training image used to create the reference, but their performances were degraded with the alternative training images. PoPEx produced the least geological artifacts but presented a rather slow convergence. ESMDA showed initially a very fast convergence which reaches a plateau, contrary to the remaining methods. DREAM-ZS was overly confident in placing some incorrect geological features but outperformed the other methods in terms of convergence.

**Keywords.** Stochastic inversion, Multiple-point statistics, Monte Carlo sampling, Posterior Population Expansion, Ensemble smoother, Groundwater flow, Scoring rules.

*Manuscript received 7 April 2022, revised 2 July 2022, accepted 12 September 2022.*

## 1. Introduction

Roughly twenty years ago, Ghislain de Marsily gave an overview of four decades of inverse problems in hydrogeology [de Marsily et al., 1999]. What is striking when reading this review is that all the methods

are aimed at inferring a continuous field of parameter values. The review highlights the evolution of the ideas in this domain and how the initial deterministic and direct methods were progressively replaced by indirect and geostatistical methods. In the epilogue of that paper, Ghislain de Marsily indicates where the research is heading with the emerging category of approaches which consists of generating images

\* Corresponding author.

of the geologic reality. Indeed, in the last 20 years, a considerable effort has been devoted to developing novel geostatistical simulation methods able to deal with categorical fields representing the spatial distribution of rock types or geological formations [de Marsily *et al.*, 2005]. In a categorical inverse problem, the aim is to identify for every location the rock type or lithology among a discrete and fixed number of possibilities.

Solving the inverse problem in the categorical case while respecting prior geological knowledge has raised new challenges and difficulties [Oliver and Chen, 2011, Linde *et al.*, 2015]. In particular, standard optimization techniques based on a gradient or adjoint-based approaches which were used successfully in the continuous case [de Marsily *et al.*, 1984] cannot be directly applied in the categorical case since the concept of “derivative” has no meaning in these situations because the possible changes in parameters are discrete. One had either to find a latent representation of the geology using an underlying continuous representation or to rely on Monte Carlo techniques that are more robust but less efficient. Many of these challenges are still open and the groundwater modeling community is actively pursuing this research. It is however not always clear what are the advantages and limitations of the different approaches. Previous intercomparison exercises [Zimmerman *et al.*, 1998, Hendricks Franssen *et al.*, 2009] did not consider the case of discrete fields with geological prior knowledge.

The aim of this paper is therefore to provide a comparison of three recent inversion methods dedicated to the categorical inverse problem. All those methods are flexible. They are based on different representations of geology that all account for a conceptual prior model and could be applied to different types of geology. They all tackle the inversion problem using a different approach and they have not yet been compared for the same inverse problem to our knowledge.

The first technique is based on a multiple-point statistics (MPS) approach to represent the categorical field. The prior knowledge is given to the algorithm by providing a training image [Journel and Zhang, 2006] which can be seen as a training data set or geological analog representing the type of patterns that are expected to occur in the region of interest. The MPS approach respects high-order statistics

and allows flexible control of heterogeneities. MPS algorithms have been extensively used in inversion frameworks. Early examples include for example the probability perturbation method [Caers and Hoffman, 2006], the blocking moving window algorithm [Alcolea and Renard, 2010, Hansen *et al.*, 2012], or the iterative spatial resampling [Mariethoz *et al.*, 2010]. These methods iteratively update MPS realizations by imposing hard or soft conditioning data either in an optimization or Monte Carlo Markov chains perspective. Here, we will use the Posterior Population Expansion (PoPEX) algorithm [Jäggli *et al.*, 2017, 2018]. It is an adaptive importance sampling (AIS) scheme that also uses hard conditioning data to iteratively expand an ensemble of models. PoPEX learns the relation between the state variables and categorical parameter values using conditional probabilities and employs this knowledge to generate new realizations that are progressively more likely to fit the data. An important feature of PoPEX is that it is highly parallelizable. Note that we present a modification in the PoPEX approach in this paper and introduce the notion of tempered weights.

The second technique uses a slightly different MPS representation of the heterogeneity allowing to use a data assimilation method [Evensen, 2009] for the parameter identification step. The assimilation approaches are known to be very efficient to infer multi-Gaussian fields from state variables. They were extended to non-Gaussian and categorical examples [Zhou *et al.*, 2014, Oliver and Chen, 2018, Kang *et al.*, 2019] but always require a continuous representation of the geology. A recent development in the MPS technology is a multiresolution algorithm [Straubhaar *et al.*, 2020]. The fine-scale geological and categorical fields are upscaled on lower-resolution grids using Gaussian pyramids. An MPS simulation can be conditioned by the values of the Gaussian pyramids, and this allowed Lam *et al.* [2020] to apply the ensemble smoother with multiple data assimilation (ESMDA) [Emerick and Reynolds, 2013] directly to MPS realizations of categorical variables. In this approach, the relation between the underlying continuous variables and the state variables are estimated using covariances.

Finally, the third technique uses a generative adversarial network (GAN) to represent geology. GAN became very popular in recent years [Goodfellow *et al.*, 2014] due to their ability to generate highly re-

alistic images provided a sufficiently large training dataset is available. One of their main interest is their flexibility and their capacity to learn the relation between a relatively low dimensional latent space representation and the final images. The representation of parameters in the latent space (which is often Gaussian) is convenient for Markov chain Monte Carlo inverse algorithms. Here, we used the DREAM-ZS algorithm [Laloy and Vrugt, 2012] combined with spatial GAN following the very successful work of Laloy *et al.* [2018].

In this paper, we first introduce the three different techniques and how they were implemented. Indeed, to ensure a fair comparison, we implemented the three methods using similar tools. The corresponding codes are available online.<sup>1</sup> To compare the performances of the three methods, it was important to have access to the reference, and therefore we designed a synthetic pumping test experiment. A geological model was generated and we simulated the pumping test. The data are then used for identifying the geology and the corresponding uncertainty with the three techniques: PoPEx, ESM DA, and DREAM-ZS. Since in practice, it is difficult to identify the proper prior model for the geology (i.e. the right training image), we also tested the inverse methods with the incorrect priors. This allows us to compare not only the performances of the inverse method in the ideal case where the prior is correct but also to check the robustness of the three techniques to incorrect priors.

## 2. Inversion algorithms

In this section, we provide a description of the three stochastic inversion algorithms and the tools used to generate the discrete random fields. Let us first recall the main notions of the stochastic formulation of the inverse problem. We will use these notations to present the three algorithms. The observed data are stored in a vector of real values  $\mathbf{d}^{\text{obs}} \in \mathbb{R}^N$ , and  $N$  is the number of observed data points. Let us consider a model manifold  $\mathfrak{M}$ . Any model  $\mathbf{m} \in \mathfrak{M}$  is supposed to describe fully the physical system. In other words, it provides sufficient input for the forward solver to simulate the data. The forward solver is an operator

$\mathbf{g}: \mathfrak{M} \rightarrow \mathbb{R}^N$ , mapping from the model manifold  $\mathfrak{M}$  to the data space. For example, the observed data can be a time series of hydraulic heads at different locations, or a time series of tracer concentrations. The model space can describe a field of geological facies in the subsurface (discrete model space) or a field of hydraulic properties (continuous model space). The forward operator can be a groundwater flow solver or transport solver. Usually, it solves a set of partial differential equations. The output of the forward solver is deterministic: given the same model, the simulated data are uniquely defined.

The probabilistic solution to the inverse problem is given by Tarantola [2005]:

$$\sigma(\mathbf{m}) = c\rho(\mathbf{m})L(\mathbf{m}; \mathbf{d}^{\text{obs}}), \quad (1)$$

with  $\sigma(\mathbf{m})$  the posterior probability distribution,  $c$  some normalization constant,  $\rho(\mathbf{m})$  the prior probability distribution, and  $L(\mathbf{m}; \mathbf{d}^{\text{obs}})$  the likelihood function. The likelihood function  $L(\mathbf{m}; \mathbf{d}^{\text{obs}})$  describes how likely is the model given the observations (it measures the mismatch between the simulated and the observed data) and it depends on the problem at hand (we indicated that the function  $L$  uses the observed data with  $L(\mathbf{m}; \mathbf{d}^{\text{obs}})$ , but we will write  $L(\mathbf{m})$  for brevity). The prior probability distribution  $\rho(\mathbf{m})$  contains knowledge that is independent of measured data. It is a domain-specific (expert) knowledge about model parameters; for example, constraining models to be generated by a specific geostatistical method. In practice, the normalization constant  $c$  does not play a role, as the model manifold is approximated using a finite-dimensional space, and solutions can be self-normalized.

The characterization of the posterior distribution  $\sigma(\mathbf{m})$  is the goal of the inversion algorithms, and any useful property can be written as a prediction in the following manner:

$$\mu = \int_{\mathfrak{M}} \sigma(\mathbf{m})f(\mathbf{m}) \, d\mathbf{m}, \quad (2)$$

where  $\mu$  represents the prediction (expected value) of the quantity of interest which is obtained using function  $f(\mathbf{m})$ . Typically, Monte Carlo methods aim to sample the posterior, and then use a subset  $\mathcal{M} \subset \mathfrak{M}$  to approximate the integral using a sum.

<sup>1</sup><https://github.com/randlab/inversion-comparison.git>.

### 2.1. PoPEx + MPS

The Posterior Population Expansion (PoPEx) algorithm [Jäggli *et al.*, 2017, 2018] is an adaptive importance sampling (AIS) technique designed for solving inverse problems in the context of categorical geostatistical fields. In this work, we use the parallelized implementation of PoPEx based on asynchronous worker processes [Jäggli *et al.*, 2018], with a modification for computing the weights for generating predictions. We do not use *corrected weights* as described by Jäggli *et al.* [2018], but instead, *tempered weights*, based on tempered likelihood, which is explained below. The motivation to use *tempered weights* instead of *corrected weights* is explained by the fact that Jäggli *et al.* [2018] and Juda and Renard [2021] had to use a subset of tracer test data to allow convergence of PoPEx. For example, Juda and Renard [2021] used 6 out of a total of 276 data points in the tracer concentration curve. This approach had to be used to increase the number of effective weights for prediction, otherwise, too few models were retained for the prediction, and uncertainty was not very well represented. While reducing the dimensionality of data in this way, might be an effective ad hoc solution, it is not generic and arbitrary. Tempered weights aim to solve this issue more generally.

#### 2.1.1. Tempered weights

Tempered weights have been inspired by other solutions to the problem of the *peakedness* of the likelihood function. Laloy *et al.* [2018] presented a case of 3-D Transient Hydraulic Tomography, where 1568 data points were used in the inversion using the DREAM-ZS algorithm [Laloy and Vrugt, 2012]. The inversion was stopped before reaching the convergence criterion. In that study, tempering of the likelihood function was implemented but limited to the burn-in. It consisted in using an inflated variance term in the likelihood function. A similar technique to tempering is also used in the context of data assimilation. Lam *et al.* [2020] used ensemble smoother with multiple data assimilation (ESMDA) for discrete inversion, where geostatistical simulation uses pyramids. ESMDA applies Kalman update repeatedly to assimilate data but introduces a factor  $\alpha$  for reducing the correction term, as the same data is assimilated multiple times [Emerick and Reynolds, 2012]. It corresponds to reducing the confidence given to

the (noisy) data at every iteration of the data assimilation.

The adaptive importance sampling provides a convenient formula, the self normalized estimator  $\hat{\mu}_{\text{sn}}$ , to approximate integrals like (2) using the following sum [Jäggli *et al.*, 2018]:

$$\hat{\mu}_{\text{sn}} = \sum_{j=1}^k f(\mathbf{m}^j) \tilde{w}^j. \quad (3)$$

The  $\tilde{w}^j$  are normalized weights:  $\tilde{w}^j = w^j / \sum_{i=1}^k w^i$ , with  $k$  representing the total number of generated models (iterations). The superscript is not used as exponent, instead, it is used for indexing iterations, and we keep this notation for consistency with the reference PoPEx paper [Jäggli *et al.*, 2018]. In the AIS framework, the weights  $w^k$  are given by:

$$w^k = \frac{\sigma(\mathbf{m}^k)}{\phi^k(\mathbf{m}^k)} = c \frac{\rho(\mathbf{m}^k)}{\phi^k(\mathbf{m}^k)} L(\mathbf{m}^k), \quad (4)$$

with a constant  $c$  (that can be ignored later due to self-normalization),  $L(\mathbf{m})$  the likelihood function, and  $\rho(\mathbf{m})$  the prior measure.  $\phi^k$  is a sampling distribution that is updated at every iteration  $k$ . The main idea of adaptive importance sampling is to update  $\phi^k$  in a way that it resembles  $\sigma$  but has heavier tails.

To resolve the problem of few significant weights, we suggest an approach based on *tempered* likelihood function. It is similar to using higher error variance in the likelihood formula. The tempering factor is adapted (optimized) based on the desired number of significant models.

Let us define the family of *tempered* likelihood function  $L_t(\mathbf{m}; f_\sigma)$ :

$$L_t(\mathbf{m}; f_\sigma) = \exp \left[ \frac{1}{f_\sigma^2} \log(L(\mathbf{m})) \right] \quad (5)$$

with the tempering factor  $f_\sigma \geq 1$ . If  $f_\sigma = 1$  we have:  $L_t(\mathbf{m}^i; 1) \equiv L(\mathbf{m}^i)$ , and the tempered likelihood becomes equivalent to the standard likelihood function. In this sense, the tempered likelihood is a generalization of the correct likelihood function for the problem at hand. The tempering factor reduces confidence in the data, it can be interpreted as a factor inflating the variance of the measurement error. Often, the standard likelihood considers the data points as non-correlated, and the tempering factor then makes sense as compensation for ignoring correlation between data points, which is difficult to evaluate.

If we use the soft likelihood in (4), we obtain a parametric formula for tempered weights:

$$w_t^k(f_\sigma) = c \frac{\rho(\mathbf{m}^k)}{\phi^k(\mathbf{m}^k)} L_t(\mathbf{m}^k; f_\sigma). \quad (6)$$

Finally, it leads to the tempered formula for the self-normalized estimator:

$$\hat{\mu}_{\text{sn,t}}(f_\sigma) = \sum_{j=1}^k f(\mathbf{m}^j) \tilde{w}_t^j(f_\sigma). \quad (7)$$

### 2.1.2. Optimal tempering factor

The tempering factor can be chosen arbitrarily, for example  $f_\sigma = \sqrt{N}$  would correspond to taking the average of log-likelihood over  $N$  observation points of the mismatch. Instead of fixing a value for  $f_\sigma$ , we propose a method to adaptively choose optimal  $f_\sigma$ . It is inspired by the formulation of *corrected* PoPEX weights.

Let us consider a set  $\mathcal{W}^k(f_\sigma)$  of  $k$  tempered weights with parameter  $f_\sigma$ :

$$\mathcal{W}^k(f_\sigma) = \{w_t^1(f_\sigma), \dots, w_t^k(f_\sigma)\}. \quad (8)$$

The effective sample size for the set  $\mathcal{W}^k(f_\sigma)$  is given by:

$$n_e(\mathcal{W}^k(f_\sigma)) = \frac{(\sum_{i=1}^k w_t^i(f_\sigma))^2}{\sum_{i=1}^k (w_t^i(f_\sigma))^2}. \quad (9)$$

Suppose that the target value of the minimal number of effective weights  $\theta$  is chosen by the user, who also specifies the value of  $f_{\text{max}}$  which will be the max bound for  $f_\sigma$ . We will define  $f_\sigma$  as optimal if it is such that the number of effective weights  $n_e$  equals at least  $\theta$  and  $f_\sigma \in [1, f_{\text{max}}]$  is as small as possible. This can be translated into the following optimization problem:

$$f_{\text{opt}} = \underset{f_\sigma \in [1, f_{\text{max}}]}{\operatorname{argmin}} (n_e(\mathcal{W}^k(f_\sigma)) - \theta)^2, \quad (10)$$

where  $f_{\text{opt}}$  is the optimal tempering factor. The set of the optimal tempered weights is given by  $\mathcal{W}^k(f_{\text{opt}})$ , and after normalization, they can be used in (7) to get the desired estimator.

The tempering framework can be summarized in the form of an algorithm. It takes as input:  $\theta$ —the target number of effective weights;  $f_{\text{max}}$ —the max bound for the tempering factor. The algorithm is as follows:

OPTIMAL-TEMPERED-WEIGHTS ( $\theta, f_{\text{max}}$ )

```

1 Minimize  $(n_e(\mathcal{W}^k(f_\sigma)) - \theta)^2$  subject to  $f_\sigma \in [1, f_{\text{max}}]$ 
2  $f_{\text{opt}} = \operatorname{argument}$  of the minimum
3 Compute  $\mathcal{W}^k(f_{\text{opt}})$ 
4 for  $j = 1, \dots, k$ 
5      $\tilde{w}_t^j = w_t^j / \sum_{i=1}^k w_t^i$ 
6 return  $\{\tilde{w}_t^1, \dots, \tilde{w}_t^k\}$ 

```

Once the models are generated (PoPEX stops after a number of steps predefined by the user), PoPEX uses weights for generating predictions. While it is possible to use the tempered likelihood instead of the exact likelihood during PoPEX sampling, we do not use this approach in this study. PoPEX is run with the correct (exact) likelihood for the problem at hand, and the tempered likelihood is only applied for computing predictive weights.

In this study and in previous ones [Jägglı et al., 2017, 2018, Dagan et al., 2020], PoPEX was coupled with the Direct Sampling (DS) MPS algorithm to generate the categorical fields. More precisely, we use the DeeSse implementation with multi-resolution features [Straubhaar et al., 2020]. The multi-resolution capability (Gaussian pyramids) is a technique allowing for improved reproduction of patterns at different scales.

### 2.2. ESMDA + DS pyramid

The second inversion method that we will compare is the one proposed by Lam et al. [2020]. It is based on the ensemble smoother with multiple data assimilation (ESMDA) coupled with DS with Gaussian pyramids. ESMDA [Emerick and Reynolds, 2013] runs for a predefined number of steps  $N_a$  (parameter given by the user, also known as number of data assimilations), and at each iteration  $k \in \{1, 2, \dots, N_a\}$  the ensemble  $N_e$  of models  $\{\mathbf{m}_1^k, \mathbf{m}_2^k, \dots, \mathbf{m}_{N_e}^k\}$  is updated to  $\{\mathbf{m}_1^{k+1}, \mathbf{m}_2^{k+1}, \dots, \mathbf{m}_{N_e}^{k+1}\}$ . We use subscript here for the model index and superscript for the iteration index. The index 1 corresponds to the initial (prior) ensemble, and the ensemble after  $N_a$  data assimilations will have index  $N_a + 1$ . We based the algorithmic implementation of the method on the paper by Emerick [2016].

Contrary to PoPEX, in the ESMDA framework, a model is a vector of real values (not discrete) and the described method concerns matching data represented by continuous values. The main novelty

of the approach proposed by Lam et al. [2020] is the way of conditioning categorical simulations with continuous variables. Therefore, the ESM DA procedure is a standard one, but the data that is assimilated is used to condition categorical simulations. In this subsection, we will review briefly how it is done. We need two ingredients: a procedure for generating an initial ensemble of models, which is a vector of continuous parameters, and a procedure to generate categorical models based on such a vector.

### 2.2.1. Coupling DS and ESM DA

The ensemble of models  $\mathbf{M}^1 = [\mathbf{m}_1^1, \mathbf{m}_2^1, \dots, \mathbf{m}_{N_e}^1]$  is generated using the following steps.

We use the multi-resolution option of the DeeSse code [Straubhaar et al., 2020] to generate unconditional realizations. The fine-scale realizations are categorical but the DeeSse simulation algorithm starts by generating a pyramid of lower-resolution continuous images over the same grid. The low-resolution continuous images guide the simulation of the higher-resolution categorical images Lam et al. [2020]. The link between the continuous and categorical variables is established on the training image using Gaussian kernels to blur and represent the field at a lower resolution. At the coarse resolution, a fraction  $f$  of the total number of cells is sampled to obtain an ensemble of pyramid values (now continuous) at fixed locations  $\{\mathbf{p}_1^1, \mathbf{p}_2^1, \dots, \mathbf{p}_{N_e}^1\}$ , with  $\mathbf{p}_i^k \in \mathbb{R}^{N_m}$ , where  $k$  is the iteration index,  $i$  the ensemble member index, and  $N_m$  the number of conditioning locations.  $\mathbf{p}_i^k[j]$  represents a Gaussian pyramid value at a location with index  $j$ . While it would be possible to use directly  $\mathbf{p}_i^k$  vectors in the ESM DA procedure, it is not a good idea, because these parameter distributions are not necessarily Gaussian and ESM DA performance will be hindered. Therefore, Lam et al. [2020] suggest using normal score transform, as proposed in the study of Zhou et al. [2011].

The normal score transfer function is constructed for each parameter in the vector  $\mathbf{p}_i^1[j]$  and is kept fixed for the entire data assimilation process. Let  $F_j$  for all  $j \in \{1, 2, \dots, N_m\}$  be the cumulative distribution function (CDF) deduced from the ensemble  $\{\mathbf{p}_1^1[j], \mathbf{p}_2^1[j], \dots, \mathbf{p}_{N_e}^1[j]\}$ . For each pyramid location,  $j$  corresponding  $F_j$  is computed with its inverse  $F_j^{-1}$  and they are stored. Now the direct normal score

transform is defined:

$$\Phi_i^{\text{direct}}(x) = G^{-1}(F_j(x)), \quad (11)$$

where  $G^{-1}$  is the inverse of normal CDF. The normal score back transform is given by:

$$\Phi_i^{\text{back}}(x) = F_j^{-1}(G(x)), \quad (12)$$

where  $F_j^{-1}$  is inverse of the pyramid CDF, and  $G$  stands for normal CDF.

Finally, the vector  $\mathbf{m}_i^1 = [\mathbf{m}_i^1[1], \mathbf{m}_i^1[2], \dots, \mathbf{m}_i^1[N_m]]^\top$  is obtained from the initial ensemble:

$$\mathbf{m}_i^1[j] = \Phi_j^{\text{direct}}(\mathbf{p}_i^1[j]) \quad (13)$$

for all  $j \in \{1, 2, \dots, N_m\}$  and  $i \in \{1, 2, \dots, N_e\}$ . At each iteration  $k$ , to transfer the parameter vector  $\mathbf{m}_i^k$  into pyramid conditioning data we employ the back transform:

$$\mathbf{p}_i^k[j] = \Phi_j^{\text{back}}(\mathbf{m}_i^k[j]), \quad (14)$$

for  $j \in \{1, 2, \dots, N_m\}$ . The set of pyramids can now be used in the DeeSse implementation to obtain the realizations by conditioning the simulations at the coarse resolution. It is important to note that the same simulation seed must be used at each iteration for the given realization index  $i$ . It ensures that the realizations are gradually improved by the inversion. Otherwise, the stochastic nature of the MPS simulation would make it impossible. Before every forward model call, the model parameters are converted to pyramid values and the subsequent realization is generated. The details of the multi-resolution Direct Sampling implementation are given in Straubhaar et al. [2020], and the details about the conditioning of Gaussian pyramids are in Lam et al. [2020].

## 2.3. DREAM-ZS + WGAN

The third inversion method is DREAM-ZS used with a Wasserstein Generative Adversarial Network (WGAN). The approach was initially proposed by Laloy et al. [2018], we use it with only a slight modification: we employ a Wasserstein GAN instead of a Spatial GAN.

### 2.3.1. DREAM-ZS

DREAM-ZS is a modified Metropolis sampler, sampling multiple chains which exchange information using an archive  $\mathcal{Z}$  of past models [Laloy et al., 2018, 2017]. Metropolis samplers generate proposals and accept them if their likelihood is higher, or with



a probability if it is lower. In a standard Metropolis sampler, chains would not communicate with each other, which makes it easily parallelizable, but requires removing outlier trajectories; this results in a slower convergence. DREAM-ZS provides a way to allow efficient parallelization and communication between the chains; hence avoiding the necessity of removing outlier chains. Usually, the sampler is run unless a convergence criterion is satisfied [for example Gelman and Rubin, 1992] but in practical cases with a large amount of data, the convergence might not be achieved in a reasonable number of iterations [Laloy *et al.*, 2018]. Therefore, we simply run here  $N_c$  chains during a predefined number of iterations  $T$  and use the two last recorded samples from each chain to form the posterior. Samples are recorded every  $K$  iterations. Our implementation uses a Wasserstein GAN (WGAN), instead of a spatial GAN (SGAN) as it was suggested by Laloy *et al.* [2018], because WGANs are known to be more stable, and easier to train for different training data sets. More details on our WGAN setup are given in the next subsection. The details of our implementation are based on the MT-DREAM-ZS paper Laloy and Vrugt [2012] but we do not use the multiple tries (MT) technique. Our implementation essentially corresponds to MT-DREAM-ZS with one trial. The algorithm for computation of crossover values is based on the work of Vrugt *et al.* [2009].

The parameter space is the latent space of the GAN,  $\mathbf{x}$  is the parameter vector of length  $d$ :  $\mathbf{x} \in \mathbb{R}^d$ . We will use  $\mathcal{Z}$  to denote the archive, which is a collection of past models used to create new proposals in the Markov chain; the archive is updated every  $K$  iteration. The posterior should be formed by taking samples from  $\mathcal{Z}$  and ignoring initial and burn-in samples. In our setting, we propose to take the last  $2N_c$  samples from  $\mathcal{Z}$ , where  $N_c$  is the number of chains. It means that the two last archived samples of each chain are conserved for the posterior.

The initial archive  $\mathcal{Z}$  is composed of  $N_p$  (number of prior samples in the archive) random normal vectors:

$$\mathcal{Z} = \left\{ \mathbf{x}_l : \mathbf{x}_l \sim \mathcal{N}(0, \mathbf{I}^d), l \in \{1, 2, \dots, N_p\} \right\}, \quad (15)$$

with  $\mathbf{I}_d$  identity matrix of size  $d \times d$ , and  $\mathcal{N}$  normal multivariate distribution. We used subscript to index different chains and not vector elements.

Similarly, the initial vector in each chain is sampled from  $\mathcal{N}(0, \mathbf{I}_d)$ .

In each chain  $i \in \{1, 2, \dots, N_c\}$ , for all  $t < T$ , ( $t$  is the iteration index) a transition from the current point  $\mathbf{x}_i$  to a new point  $\mathbf{x}'_i$  is proposed. There are two ways to generate a proposal point in DREAM-ZS: either by the parallel update or by the snooker update. The snooker update is applied with a certain frequency ( $f_s$ ), otherwise, parallel update is applied. The proposed point is always accepted if its likelihood  $L(\mathbf{x}'_i)$  is higher than  $L(\mathbf{x}_i)$ , otherwise it is accepted with probability  $L(\mathbf{x}'_i)/L(\mathbf{x}_i)$ . If the point is accepted, we set  $\mathbf{x}_i = \mathbf{x}'_i$ , otherwise, the state of the chain remains unchanged. Every  $K$  iterations the archive  $\mathcal{Z}$  is updated:

$$\mathcal{Z} = \mathcal{Z} \cup \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{N_c}\}. \quad (16)$$

The snooker update was described by ter Braak and Vrugt [2008] and parallel update by Laloy and Vrugt [2012]: Laloy and Vrugt [2012] suggested that every fifth iteration, the jump size  $\gamma$  is set to 1, in our implementation, it is set every fifth iteration on average.

The final piece of the DREAM-ZS algorithm is the implementation of crossover values, which has two ingredients: determination of  $CR$  value for each chain, and the  $CR$  distribution improvement. Unlike Vrugt *et al.* [2009], we improve the  $CR$  distribution at every iteration until the end of the DREAM-ZS algorithm.

The determination of  $CR$  value for the chain  $i$  proceeds as follows. Supposing that we have a probability vector  $p \in \mathbb{R}^{n_{CR}}$  such that  $p_m \in [0, 1]$  and  $\sum_{m=1}^{n_{CR}} p_m = 1$ . The value  $v_i$  is drawn from a categorical distribution with possible values  $\{1, 2, \dots, n_{CR}\}$  and corresponding probabilities  $\mathbf{p}$ . Let us use  $\mathcal{M}(\{1, 2, \dots, n_{CR}\}, \mathbf{p})$  to denote such a categorical distribution. The corresponding  $CR$  value is set as:  $CR = 1/v_i$ .

Now, we need a way to update the vector  $\mathbf{p}$ . The initial values of elements for the vector are  $p_m = 1/n_{CR}$  for  $m \in \{1, 2, \dots, n_{CR}\}$  and these values are recalculated after each DREAM-ZS iteration. Let  $\mathbf{v} \in \{1, 2, \dots, n_{CR}\}^{N_c}$  be the vector with the sampled  $CR$  values for each chain. Let us define vectors  $\Delta, \mathbf{L} \in \mathbb{R}^{n_{CR}}$ , initialized with zero vectors, whose elements

**Table 1.** Layers of the convolutional neural network used for generator

Layer type	Kernel	Stride	Padding	Output shape
Input				50
2D transp. conv.	$4 \times 4$	$1 \times 1$	$0 \times 0$	
Batch norm, ReLU				$1024 \times 4 \times 4$
2D transp. conv.	$4 \times 4$	$2 \times 2$	$1 \times 1$	
Batch norm, ReLU				$512 \times 8 \times 8$
2D transp. conv.	$4 \times 4$	$2 \times 2$	$1 \times 1$	
Batch norm, ReLU				$256 \times 16 \times 16$
2D transp. conv.	$4 \times 4$	$2 \times 2$	$1 \times 1$	
Batch norm, ReLU				$128 \times 32 \times 32$
2D transp. conv.	$4 \times 4$	$2 \times 2$	$1 \times 1$	
Batch norm, ReLU				$64 \times 64 \times 64$
2D transp. conv.	$4 \times 4$	$2 \times 2$	$1 \times 1$	$1 \times 128 \times 128$
Tanh				$1 \times 128 \times 128$

are updated as follows:

$$\Delta_m = \Delta_m + \sum_{i=1}^{Nc} \mathbb{1}_m(v_i) \sum_{j=1}^d ((\mathbf{x}'_i[j] - \mathbf{x}_i[j])^2 / r_j^2),$$

$$m \in \{1, 2, \dots, n_{CR}\},$$

with  $r_j^2 = \text{Var}(\{\mathbf{x}_1^t[j], \mathbf{x}_2^t[j], \dots, \mathbf{x}_{Nc}^t[j]\})$  the variance of the parameter  $j$  of the state vector among all chains. The square brackets serve to obtain elements of a vector  $\mathbf{x}$ . The vector  $\mathbf{L}$  counts how many times each  $CR$  value was drawn:

$$L_m = L_m + \sum_{i=1}^{Nc} \mathbb{1}_m(v_i), \quad m \in \{1, 2, \dots, n_{CR}\}. \quad (17)$$

Finally, we can state the update of the vector of probabilities:

$$p_m = \frac{\Delta_m / L_m}{\sum_{j=1}^{n_{CR}} (\Delta_j / L_j)}, \quad m \in \{1, 2, \dots, n_{CR}\}. \quad (18)$$

### 2.3.2. WGAN

Generative adversarial neural networks (GAN) can learn a complex mapping between a latent space and the space of two-dimensional images [Goodfellow *et al.*, 2016]. In this work, we decided to use the Wasserstein GAN [Arjovsky *et al.*, 2017] with gradient penalty term, which is claimed to be robust for changing architecture of the network [Gulrajani *et al.*, 2017]. GANs are composed of two neural networks: a critic (discriminator) and a generator. The generator maps the latent space vectors to the image space. The critic is fed by the output from the generator or real

images (the images from the training set) and predicts if the images are fake (generated images) or not. The goal of the generator is to deceive the critic so that it cannot distinguish the generated images from the images of the training set. Typically, for an epoch (GAN training iteration), the critic is optimized several times after a single generator training.

In our case, the latent space has  $d$  dimensions and the images represent the geology. While the generated images (GAN output) have values between  $[-1, 1]$ , they can be converted to binary images by applying a threshold (0). However, the threshold is not applied when evaluating the likelihood of the model. Instead, the physical parameters are linearly transformed from the pixel values, with  $-1$  and  $1$  corresponding to the exact values according to facies. The training set contains all the possible extractions from the TI of the size  $128 \times 128$ . The training batch size is 64, the learning rate  $1 \times 10^{-4}$ , the ADAM optimizer was used with beta parameters: 0.5 and 0.999. There are 5 critic iterations per generator iteration, and the lambda term for the gradient penalty was set to 10. The architectures of the generator and the critic are shown in Tables 1 and 2, respectively.

**Table 2.** Layers of the convolutional neural network used for critic

Layer type	Kernel	Stride	Padding	Output shape
Input				$1 \times 128 \times 128$
2D conv.	$4 \times 4$	$2 \times 2$	$1 \times 1$	
Instance norm, leakyReLU				$64 \times 64 \times 64$
2D conv.	$4 \times 4$	$2 \times 2$	$1 \times 1$	
Instance norm, leakyReLU				$128 \times 32 \times 32$
2D conv.	$4 \times 4$	$2 \times 2$	$1 \times 1$	
Instance norm, leakyReLU				$256 \times 16 \times 16$
2D conv.	$4 \times 4$	$2 \times 2$	$1 \times 1$	
Instance norm, leakyReLU				$512 \times 8 \times 8$
2D conv.	$4 \times 4$	$2 \times 2$	$1 \times 1$	
Instance norm, leakyReLU				$1024 \times 4 \times 4$
2D conv.	$4 \times 4$	$1 \times 1$	$0 \times 0$	1

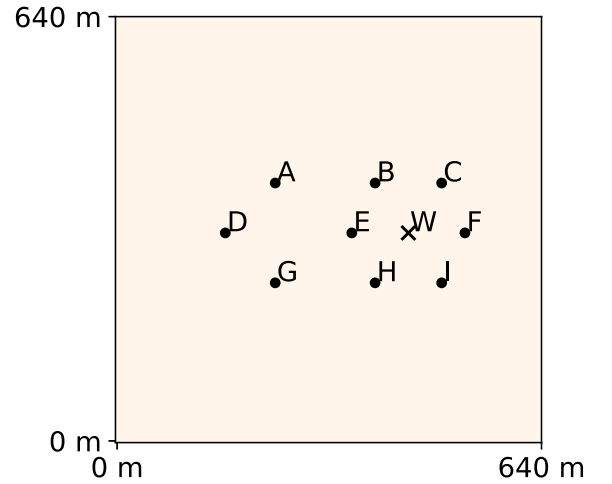
### 3. Test case

#### 3.1. The inverse problem

We consider a pumping test in a confined aquifer of thickness 10 m. At the beginning of the test, the hydraulic heads are uniform and constant at 0 m. Water is pumped with a constant discharge rate of  $0.08 \text{ m}^3/\text{s}$  during 2 h. The hydraulic heads are recorded in the pumping well and nine piezometers in the vicinity of the pumping well (Figure 1 and Table 3). The hydraulic heads are recorded every 100 s, so that 72 measurements are available at each of the 10 locations, which makes up for a total of  $N = 720$  measurement points (Figure 2).

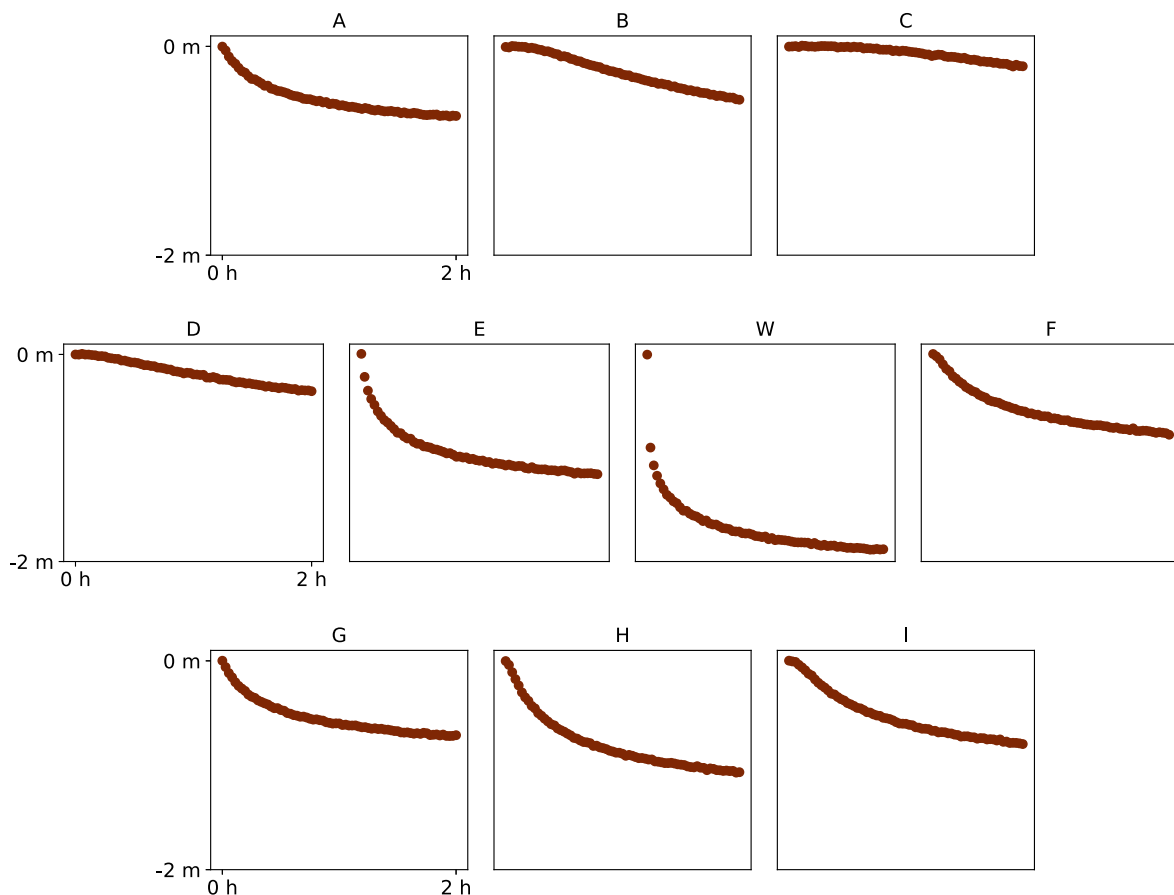
#### 3.2. The reference set-up

The data presented in the previous section were obtained from a synthetic setup. We will refer to it as the reference. It is not the solution to the inverse problem framed in a probabilistic manner. It is rather a model which has a very high likelihood, given the data. The domain has an extension of 640 m by 640 m. The petrophysical parameters are modeled using a categorical 2D field with two geological facies: a permeable (channels or ellipsoidal deposits, labeled with 1) and a less permeable matrix (labeled with 0). The area is discretized using a regular grid with cells of size 5 m by 5 m, thus the grid contains 128 by 128



**Figure 1.** Position of the pumping well (W) and nine piezometers (A–I) in the pumping test.

cells. The two geological facies have constant hydrogeological parameters (Table 4). The boundary conditions are constant head values at all edges, equal to 0 m, and at the beginning of the pumping test, hydraulic charge equals to 0 m everywhere in the domain. The reference field was created using the DeeSse software, which is an implementation of Direct Sampling algorithm with pyramids [Straubhaar *et al.*, 2020]. An extended image of a channelized aquifer was used as the training image (Figure 3). We chose to run the DeeSse in the Direct Sampling Best Candidate (DSBC) mode, which boils down to choos-



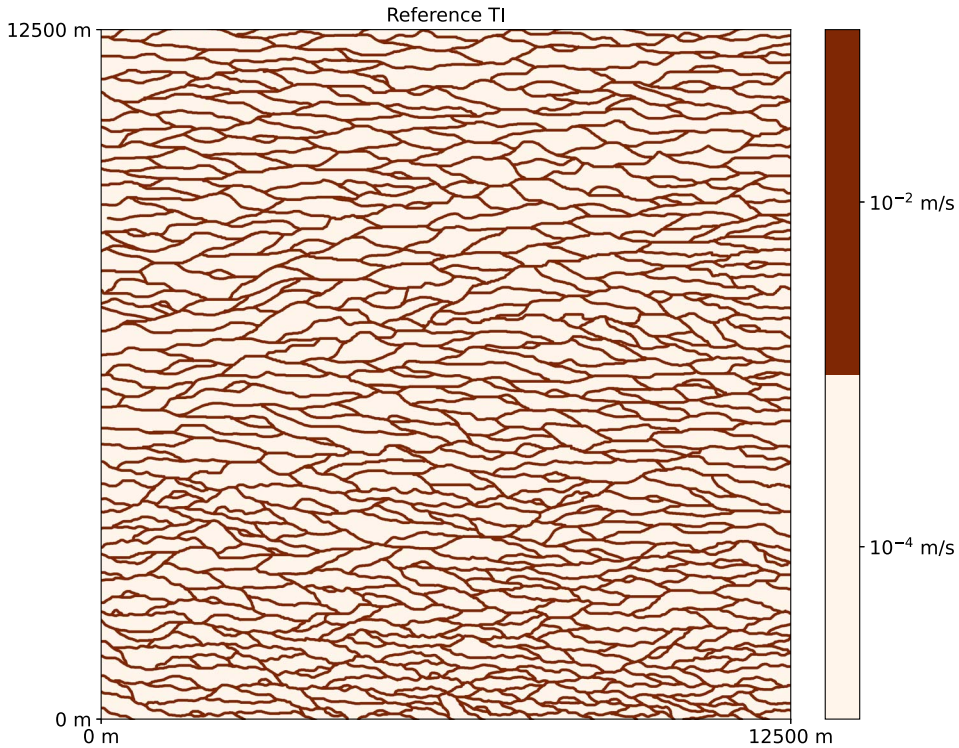
**Figure 2.** Time series of hydraulic heads recorded at nine piezometers (A–I) and the pumping well (W).

**Table 3.** Positions of piezometers ( $x$ ,  $y$  coordinates), corresponding columns and row indexes, and labels

$x$ (m)	$y$ (m)	Column index	Row index	Label
242.5	392.5	48	78	A
392.5	392.5	78	78	B
492.5	392.5	98	78	C
167.5	317.5	33	63	D
357.5	317.5	71	63	E
442.5	317.5	88	63	W
527.5	317.5	105	63	F
242.5	242.5	48	48	G
392.5	242.5	78	48	H
492.5	242.5	98	48	I

ing a threshold of 0 in the standard Direct Sampling, or very small, close to 0, if the software does not allow for non-positive input. The maximal scan fraction was set to 0.01 and the number of neighboring nodes to 40. Two pyramid levels were used, with the reduction by 2 in each direction at every level. The groundwater flow was simulated using the FloPy python package [Bakker et al., 2016], which is a wrapper for the MODFLOW software [Hughes et al., 2017]. To emulate the measurement error, the obtained values of hydraulic heads were corrupted with Gaussian noise with mean 0 m and standard deviation 0.005 m. These data will be used as input for the different inversion procedures.

To evaluate the quality of the inversion methods, we use in addition a prediction problem. The prediction data will not be used by the inversion algo-



**Figure 3.** Reference (true) training image, used to generate the reference (“true”) field. Data from Zahner et al. [2016].

**Table 4.** Hydrogeological parameters of different geological facies considered in the study

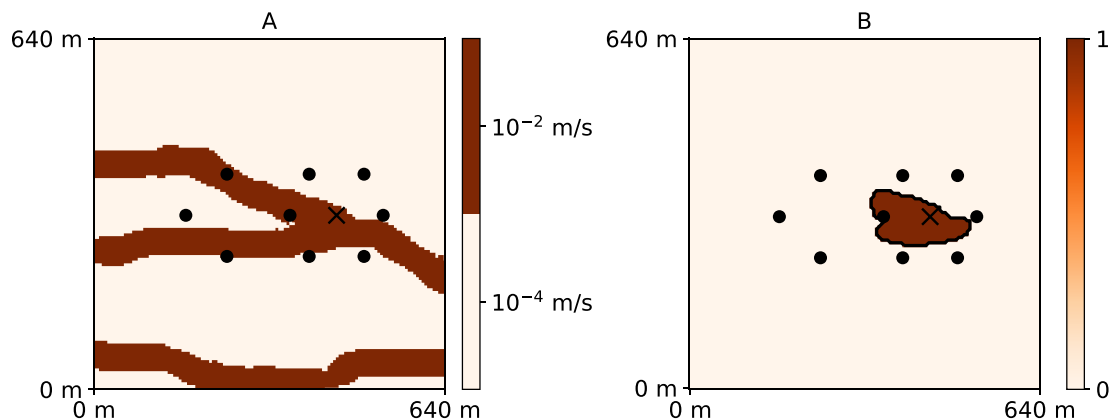
	Less permeable (0)	More permeable (1)
Hydraulic conductivity (m/s)	$1 \times 10^{-4}$	$1 \times 10^{-2}$
Specific storage ( $m^{-1}$ )	$5 \times 10^{-4}$	$5 \times 10^{-5}$
Porosity	0.4	0.3

rhythms. Here we consider, the prediction of the 10-day groundwater protection zone, it is also referred to as the 10-day capture zone [van Leeuwen et al., 1998]. It is calculated in a slightly different setup but with the same geological model and its properties. The boundary conditions are prescribed heads equal to 1 m on the left boundary, 0 m on the right boundary and interpolated between those two values on the upper and the lower boundaries. A constant pumping rate of  $0.04 \text{ m}^3/\text{s}$  is imposed at the well and forward particle tracing is performed on the steady-state solution of groundwater flow. The 10-day zone contains each location (pixel), from where

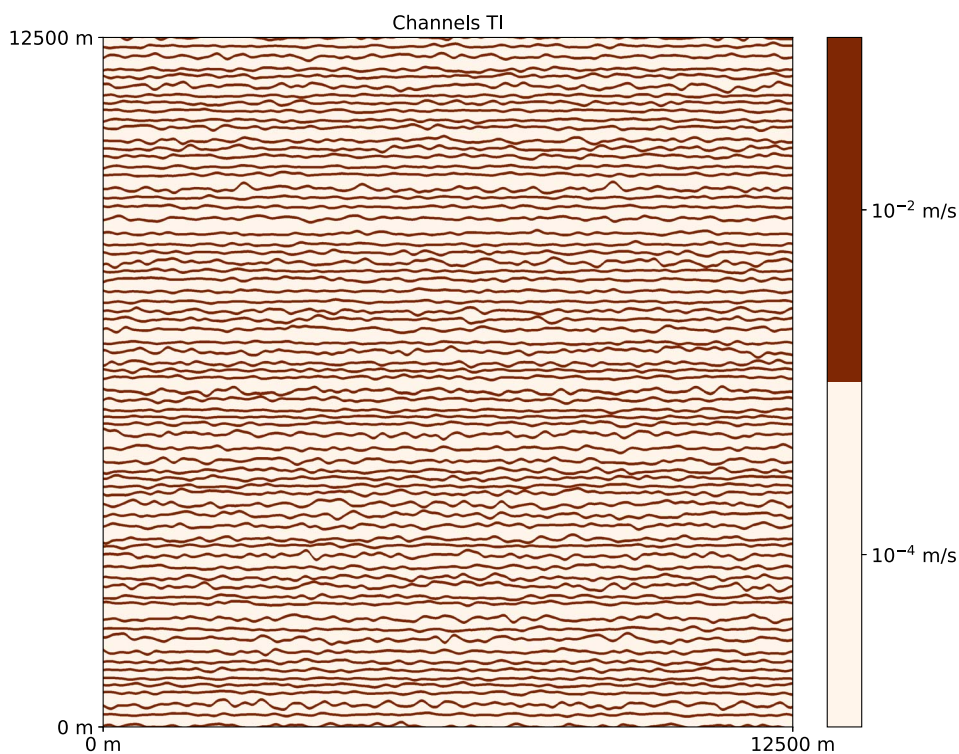
groundwater reaches the pumping well in less than 10 days.

### 3.3. Inversion set-up

We will perform the inversion three times for each method. Every time, we use a different training image. The baseline case uses the reference TI, the two other cases use a different one. In this way, the robustness of the inversion methods can be tested. The two other training images were generated with the TI generator tool [Maharaja, 2008] in the AR2GEMS software. Their size is identical to the one of the



**Figure 4.** The reference field (A) used as the synthetic reality (considered unknown) and the corresponding 10-day groundwater protection zone (B).

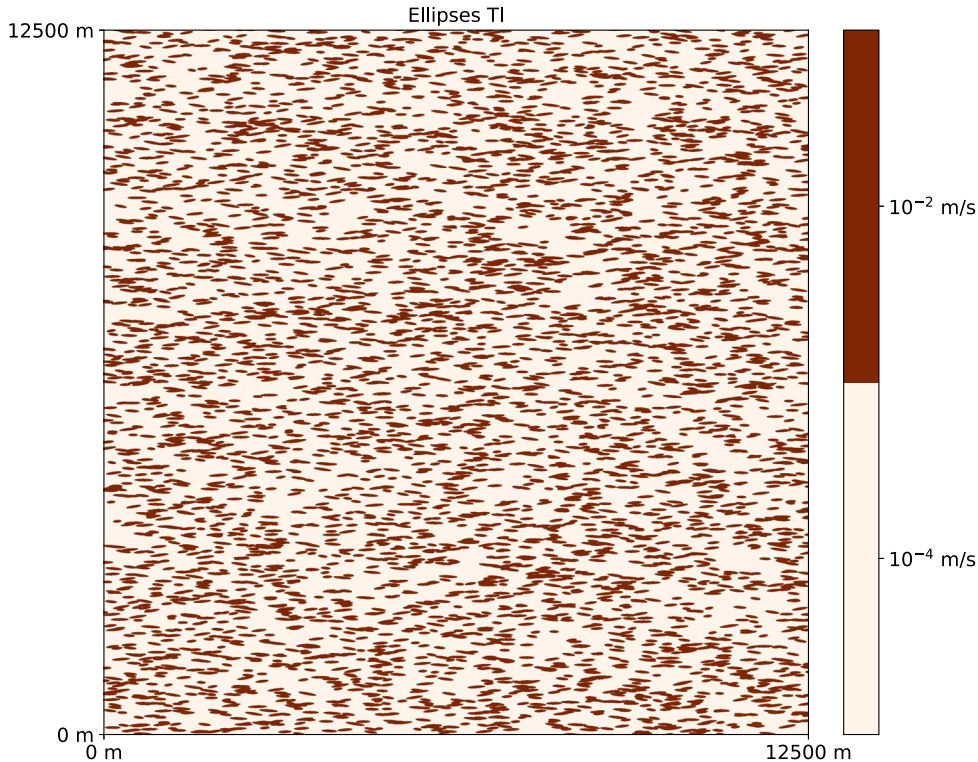


**Figure 5.** “Channels TI”, an alternative training image.

original TI. The first is a channelized medium with disjoint channels, and the second represents ellipsoidal deposits. TIs have the same proportion of facies as the original TI. We will refer to them as “Channels TI” (Figure 5) and “Ellipses TI” (Figure 6).

Both PoPEX and ESMDA benefit from the same

geostatistical engine as the original reference; it means that the same DeeSse parameters are used in PoPEX + MPS and in ESMDA + MPS to generate the reference. In theory, it is possible (but highly unlikely) that PoPEX samples the same model as the reference. It might not be the case for ESMDA, as it im-



**Figure 6.** “Ellipses TI”, an alternative training image.

poses dense conditioning on a coarse grid, but very similar models can be produced. The fact that PoPEX and ESM DA use DeeSse, gives them an edge compared to DREAM-ZS+GAN, as the reference is a realization generated by DeeSse (therefore not present in the TI). GAN is trained on images cut from the TI and learns to develop similar realizations, but it does not have access to samples simulated with MPS.

It is important to note that the three inverse methods identify a different number of unknowns. For PoPEX, the number of unknowns is simply the number of grid cells in the domain, i.e.  $(128 \times 128) = 16,384$ . For ESM DA, the number of unknowns is reduced, as compared to PoPEX, because they are the continuous values on the low-resolution map used to constrain the MPS realizations. In the example, two pyramid levels are added, each one being obtained by dividing by 2 the number of cells along each axis, and 20% of the cells in the coarse level are updated by the procedure, which results in a total of  $0.2 \times (128 \times 128) / (4 \times 4) \approx 204$  unknowns. Finally, for DREAM-ZS+GAN, the number of unknowns is the

size of the latent Gaussian vector used as input to the GAN (Table 1), and it is only 50. However, note that the above calculation omits the fact that, for the two first approaches, the values in the grid cells are correlated via the MPS statistics, and it is therefore difficult to estimate the actual dimension of the underlying parameter space.

PoPEX was run with the following parameters: 32 parallel processes, and for a total of 50,000 iterations (50,000 forward runs), the maximal number of conditioning points is 10. The choice of the number of parallel processes depends on available computing resources, in our case it was adjusted to the number of cores in a computing node. The number of forward runs and conditioning points are similar to values suggested in the paper introducing the parallel PoPEX algorithm [Jäggli *et al.*, 2018], where the problem size was of comparable dimensions. ESM DA was run for 16 iterations (16 data assimilations), the size of ensemble is 128 (a total of 2048 forward runs) and number of parallel processes 32. The chosen parameters are again close to those used in

the paper which introduced the method Lam [2019], which also suggested that using larger number of iterations would not lead to better results. DREAM-ZS was run for total  $T = 5000$  iterations with 32 parallel chains (160,000 forward runs). The initial archive size is set to 500. The size of the latent vector is  $d = 50$ . Frequency of snooker update is 0.1,  $\delta_{\max} = 1$ , and  $n_{CR} = 3$ , according to recommendations for standard parameters. To study the convergence, we set a fairly low value for  $K$ , equal to 10, but the value 100 should be sufficient for the chosen number of iterations. Moreover, we set  $b = 0.05$  and  $b^* = 1 \times 10^{-6}$ . These parameters have been tested in the paper presenting the DREAM-ZS with GAN for a hydrogeological problem [Laloy et al., 2018]. For practical reasons, to avoid too long computing times, we chose a smaller size of the latent vector and set a limit on the number of iterations.

#### 4. Comparing the results

Since our numerical experiments involve a large number of results, we need to use summary statistics to compare the different methods efficiently. For that purpose, we will use three metrics for comparing the quality of the results:

- (1) The first metric indicates how well the measurements are reproduced by the simulation ensemble. Each continuous measured value will be compared with the ensemble of simulated values. A score that allows for such a comparison is the continuous ranked probability score (CRPS) used to assess probabilistic forecasts [Gneiting et al., 2007, Gneiting and Raftery, 2007].
- (2) The second metric indicates how well the protection zone is predicted. In this case, a discrete (categorical) value is compared with a probability for each point using a quadratic score [Gneiting et al., 2007].
- (3) Finally, the third metric indicates how well the geology is identified. For that purpose, we use the average of pixel-by-pixel quadratic scores of predicted facies.

Below, we recall the definition of the CRPS and quadratic (Brier) scores.

##### 4.1. CRPS score

The continuous ranked probability score (CRPS) is given by:

$$crps(F, x) = \int_{-\infty}^{\infty} (F(y) - \mathbb{1}(y \geq x))^2 dy, \quad (19)$$

where  $F$  is the predictive cumulative distribution function (CDF), and  $x$  the true (observed) value, and  $\mathbb{1}(y \geq x) = 1$  if  $y \geq x$  and 0 otherwise. The advantage of the CRPS score is that it is expressed in the same units as the observations, and it generalizes absolute error, as when the CDF becomes a point observation, it equals the absolute error. In this way, it also provides a way to compare probabilistic and deterministic forecasts.

In our context, the *crps* score will be averaged for all measurement points:

$$CRPS = \frac{1}{N} \sum_{i=1}^N crps(F_i, d_i^{\text{obs}}), \quad (20)$$

where  $F_i$  is the predicted CDF corresponding to the measurement point  $d_i^{\text{obs}}$ .

##### 4.2. Quadratic (Brier) score

The quadratic score was first introduced by Brier [1950] to quantify forecasts of categorical variables expressed in terms of probabilities. The quadratic (aka Brier) scoring rule is given by:

$$bs(\mathbf{p}, i) = \sum_{j=1}^M (\delta_{ij} - p_j)^2, \quad (21)$$

where  $\mathbf{p}$  is a discrete probability distribution, i.e.  $\sum_{k=1}^M p_k = 1$ , and  $p_k \geq 0$  for all  $k \in \{1, 2, \dots, M\}$ ,  $M$  the total number of categories, in this case equal to 2.  $\delta_{ij}$  is the Kronecker delta  $\delta_{ij} = 1$  if  $i = j$ , and 0 otherwise.

In our context, we will average the *Brier* score over all locations where it is predicted if that location is in the groundwater protection zone. Let  $\mathbf{p}_k$  be a probabilistic forecast if point  $k$  is in groundwater protection zone, and  $l_k$  true category (in this case binary) of point  $k$ . The average Brier score is then given by:

$$BS = \frac{1}{N_l} \sum_{k=1}^{N_l} bs(\mathbf{p}_k, l_k), \quad (22)$$

with  $N_l$  the total number of locations. In the case of geology, the score is calculated in the same manner, but the facies data is used instead of the protection zone.



## 5. Results

For each of the methods and for each of the TIs, we report prior data coverage and posterior data match, and corresponding prior and posterior probability maps (groundwater protection zone, facies) with examples of realizations. We also compare the convergence of the methods with respect to the number of forward calls.

### 5.1. *Prior distributions*

The prior groundwater protection zones are essentially circular (Figure 7), as the most probable event is that the well is placed in the less impermeable facies. Such a configuration results in a large (and unrealistic) drawdown. If we look more closely at the prior groundwater protection zone maps, we can see a second mode, which is an ellipse. This occurs when the pumping well intersects a channel. Such a zone shape is visible for the Channels TIs (Figure 7).

The channel prior probability maps are roughly homogeneous for all the simulation methods and training images (for example, see Figure 7). This is expected as we did not impose any prior conditioning data, the prior probability of a channel corresponds then to the proportion of channels in the training image.

For the drawdown curves in the piezometers and pumping well, the prior distribution shows wide confidence intervals, as can be seen for DREAM-ZS and Ellipses TI for example (Figure 10). The true data are usually within the 95% confidence intervals, but there are some exceptions, and it happens that some head data lie outside of this range.

### 5.2. *Posterior distributions*

The confidence intervals for the drawdown curves are very much reduced, and they mostly match the data well (Figures 8–10).

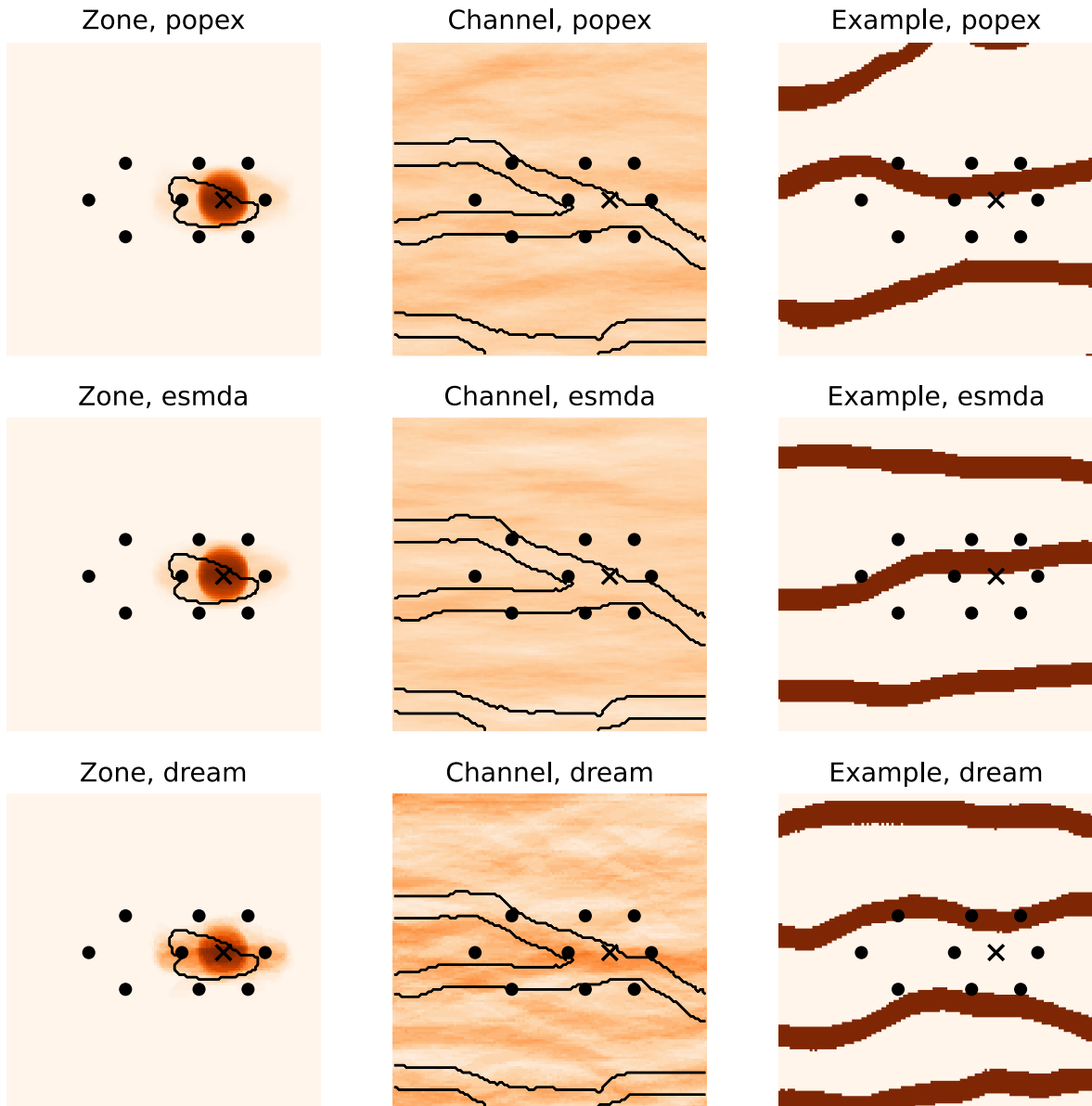
Even if we do not show the figures here for the sake of brevity, the posterior distribution computed using the reference TI with all the methods achieved a satisfactory data fit, with DREAM-ZS and PoPEX producing very narrow confidence intervals and an excellent match and ESDMA producing wider confidence intervals.

With the Channels TI, PoPEX produced wider confidence intervals, and some piezometric data (A, F, H, I) are not matched perfectly, but the fit is reasonable. ESDMA achieved slightly worse piezometer data fit, but the well data is reproduced poorly, giving a high probability of the well placement in (or close to) the less permeable region. DREAM-ZS achieves a very close fit and provides narrow confidence intervals.

The Ellipses TI is the most difficult one. PoPEX does not match the data well for some piezometers (A,C,F) and produces quite wide confidence intervals (Figure 8). ESDMA produces wide confidence intervals, and the head data in the pumping well are poorly represented (Figure 9). DREAM-ZS produced very narrow confidence intervals, but the data are sometimes not very well matched (Figure 10F).

In terms of posterior probability maps, all methods solved reasonably well the inverse problem in the reference case. The protection zone probability maps are very close to the reference protection zone for all methods (Figure 11). The permeable facies probability maps were able to represent the bifurcation. We note however some general trends to predict with over-confidence certain geological features. For PoPEX, on the right side of the map, the channel goes straight with a high probability, while in the reference it goes slightly to the bottom. For the ESDMA algorithm, the posterior probability map suggests some “eye” feature to the left of the bifurcation. This pattern is not suggested by PoPEX. For DREAM-ZS, the “eye” structure is even more pronounced and the same map indicates with high probability a channel at the bottom, which only partially coincides with the channel in the reference realization. These rather high posterior probabilities of the presence of certain geological features that are not present in the reference seem to correspond to some artifacts of the methods and not to features suggested by the statistics of the TI.

The case of the Channels TI (Figure 12) posed more challenges for the inversion algorithms. The protection zone was only very well represented by the DREAM-ZS method. The zone obtained with PoPEX is more elongated and less influenced by the bifurcation. Indeed, PoPEX indicates a higher probability of channel only in the lower branch. Due to this TI with disjoint channels, it had difficulties in reproducing the branching of channels. ESDMA did not attribute a high probability of permeable facies near

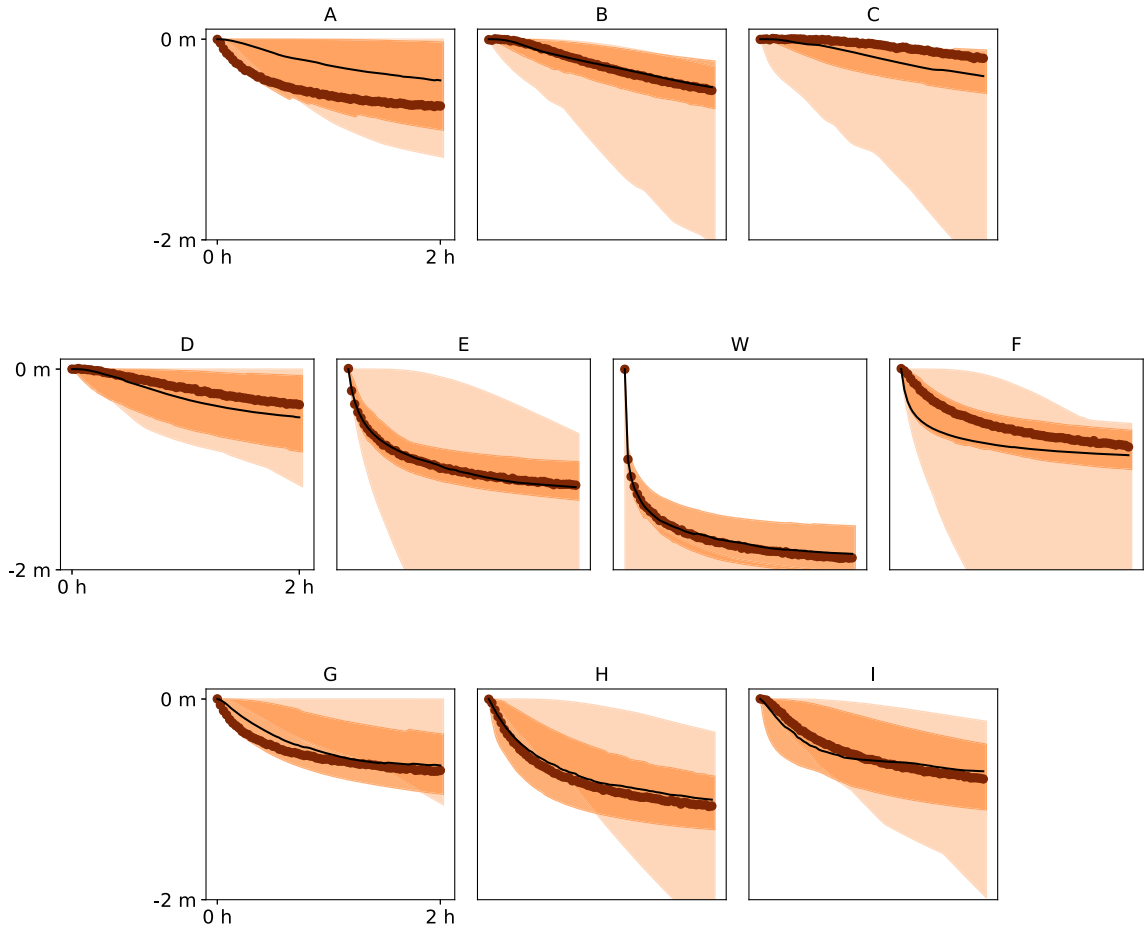


**Figure 7.** Prior probability distributions obtained with the Channels TI for the groundwater protection zone (left column), and the channel occurrence (middle column). The black contours correspond to the reference. The right column shows one example of realization. The top row shows results obtained with PoPEX, the middle row shows results obtained with ESM DA, and the bottom row shows results obtained with DREAM-ZS. The colormap is the same as for Figure 4B.

the pumping well. However, it managed to find the double channel on the left side of the field. DREAM-ZS performed best in reproducing the bifurcation, but it also displays channel artifacts in the upper part of the image.

The Ellipses TI can be considered as the hard-

est case, due to the disconnected nature of the high permeability features (the ellipses). The solution provided by PoPEX can be thought of as the most conservative, as it only places a blurred region of higher channel probabilities around the pumping well (Figure 13). The advantage of this solution is



**Figure 8.** Prior and posterior distributions of the drawdown curves obtained with PoPEX and the Ellipses TI at the nine piezometers (A–I) and pumping well (W). The observed data is marked with thick brown dots, the median of the posterior distribution with thin black line, the 95% confidence intervals of the posterior are shown as dark shaded region, and those of the prior distribution are shown as the light shaded region.

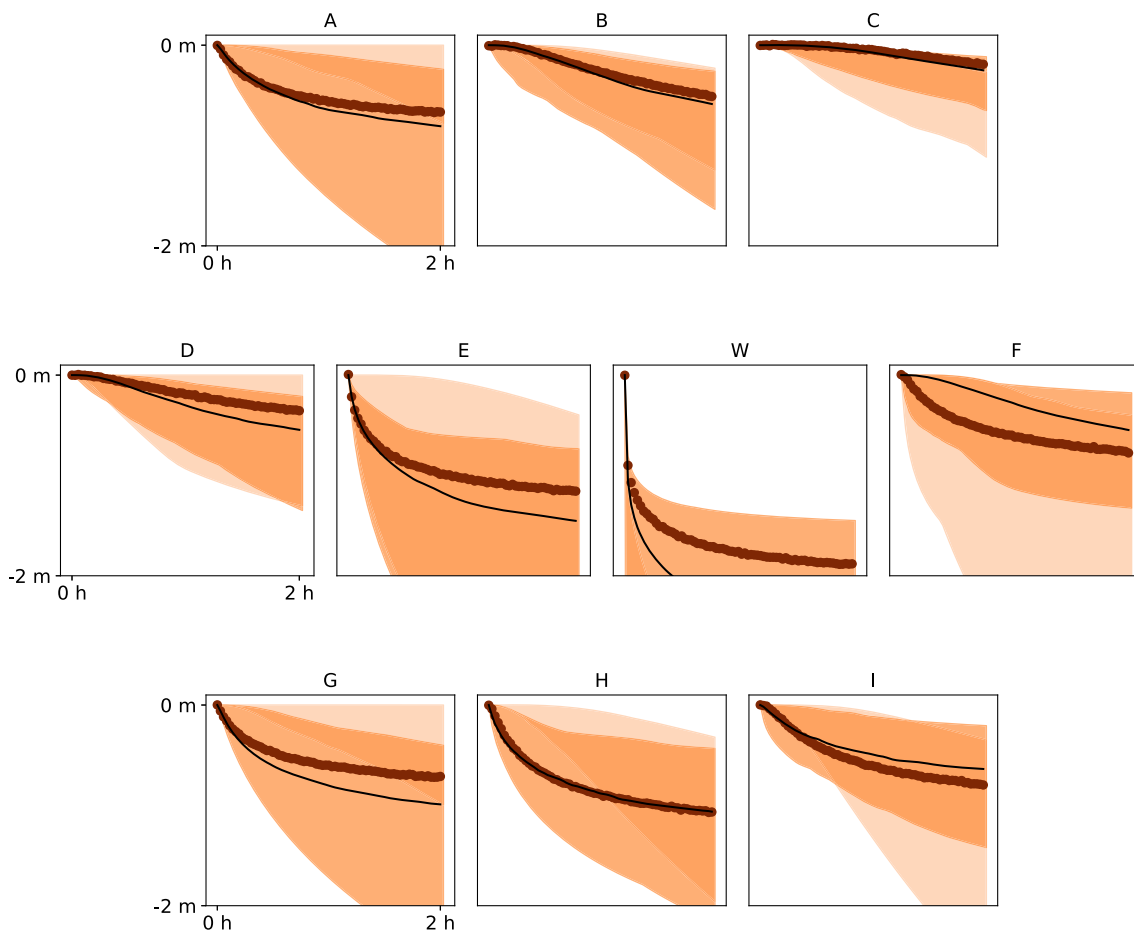
the absence of a high probability of channels in spurious zones. Nevertheless, the protection zone is not accurately reproduced. ESMDA placed a smaller protection zone than it should be and indicated with high certainty the presence of channels in regions where they should not be. DREAM-ZS provides the most contrasted probability maps and places incorrect geological features in the whole area.

### 5.3. Convergence and quality metrics

The previous comparison of the posterior distributions shows that it is not simple to compare visually

the results. Therefore, in this section, we compare the convergence of three quality metrics to get a better understanding of the performances of the methods. We grouped the convergence plots by methods in Figure 14. The quality metrics are plotted as a function of the number of forward model runs since these runs constitute the most expensive computational cost during the inversion.

Figure 14 shows that all the methods produced the best results (smaller values of the quality indicators) for the true training image. The methods also converge, i.e. the errors diminish, with the number of forward calls (iterations) for the true training image. Generally, the scores are similar or better for

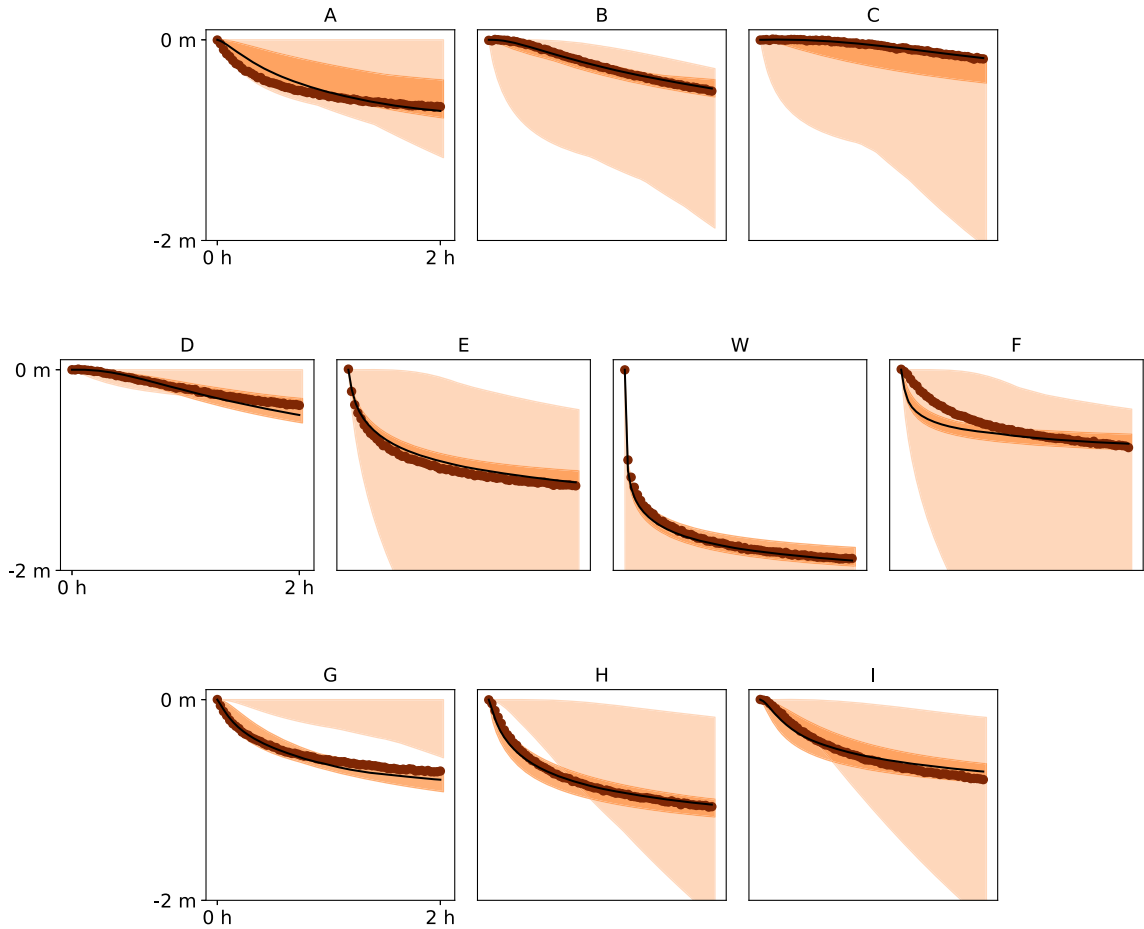


**Figure 9.** Prior and posterior distributions of the drawdown curves obtained with ESMDA and the Ellipses TI at the nine piezometers (A–I) and pumping well (W). The observed data is marked with thick brown dots, the median of the posterior distribution with thin black line, the 95% confidence intervals of the posterior are shown as dark shaded region, and those of the prior distribution are shown as the light shaded region.

the channel TI than for the Ellipses TI. While different scores show similar trends for the same case, a good CRPS score does not necessarily imply a good Brier score. The most notable example of the lack of correlation is the case of the DREAM-ZS algorithm. The CRPS score goes down for all TIs rather fast, and the scores for Channels TI and Ellipses TI are close. However, when comparing the BS-zone scores, the Ellipses TI results become significantly worse than those with the Channels TI. Moreover, the BS-channel score increases after  $1 \times 10^4$  forward calls. It seems that the realizations collapse on a very similar (and not fully correct) model realization, and produce geological artifacts which are highlighted by this

score, while the CRPS score on data match remains low.

When comparing the different methods for the same TIs, we note that ESMDA exhibits a very fast convergence at the beginning and then reaches a plateau. The method stops after relatively few forward runs (as compared to the other methods), and this is related to the choice of the parameter governing the number of data assimilations. Here, the default value of 16 is used [Lam *et al.*, 2020] and is based on recommendations of Emerick and Reynolds [2013]. In a sense, it is the least computationally expensive method, but the achievable quality is limited. Other methods can provide results of better

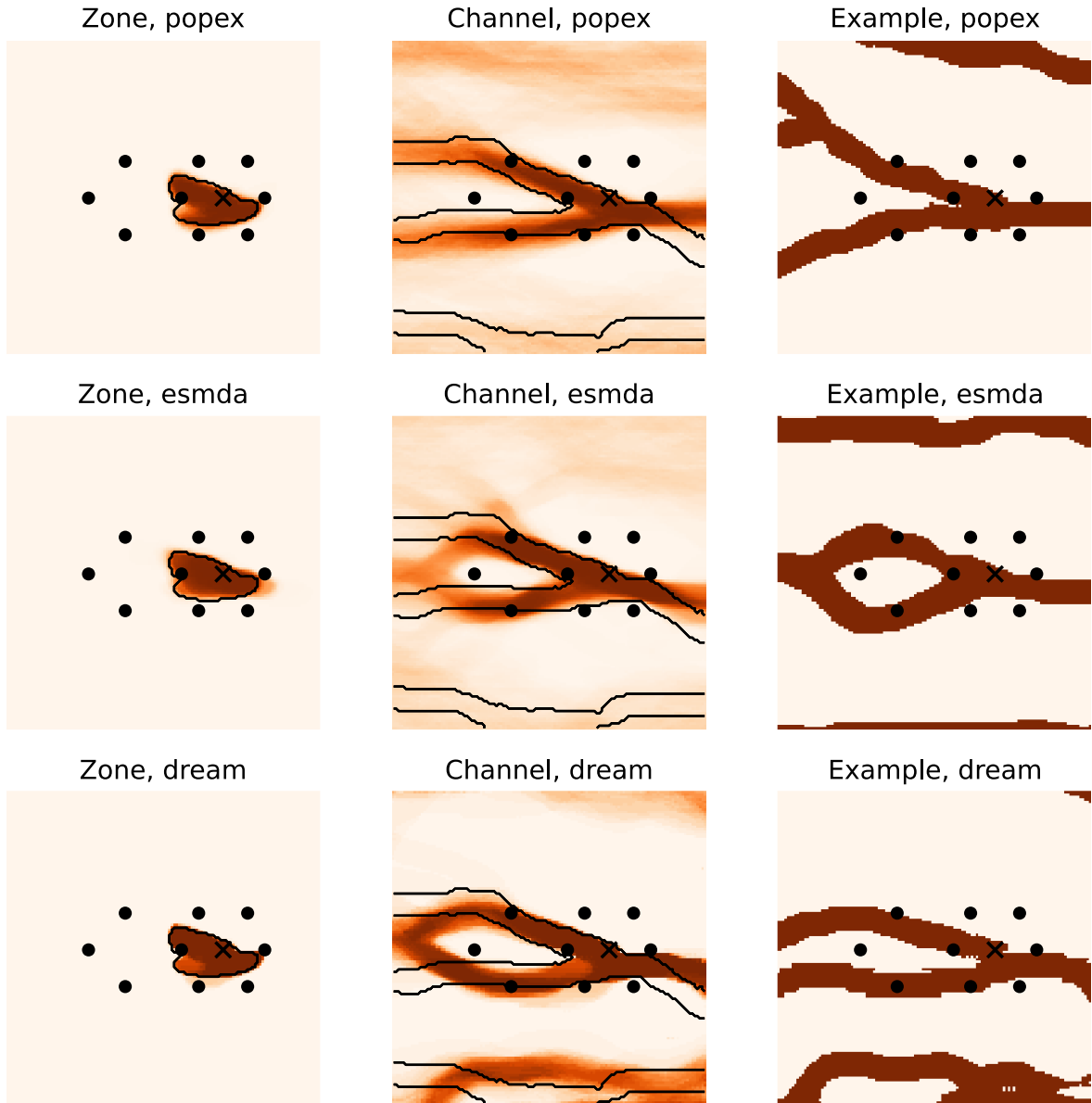


**Figure 10.** Prior and posterior distributions of the drawdown curves obtained with DREAM-ZS and the Ellipses TI at the nine piezometers (A–I) and pumping well (W). The observed data is marked with thick brown dots, the median of the posterior distribution with thin black line, the 95% confidence intervals of the posterior are shown as dark shaded region, and those of the prior distribution are shown as the light shaded region.

quality, but they need more iterations. DREAM-ZS shows the best convergence for the data match and protection zone reproduction for the Channels TI, but the error diverges for BS-channel and becomes larger than those of PoPEX and ESMDA for the largest number of iterations. We note a slow but steady convergence of PoPEX when using the Ellipses TI, as opposed to ESMDA, which stagnates after the first iterations, and DREAM-ZS, which matches the piezometric data very well but has slow convergence on the BS-zone and even diverges on the estimation of the geological features (BS-channel).

In summary, PoPEX is the method that seems to

be the most robust. It converges steadily with the number of forward model calls, and rather fast when the proper training image is given but slowly if a wrong training image is provided. On the contrary, ESMDA always improves the solution very fast for the first iterations even with a wrong training image, but it stagnates rapidly after a few iterations. DREAM-ZS has an intermediate behavior, it can be faster than PoPEX and continues to improve the results when ESMDA stagnates even with a wrong TI. But there are cases where the DREAM-ZS method diverges and the error increases with additional iterations.

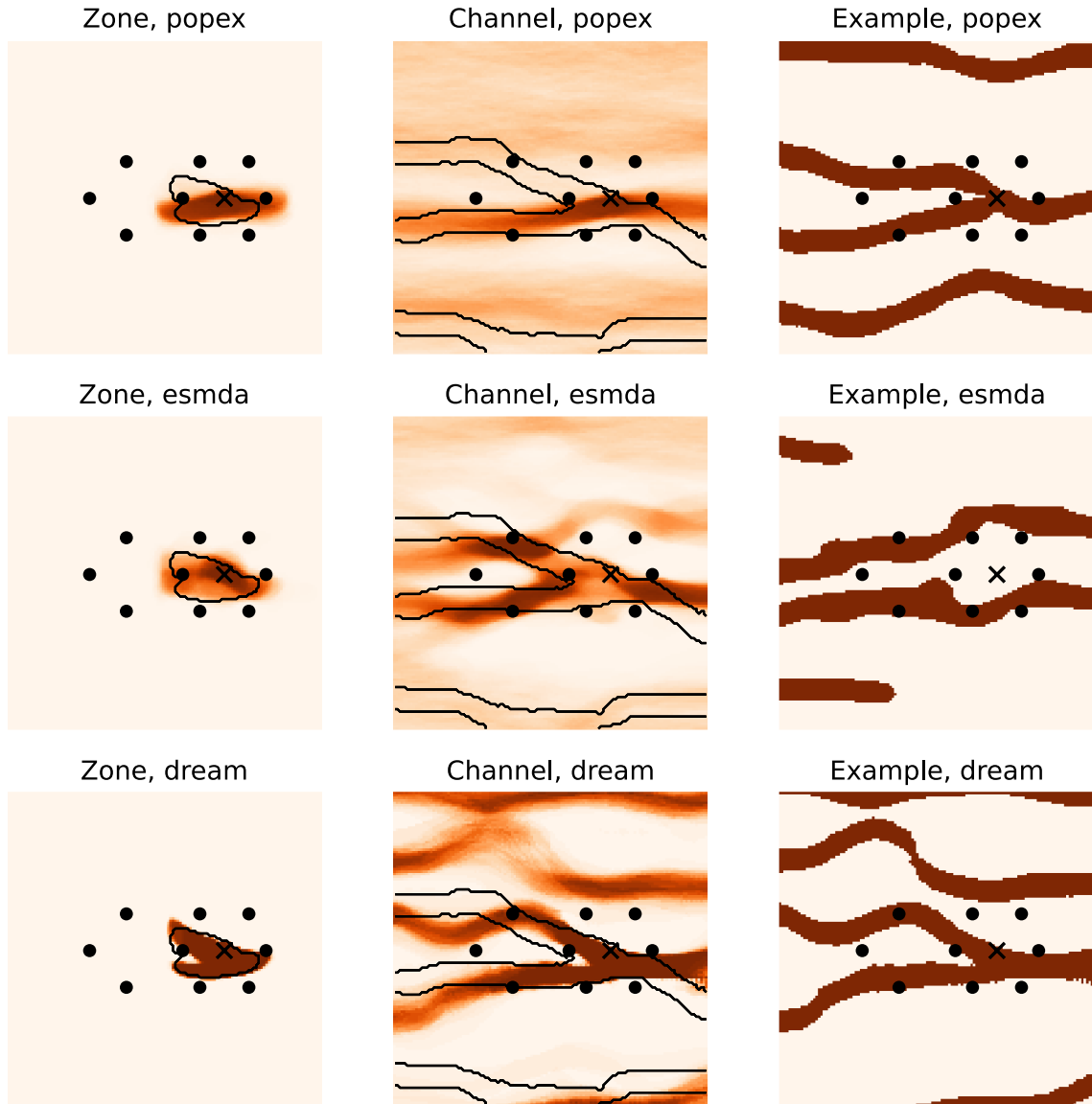


**Figure 11.** Posterior probability maps obtained with the true TI for the groundwater protection zone (left column), and the channel occurrence (middle column). The black contours correspond to the reference. The right column shows one example realization. The colormap is the same as for Figure 4B.

## 6. Conclusions

In this study, we compared three recent stochastic inversion methods capable of inverting categorical fields: Posterior Population Expansion (PoPEX) with multiple-point statistics (MPS), Ensemble Smoother Multiple Data Assimilation (ESMDA) with MPS pyra-

mids, and DREAM-ZS with Wasserstein Generative Adversarial Network (WGAN). A synthetic test case with hydrogeological data (time series of hydraulic heads) was used, and two geological facies were considered. The results were analyzed both for the inversion and for a prediction of the 10-day groundwater protection zone. The quality indicators took into



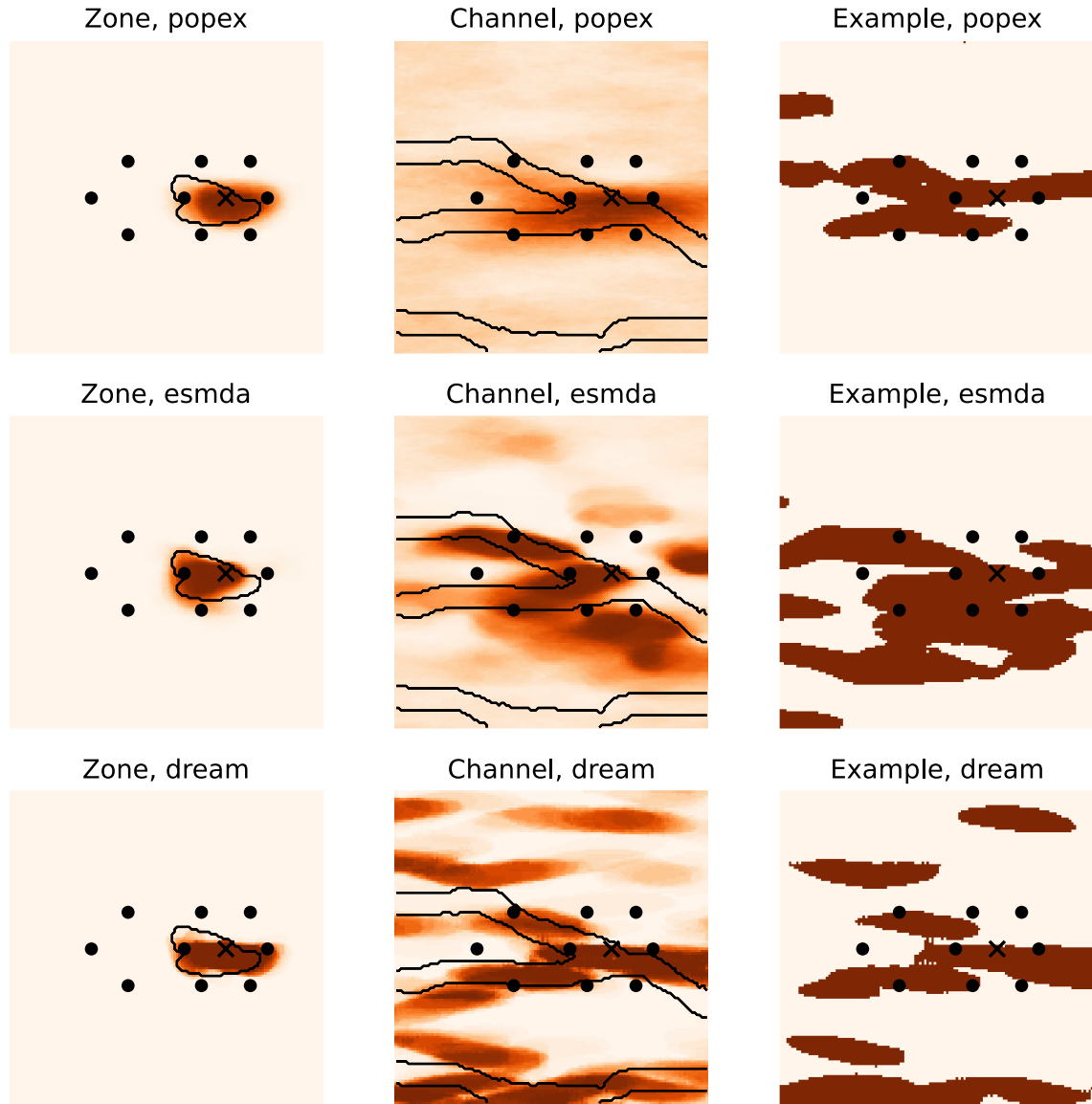
**Figure 12.** Posterior probability maps obtained and example realizations with the Channels TI.

account the reference solution (ground truth represented by the reference solutions) with probability forecasts.

Our main finding is that when the methods were given the correct prior information (represented by a training image), they all achieved reasonable convergence. Even with the wrong priors, some acceptable solutions were obtained. However, the choice of the prior is essential. The convergence was negatively affected when the TI with lenticular deposits

was used. The TI with disjoint channels (as opposed to the original TI with bifurcating channels) provided slightly better results than ellipsoidal deposits, but the two wrong TIs deteriorated the results and introduced artifacts. As previously discussed in the inter-comparison exercise of Zimmerman *et al.* [1998], we also observed that none of the methods performed systematically better than the others for all the criteria that we studied.

The advantage of PoPEX is that it presented the



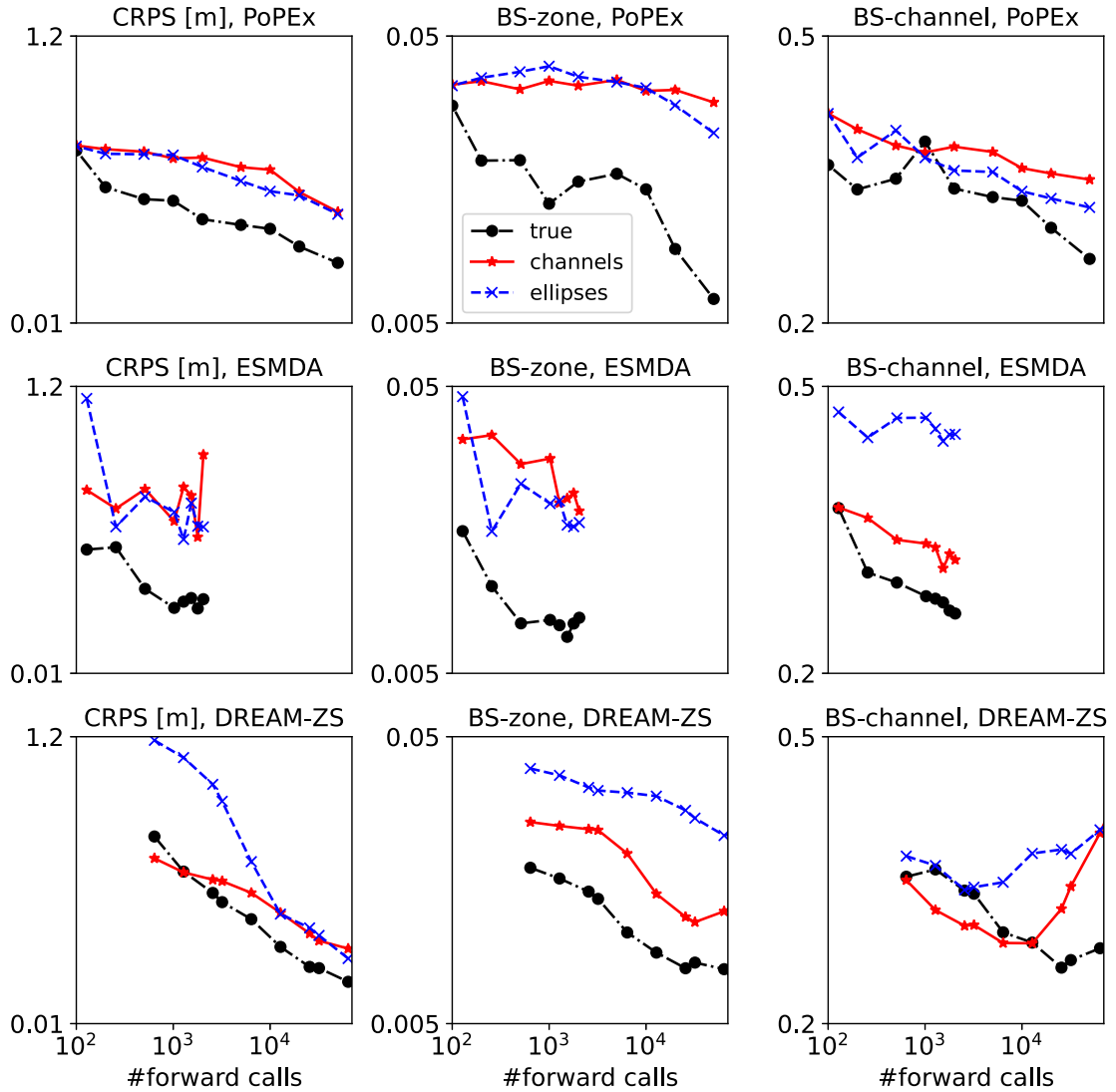
**Figure 13.** Posterior probability maps and example realizations obtained with the Ellipses TI.

most steady convergence compared with ESMDA and DREAM-ZS, whose scores fluctuated (ESMDA) or even increased (DREAM-ZS) with the number of iterations. PoPEX can also be thought of as the most “conservative”, as it does not introduce “artifacts”, e.g. geological features which are unlikely (outside the informed zone). However, this “conservative” approach leads to poorer data fit and worse predictions of the protection zone in certain cases. In the case of a wrong prior, it rarely performed better than the

other methods.

The advantage of ESMDA is its fast convergence. It was able to reasonably identify the permeable zones between the piezometers even with wrong priors, producing patterns not present in the TI, which was needed to match the data. But it also indicated with high certainty some permeability patterns outside this zone that were incorrect. A surprising point, and drawback of the method, is that it did not always place the well in the permeable zone when a wrong





**Figure 14.** Convergence of the posterior predictions, as measured by different scores (columns) for every method (rows). Each plot compares the curves for different TIs.

prior was used. It resulted in overestimated confidence interval for the well data.

DREAM-ZS often matched the data the best and provided very good protection zone estimates. It was able to generate realizations that had bifurcation patterns without seeing them in the training set. Outside the informed zone, it was often overconfident in placing geological patterns, even to a greater extent than ESMDA. However, the good performance of DREAM-ZS is remarkable, as the reference data was generated using the MPS tool employed by PoPEX

and ESMDA. Identifying these geometries should be easier for these methods than for DREAM-ZS. For a fairer comparison, an additional reference realization generated with GAN could be added, but it would require repeating the whole study.

Note that, this study might be extended by including other methods coupling ESMDA with GAN [Bao *et al.*, 2020, Canchumuni *et al.*, 2021], which would be interesting to compare especially with ESMDA + MPS and DREAM-ZS + GAN. Another point is that the results are compared to a single reference field,

and not to a probabilistic reference solution, as for example done by Jägglı et al. [2017]. However, such a probabilistic reference solution requires a very large ensemble of unconditional realizations and depends on the simulation technique and the prior. Since we are modifying the prior in our tests, it would be necessary to consider several probabilistic reference solutions to make the comparisons, implying an even higher computational cost.

More generally, the application of these methods to real case studies still needs to be explored before we could give recommendations to the practitioners. First, one needs to extend the methods to the multi-categorical case. PoPEX is capable of handling it, adapting GAN should be straightforward, but ESMDA + MPS must be adapted to account for multiple Gaussian pyramids. In a practical application, identifying the prior or priors requires geological exploration and testing the different priors using for example K-fold cross-validation strategies as shown in Juda et al. [2020]. This is feasible in theory, but may not always be possible because of limited time. Another important unknown is how these different methods would behave when adding borehole data. Making such a comparison was not possible, because the WGAN that we used is not able to generate geological simulations conditioned by borehole data. Finding ways to condition efficiently the GAN is still a research topic. We also did not evaluate how the different methods perform when the quantity of information varies. All these questions show that research on categorical inverse problems is still very open.

To conclude, the approaches presented and compared in this paper include a statistical and spatial model of a categorical field and a stochastic technique for generating or identifying a set of realizations that could match the head data. This is directly inspired by the philosophy that Ghislain de Marsily taught us (see companion paper by White and Lavenu [2022]). We are still pursuing this track with passion, but the road is bumpy and the destination is far from sight.

### Conflicts of interest

The authors declare no competing financial interest.

### Acknowledgments

The authors thank Augustin Gouy for the preliminary study that he conducted and for sharing his synthetic inversion set-up, which inspired the test case presented in this article.

### References

- Alcolea, A. and Renard, P. (2010). Blocking moving window algorithm: Conditioning multiple-point simulations to hydrogeological data. *Water Resour. Res.*, 46(8), article no. W08511.
- Arjovsky, M., Chintala, S., and Bottou, L. (2017). Wasserstein generative adversarial networks. In *Proceedings of the 34th International Conference on Machine Learning*, pages 214–223. PMLR, Cambridge, MA, <https://proceedings.mlr.press/v70/arjovsky17a.html>.
- Bakker, M., Post, V., Langevin, C. D., Hughes, J. D., White, J. T., Starn, J. J., and Fienen, M. N. (2016). Scripting MODFLOW model development using python and FloPy. *Groundwater*, 54(5), 733–739.
- Bao, J., Li, L., and Redoloza, F. (2020). Coupling ensemble smoother and deep learning with generative adversarial networks to deal with non-gaussianity in flow and transport data assimilation. *J. Hydrol.*, 590, article no. 125443.
- Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Rev.*, 78(1), 1–3.
- Caers, J. and Hoffman, T. (2006). The probability perturbation method: A new look at Bayesian inverse modeling. *Math. Geol.*, 38(1), 81–100.
- Canchumuni, S. W. A., Castro, J. D. B., Potratz, J., Emerick, A. A., and Pacheco, M. A. C. (2021). Recent developments combining ensemble smoother and deep generative networks for facies history matching. *Comput. Geosci.*, 25(1), 433–466.
- Dagasan, Y., Juda, P., and Renard, P. (2020). Using generative adversarial networks as a fast forward operator for hydrogeological inverse problems. *Groundwater*, 58(6), 938–950.
- de Marsily, G., Delay, F., Gonçalvès, J., Renard, P., Teles, V., and Violette, S. (2005). Dealing with spatial heterogeneity. *Hydrogeol. J.*, 13(1), 161–183.
- de Marsily, G., Delhomme, J.-P., Delay, F., and Buoro, A. (1999). Regards sur 40 ans de problèmes inverses

- en hydrogéologie. *C. R. Acad. Sci. Ser. IIA Earth Planet. Sci.*, 329(2), 73–87.
- de Marsily, G., Lavedan, G., Boucher, M., and Fasamino, G. (1984). Interpretation of interference tests in a well field using geostatistical techniques to fit the permeability distribution in a reservoir model. In *Geostatistics for Natural Resources Characterization*, pages 831–849. D. Reidel Pub. Co., Dordrecht.
- Emerick, A. A. (2016). Analysis of the performance of ensemble-based assimilation of production and seismic data. *J. Pet. Sci. Eng.*, 139, 219–239.
- Emerick, A. A. and Reynolds, A. C. (2012). History matching time-lapse seismic data using the ensemble Kalman filter with multiple data assimilations. *Comput. Geosci.*, 16(3), 639–659.
- Emerick, A. A. and Reynolds, A. C. (2013). Ensemble smoother with multiple data assimilation. *Comput. Geosci.*, 55, 3–15.
- Evensen, G. (2009). *Data Assimilation: The Ensemble Kalman Filter*. Springer, Berlin, Heidelberg.
- Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Stat. Sci.*, 7(4), 457–472.
- Gneiting, T., Balabdaoui, F., and Raftery, A. E. (2007). Probabilistic forecasts, calibration and sharpness. *J. R. Stat. Soc.: Ser. B (Stat. Methodol.)*, 69(2), 243–268.
- Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *J. Am. Stat. Assoc.*, 102(477), 359–378.
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press, Cambridge, MA.
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial nets. In *Proceedings of the 27th International Conference on Neural Information Processing Systems – Volume 2, NIPS’14*, pages 2672–2680. MIT Press, Cambridge, MA.
- Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., and Courville, A. C. (2017). Improved Training of Wasserstein GANs. In Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R., editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc, <https://proceedings.neurips.cc/paper/2017/file/892c3b1c6dccc52936e27cbd0ff683d6-Paper.pdf>.
- Hansen, T. M., Cordua, K. S., and Mosegaard, K. (2012). Inverse problems with non-trivial priors: Efficient solution through sequential Gibbs sampling. *Comput. Geosci.*, 16(3), 593–611.
- Hendricks Franssen, H.-J., Alcolea, A., Riva, M., Bakr, M., Van der Wiel, N., Stauffer, F., and Guadagnini, A. (2009). A comparison of seven methods for the inverse modelling of groundwater flow. Application to the characterisation of well catchments. *Adv. Water Resour.*, 32(6), 851–872.
- Hughes, J. D., Langevin, C. D., and Banta, E. R. (2017). Documentation for the MODFLOW 6 framework. In *Techniques and Methods Report 6-A57*. U.S. Department of the Interior, U.S. Geological Survey, Reston, VA.
- Jäggli, C., Straubhaar, J., and Renard, P. (2017). Posterior population expansion for solving inverse problems. *Water Resour. Res.*, 53(4), 2902–2916.
- Jäggli, C., Straubhaar, J., and Renard, P. (2018). Parallelized adaptive importance sampling for solving inverse problems. *Front. Earth Sci.*, 6, article no. 203.
- Journel, A. and Zhang, T. (2006). The necessity of a multiple-point prior model. *Math. Geol.*, 38(5), 591–610.
- Juda, P. and Renard, P. (2021). An attempt to boost posterior population expansion using fast machine learning algorithms. *Front. Artif. Intell.*, 4, article no. 624629.
- Juda, P., Renard, P., and Straubhaar, J. (2020). A framework for the cross-validation of categorical geostatistical simulations. *Earth Space Sci.*, 7(8), article no. e2020EA001152.
- Kang, X., Shi, X., Revil, A., Cao, Z., Li, L., Lan, T., and Wu, J. (2019). Coupled hydrogeophysical inversion to identify non-Gaussian hydraulic conductivity field by jointly assimilating geochemical and time-lapse geophysical data. *J. Hydrol.*, 578, article no. 124092.
- Laloy, E., Héroult, R., Jacques, D., and Linde, N. (2018). Training-image based geostatistical inversion using a spatial generative adversarial neural network. *Water Resour. Res.*, 54(1), 381–406.
- Laloy, E., Héroult, R., Lee, J., Jacques, D., and Linde, N. (2017). Inversion using a new low-dimensional representation of complex binary geological media based on a deep neural network. *Adv. Water Resour.*, 110, 387–405.
- Laloy, E. and Vrugt, J. A. (2012). High-dimensional

- posterior exploration of hydrologic models using multiple-try DREAM(ZS) and high-performance computing. *Water Resour. Res.*, 48(1), article no. W01526.
- Lam, D. T. (2019). *Conditioning Groundwater Flow Parameters with Iterative Ensemble Smoothers: Analysis and Approaches in the Continuous and the Discrete Cases*. PhD thesis, Université de Neuchâtel, Faculté des Sciences, Hydrogéologie et géothermie.
- Lam, D.-T., Renard, P., Straubhaar, J., and Kerrou, J. (2020). Multiresolution approach to condition categorical multiple-point realizations to dynamic data with iterative ensemble smoothing. *Water Resour. Res.*, 56(2), article no. e2019WR025875.
- Linde, N., Renard, P., Mukerji, T., and Caers, J. (2015). Geological realism in hydrogeological and geophysical inverse modeling: A review. *Adv. Water Resour.*, 86, 86–101.
- Maharaja, A. (2008). TiGenerator: Object-based training image generator. *Comput. Geosci.*, 34(12), 1753–1761.
- Mariethoz, G., Renard, P., and Caers, J. (2010). Bayesian inverse problem and optimization with iterative spatial resampling. *Water Resour. Res.*, 46(11), article no. W11530.
- Oliver, D. S. and Chen, Y. (2011). Recent progress on reservoir history matching: a review. *Comput. Geosci.*, 15(1), 185–221.
- Oliver, D. S. and Chen, Y. (2018). Data assimilation in truncated plurigaussian models: Impact of the truncation map. *Math. Geosci.*, 50(8), 867–893.
- Straubhaar, J., Renard, P., and Chugunova, T. (2020). Multiple-point statistics using multi-resolution images. *Stoch. Environ. Res. Risk Assess.*, 34, 251–273.
- Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM Society for Industrial and Applied Mathematics, Philadelphia.
- ter Braak, C. J. F. and Vrugt, J. A. (2008). Differential Evolution Markov Chain with snooker updater and fewer chains. *Stat. Comput.*, 18(4), 435–446.
- van Leeuwen, M., te Stroet, C. B. M., Butler, A. P., and Tompkins, J. A. (1998). Stochastic determination of well capture zones. *Water Resour. Res.*, 34(9), 2215–2223.
- Vrugt, J. A., ter Braak, C., Diks, C., Robinson, B. A., Hyman, J. M., and Higdon, D. (2009). Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. *Int. J. Nonlinear Sci. Numer. Simul.*, 10(3), 273–290.
- White, J. T. and Lavenue, A. M. (2022). Advances in the pilot point inverse method: OÙ en sommes-nous maintenant? *C. R. Geosc.* Forthcoming.
- Zahner, T., Lochbühler, T., Mariethoz, G., and Linde, N. (2016). Image synthesis with graph cuts: A fast model proposal mechanism in probabilistic inversion. *Geophys. J. Int.*, 204(2), 1179–1190.
- Zhou, H., Gómez-Hernández, J. J., Hendricks Franssen, H.-J., and Li, L. (2011). An approach to handling non-Gaussianity of parameters and state variables in ensemble Kalman filtering. *Adv. Water Resour.*, 34(7), 844–864.
- Zhou, H., Gómez-Hernández, J. J., and Li, L. (2014). Inverse methods in hydrogeology: Evolution and recent trends. *Ad. Water Resour.*, 63, 22–37.
- Zimmerman, D., De Marsily, G., Gotway, C. A., Marietta, M. G., Axness, C. L., Beauheim, R. L., Bras, R. L., Carrera, J., Dagan, G., Davies, P. B., et al. (1998). A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow. *Water Resour. Res.*, 34(6), 1373–1413.



Research article

Geo-hydrological Data & Models

# Identification of aquifer heterogeneity through inverse methods

Philippe Ackerer<sup>Ⓢ,\*</sup>, Jesus Carrera<sup>Ⓢ,b</sup> and Frédéric Delay<sup>Ⓢ,a</sup>

<sup>a</sup> Institut Terre et Environnement de Strasbourg, Université de Strasbourg, CNRS, ENGEES, UMR 7063, Strasbourg, France

<sup>b</sup> Institute of Environmental Assessment and Water Research (IDAEA), CSIC, Spain

E-mails: ackerer@unistra.fr (P. Ackerer), jesus.carrera.ramirez@gmail.com (J. Carrera), fdelay@unistra.fr (F. Delay)

**Abstract.** The paper underlines the contributions of Ghislain de Marsily (GdM) to the identification of aquifers heterogeneity using inverse methods mainly for modeling subsurface flow. Inverse methods require an objective function to express the goodness of fit of the chosen model, a parameterization to describe the spatial distribution of model parameters, and a minimization algorithm. The resulting inverse problem, which consists in seeking model parameters' values that render model outputs close to the observations, is usually unstable. GdM developed seminal ideas for the two key inversion issues that are: to stabilize the inverse problem through regularization, and to parameterize it to reproduce the natural heterogeneity of the subsurface with a limited number of parameters. GdM conducted pioneering works that are the basis of current parameterization methods relying upon adaptive zonation and/or interpolation based on pilot points. We take here the opportunity to highlight the GdM's contributions inspiring currently used techniques.

**Keywords.** Groundwater, Inverse methods, Modelling, Parameter identification, Parameterization.

**Quotations.** "Verification will certainly never be perfect, but again we are not striving for certainty and perfection, only to do our level best". [de Marsily et al., 1992].

*Manuscript received 19 April 2022, revised 13 July 2022, accepted 12 September 2022.*

## 1. Introduction

As soon as anyone is tempted to find the causes of some observations, an inverse problem emerges. We face inverse problems daily and most of the time, without being fully aware of this. A quite general definition of an inverse problem is finding the causes of observations, while the direct (or the forward) problem would be to assess the effects of some causes. One of the most popular successes in inverse problems is probably the discovery of Neptune by Le Verrier [1846]. His manuscript started with "*Je me pro-*

*pose ... d'étudier la nature des irrégularités du mouvement d'Uranus; et de remonter à leurs causes*" (I propose ... to study the nature of the irregularities of the movement of Uranus, and to retrieve their causes). Based on observations of unexplained trajectories of Uranus, Le Verrier applied two inverse methods: (i) a parameter estimation trying to find the best parameter set for the mathematical model calculating the interactions between Uranus, Saturn, Jupiter, and the Sun; (ii) a change in the model by adding a new planet (Neptune) and finding its position and mass. He also questioned the uncertainty of existing data and their potential impact on his conclusions. In his quest to explain why Uranus had irregular unexplained trajectories, he expressed the two basic is-

\* Corresponding author.

sues arising when formulating an inverse problem: to find parameters for a given model (parameter estimation), and to find both the model and its parameters (model identification).

Most inverse problems are considered as ill-posed in that they contradict, to some extent, the principles of a well-posed problem in the sense of Hadamard [1902]:

- for all admissible data, a solution exists;
- for all admissible data, the solution is unique;
- the solution depends continuously on data.

It can be noticed that these criteria lack mathematical rigor. What is “admissible”? What kind of continuity is required? Even worse, in Geosciences, a unique and stable solution does not need to be a good representation of reality. This might be one of the reasons why inverse methods have been mostly initiated by intuitive reasoning rather than grounded in mathematically sound deductions. In fact, an ill-posed problem should lead to its reformulation. However, some problems can be considered as “morally ill-posed” [Jaynes, 1984] due to practical difficulties such as determining the subsurface structure from surface seismic data, or because of small sensitivities such as determining the transmissivity variability from piezometric heads in groundwater modeling, or because of the complexity (chaotic character, i.e., instability) of interactions between individuals or groups such as the mechanics of billiard balls. Many ill-posed problems are thoroughly described in Bertrand [1889]. In hydrogeology, we know that a solution usually exists, but whether it depends continuously on data or is unique has been the topic of many research activities over the last seventy years. The outcome was that the well-posed character of a problem greatly depends on how the problem is formulated.

The first attempt of parameter identification through an inverse method in hydrogeology can be attributed to Bennett and Meyzer [1952] who used a flow net approach to estimate transmissivity. Stallman [1956] developed the first numerical model for parameter estimation by relying upon this flow net approach. This type of approach continued during the 1960s, until Emsellem and de Marsily [1971] showed that the method led to non-unique (arbitrary, as we will see) and unstable solutions. They realized that a well-posed problem was needed for reg-

ularization, and that extra information was required to render the solution realistic. These issues were formalized in the GdM’s “thèse d’Etat” [1978], which led to formulating the inverse problem as an optimization procedure minimizing an objective function that penalizes discrepancies between observations and model outputs. However, regularization was still needed.

The goal of regularization is to improve the conditioning of the inverse problem. Numerous approaches can be adopted. But, to add realism beyond conditioning, they should integrate prior information (i.e., a knowledge about model parameters stemming from sources other than head data). The incorporation of prior information was initially introduced by restricting the size of the domain in the parameter space where the solution was supposed to be (i.e., constraints on the parameter values). However, this led to parameters fluctuating spatially between their bounds, which inclined Neuman [1973] to propose an additional “plausibility criterion”. Tikhonov [1963b,a] had shown that by simply adding the quadratic norm of the parameters to the objective function, linear regression could become well-posed. Even though this added criterion did not include prior information, it opened the way of using metrics accounting for priors. Note that strict Tikhonov regularization simply promotes smoothness. Carrera and Neuman [1986b] argue that this limitation is fictitious because Tikhonov’s regularization can be used with any appropriate formulation i.e., Tikhonov’s results are valid if the regularization term penalizes any norm of the parameter departures with respect to the prior information. See Engl *et al.* [1996] for a general overview of Tikhonov regularization.

Part of the instability problem results from trying to estimate too many parameters from a limited amount of data. Therefore, another line of research has been trying to represent permeability fields with a limited number of unknown parameters. This is called parameterization, which can also be seen as a particular regularization method, in the sense that it improves the conditioning of the inverse problem by reducing the number of degrees of freedom (i.e., the number of parameters to estimate). In fact, an adequate parameterization allows for incorporating prior information, both in terms of parameter values and their statistical and spatial distributions. Parameterization can compile “hard” informa-

tion such as measured parameter values and subjective knowledge (“soft” information) coming from the investigations of different disciplines including geology, geophysics, and hydrology. This “soft” information may also be the result of field works (geologic maps, drilling logs, pumping tests, geophysical surveys) that may provide a reasonably good insight into the acceptable range of parameters and their spatial distribution. Again, words like “reasonably” and “acceptable” cover subjective notions which allow for a lot of creativity. As we will see, this may have been the most lasting contribution of GdM.

The aim of this paper is not to review parameter estimation by inverse methods. Numerous excellent reviews targeting the inverse problem in hydrogeology can be found in Yeh [1986], Carrera [1988], Ginn and Cushman [1990], de Marsily *et al.* [2000], Carrera *et al.* [2005], McLaughlin and Townley [1996], Kitaniidis [2007], Oliver and Chen [2010], Zhou *et al.* [2014], Yeh [2015]. Instead, our goal is to highlight the contributions of GdM in solving the above-mentioned problems (proper formulation, regularization, and parameterization), which are keys for model calibration and predictive uncertainty [Cui *et al.*, 2021].

The paper is organized along two lines. First, we outline the evolution of the formulation of the problem. Second, we review the four main parameterization strategies: zonation, interpolation such as pilot points, embedding geophysical information, and lithological information from modeling. For each strategy, we start from GdM’s work and provide a short review of how his ideas percolated until today. For simplicity (for the readers and the authors), we will only focus on the identification of hydraulic transmissivity or conductivity.

## 2. Formulation of the inverse problem: early attempts and stability issues

Early modelers of groundwater flow recognized the need for inversion [Stallman, 1956]. The solution to the inverse problem was initially posed as obtaining transmissivity, given hydraulic heads, assumed as known everywhere. In the steady-state case, this leads to a Cauchy problem [Nelson, 1960], as outlined in Figure 1. As mentioned in the introduction, Hadamard [1902] had stated that, for a problem to be well-posed, its solution should exist, be unique and stable. Emsellem and de Marsily [1971] showed in a

very simple way that the last two conditions might not be met in the estimation of transmissivities by considering a simple steady-state flow problem, governed by

$$\nabla(T\nabla h) + q = 0 \quad (1)$$

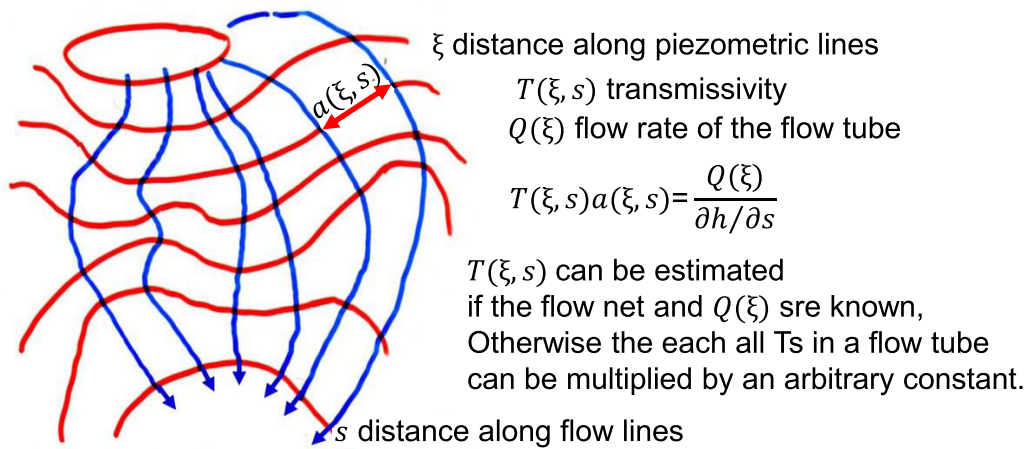
where  $T = T(\mathbf{x})$  and  $q = q(\mathbf{x})$  ( $\mathbf{x}$  a given location) are the transmissivity and recharge fields, respectively, and  $h = h(\mathbf{x})$  is the head, solution of the flow problem, if subjected to appropriate boundary conditions. Then, if a field  $T'$  can be found satisfying

$$\nabla(T'\nabla h) = 0 \quad (2)$$

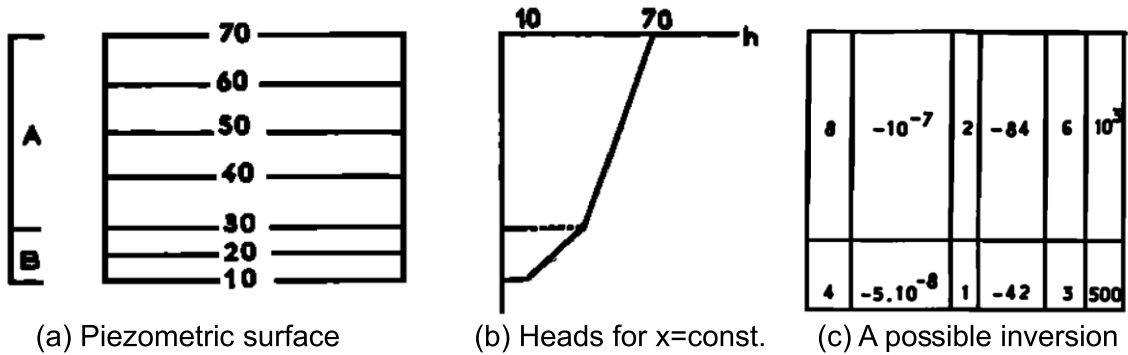
under homogeneous boundary conditions, then  $T + T'$  will also satisfy (1). For transient problems, finding  $T'$  may be difficult, but for steady-state problems it is not. Therefore, the solution to the problem is non-unique. Worse, as shown in Figure 1, the solution obtained with the initial Cauchy method is arbitrary. The calculated solution is defined up to the multiplication by a constant that can be chosen arbitrarily for each integration flow tube. This leads to “odd” solutions, like the one in Figure 2.

To overcome these difficulties, Emsellem and de Marsily [1971] raised and then did three things that would mark future developments:

- The need for regularization. Since the solution can lead to arbitrary jumps, one should adopt a “smoothing” criterion to reduce arbitrariness, while ensuring existence and stability of the solution (i.e., regularization).
- The need for data other than heads to ensure resemblance to the actual aquifer i.e., prior information and adequate parameterization.
- To propose a general solution method based on minimizing mass balance errors when selecting the  $T$  field. While their method still required a full knowledge of heads, it was generalized in that it could be applied to transient problems with internal sinks and sources (even if uncertain). The full knowledge of heads may sound arbitrary nowadays, but one must bear in mind that automatic interpolation was not a standard tool in the 1960s. Therefore, hydrogeologists were used to draw piezometric surfaces with hydrogeological criteria, so that a good deal of conceptual understanding went into these maps.



**Figure 1.** Early formulation of the inverse problem as integration of the flow equation along streamlines (picture taken from a lecture of GdM on the topic).



**Figure 2.** Example of Emsellem and de Marsily [1971] to illustrate the non-uniqueness of the inverse problem: (a) a medium with homogeneous  $T$  zones A and B (A twice as transmissive as B), will yield a piezometric surface with (b) twice head gradient in B than in A. Any  $T$  field as in (c), with along vertical lines  $T_A = 2T_B$ , will reproduce heads exactly [modified from Emsellem and de Marsily, 1971].

These concepts marked future developments. The need for regularization was explicitly incorporated by Neuman [1973], who did two things. First, he formulated the problem in terms of head errors (minimizing the sum of squared errors on heads), which he termed the indirect formulation of the inverse problem [see also Yeh, 1975]. Second, he formalized the smoothing criterion so that inversion became a multi-objective problem, with a model error criterion (e.g., the sum of squared errors) and a plausibility criterion incorporating prior information. The work of many others followed this path. Statistical formulations of the inverse problem were the immediate next step [Neuman, 1980, Kitanidis and Vomvoris, 1983,

Carrera and Neuman, 1986a, Rubin and Dagan, 1987, and many others]. These formulations allowed formalizing the regularization requirement as a statistical problem. The issues of existence and identifiability were formalized by Carrera and Neuman [1986b].

Still, the instability problem remained. To the point that the inverse problem of groundwater hydrology was purported to be intrinsically unstable [Yakowitz and Duckstein, 1980], i.e., a “morally ill-posed” problem. A large part of the problem rested on trying to estimate  $T$  everywhere. The discussions above (and the examples in Figures 1 and 2) make it clear that trying to estimate  $T(\mathbf{x})$  for all locations  $\mathbf{x}$  within the domain will lead to an ill-posed



problem, even if  $h(\mathbf{x})$  is accurately known everywhere. Since flow through permeable media is essentially dissipative, heads do not contain sufficient information about the small scale variability patterns of hydraulic conductivity. This feature motivated regularization methods, which initially were seeking smoothness. Obviously, nature is not necessarily smooth and a smooth solution does not need to resemble natural variability, which may become essential to other problems (notably, solute transport). Additional information on the nature of spatial variability is needed. Therefore, Emsellem and de Marsily [1971] sought expressing  $T(\mathbf{x})$  in terms of a number of parameters, while resembling natural variability. Note that, in doing so, the inverse problem moves from parameter estimation to model identification.

### 3. Parameterization

The problem of expressing the complete transmissivity field in terms of a (hopefully small) number of scalars is termed parameterization. The immediate option is to divide the aquifer into “zones” based on geological understanding. And this is still a basic option. However, this approach encounters numerous difficulties, from both theoretical and practical points of view. Geology is usually ambiguous, so it is not easy to define geological units accurately. Even if this was possible, geological units are not homogeneous.

A proper framework to study spatial variability is geostatistics, the study of random fields, which was introduced by Matheron [1967]. Two geostatistical tools are relevant to our problem. The first one, kriging, solves the problem of estimating a random function at a location, given measurements elsewhere. Since  $T(\mathbf{x})$  is generally assumed to be log-normally distributed, it is best to work with  $Y(\mathbf{x}) = \ln T(\mathbf{x})$ . Then, the estimate of  $Y(\mathbf{x})$  at a location  $\mathbf{x}$ , assuming that its values  $Y_i$  are known at a set of measurement points  $\mathbf{x}_i$  is:

$$Y_K(\mathbf{x}) = \sum_{i=1}^N \lambda_i(\mathbf{x}, \mathbf{x}_i) Y_i \quad (3)$$

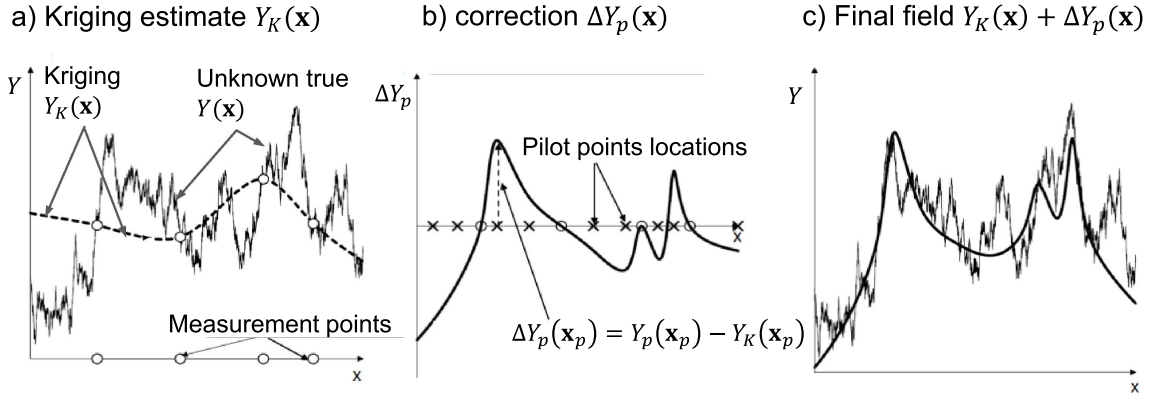
where the subscript  $K$  stands for kriging estimate, and the “kriging weights”,  $\lambda_i$ , are usually obtained by ensuring that  $Y_K$  is a minimum variance unbiased estimate of  $Y(\mathbf{x})$ . The weights can also be viewed as the Gaussian conditional estimation weights, which

is convenient for generalizing. A problem with conditional estimation is that it yields unrealistically smooth fields (Figure 3). A second geostatistical tool is conditional simulation, whose formulation is similar to the kriging problem, except that instead of seeking the “best” estimate, it seeks a random field that is statistically consistent with available data. These tools are basics for most parameterization methods, which are reviewed below.

#### 3.1. Zonation parameterization

Zonation was the first parameterization strategy used for parameter estimation. It consists in partitioning the complete domain into zones where the parameter variability can be described by the same model (uniform value, statistical properties, etc.). The most common zonation for parameters is based on uniform values over a given subarea, but varying between diverse subareas. Pioneer works on zonation can be found in Jacquard and Jain [1965], and Jahns [1966]. In the work by Jahns, the inverse problem was solved using a regression problem (minimization of an objective function with a Gauss–Newton algorithm) based on a cost function as the quadratic norm of the vector of residuals (i.e., the squared difference between computed and measured heads). The number of zones was increased during the identification procedure and the reliability of the parameters was discussed in detail, addressing non-uniqueness due to Lorentz reciprocity in flow [see, e.g., Delay *et al.*, 2011, Marinoni *et al.*, 2016] and correlations between parameters. Jahns [1966] provided also computation times: 20 min for the estimation of 10 parameters with a  $20 \times 20$  grid and 25 time steps.

Emsellem and de Marsily [1971] suggested adding a regularization term to the objective function by introducing a “smoothing” term in the minimization procedure to reduce the number of possible solutions. Adaptive parameterization was also described by successive refinements of the initial zonation. To our knowledge, for the first time, parameterization was embedded in the minimization procedure by adding to the cost function the quadratic norm of the transmissivity vector multiplied by a weighting coefficient. In this way, the parameterization was automatically accounted for in the inversion through the minimization procedure. Zonation can also be different for the same inverse problem, according to the



**Figure 3.** Schematic description on the pilot points method (PPM). (a) Kriging honors measurements but not variability away from measurements. (b) The PPM can be viewed as a correction of kriging estimation that can yield (c) adequate large scale variability (conditional simulation, instead of kriging, can be used for small scale variability).

type of sought parameters. For example, Zhang *et al.* [2014] proposed two different zonations for transmissivity and for elastic and inelastic specific storage parameters.

The main drawback of a raw zonation technique is the prior definition of the number of zones and their shape. The number of zones can be increased stepwise as suggested by Jahns [1966] and Emsellem and de Marsily [1971]. Carrera and Neuman [1986a], and Sun *et al.* [1998] also suggested some metrics to estimate the potential benefits of increasing the number of zones. These metrics were also used by Tung and Tan [2005] with a zonation based on Voronoï diagrams.

Several methods have been developed during the last twenty years to adapt the shape of the zones and the parameter values. Ben Ameer *et al.* [2002], Grimstad *et al.* [2003], and Hayek and Ackerer [2006] suggested the calculation of refinement indicators for the definition of the shape and the number of zones. Transmissivity is assumed to be a piecewise constant space function and unknowns are both the transmissivity values and the shape of the zones. The shape of each zone is adapted by relying upon refinement indicators easily computed from the gradient of the cost function. These refinement indicators define where to split a zone, rendering then the largest reduction in the cost function after the corresponding estimation. Level-set corrections were suggested by Lu and Robinson [2006] and Berre *et al.* [2007], for example, to deform the boundaries between zones

to obtain a better zonation. The local deformation of the boundaries is proportional to the permeability contrast between the two sides of a boundary, the sensitivity of heads to transmissivity, and the residual between the simulated and observed heads.

### 3.2. Pilot points parametrization

Kriging represents the BLUE (Best Linear Unbiased Estimator) of  $Y(\mathbf{x})$ . Therefore, a natural parameterization consists of extending the kriging equations (3) beyond the original set of measurement points,  $\mathbf{x}_i$ , by adding a set of pilot points,  $\mathbf{x}_p$ . The kriging estimate becomes:

$$Y_K(\mathbf{x}) = \sum_{i=1}^N \lambda_i(\mathbf{x}, \mathbf{x}_i) Y_i + \sum_{p=1}^{N_p} \lambda_p(\mathbf{x}, \mathbf{x}_p) Y_p \quad (4)$$

where  $\lambda_p$  are the kriging weights of the pilot points computed at any location  $\mathbf{x}$ . Obviously,  $Y_p$  are unknown. The pilot points method [PPM, de Marsily, 1978] consists of estimating  $Y_p$  so as to obtain an optimum fit of simulated heads to head measurements. The PPM can also be viewed as adding a correction to the kriging estimate, which is more convenient computationally (Figure 3). Obviously, the original form of the method suffered several problems. For instance, the location of the pilot points was arbitrary and had to be chosen by trial and error [de Marsily *et al.*, 1984]. This was overcome by Ramarao *et al.* [1995], who sequentially increased the

number of pilot points by adding new ones in the locations where they yielded the best improvement to the objective function.

A second difficulty in the original form was that pilot point estimates were not regularized, which caused instability and led modelers to emphasize both a “prudent” increase in the number of pilot points, and to bound the values of estimated  $Y_p$ . Doherty [2003] included a regularization criterion penalizing non-homogeneity of the model parameters, while not using prior information [see details by Doherty *et al.*, 2010]. Neglecting prior information results from viewing regularization in a “strict” Tikhonov sense (recall discussion on Tikhonov regularization in the introduction). As a consequence, limiting the number of pilot points remained a barrier and, for a while, it seemed that the method of choice was the self-calibration approach of Capilla *et al.* [1998] and Gómez-Hernández *et al.* [1997], which was able to reproduce quite accurately complex fields. The problem with the traditional formulation of the PPM was that direct measurements were used for the first term of (4), but were disregarded during the inversion process. Alcolea *et al.* [2006] overcame this problem by formulating the PPM in a fully geostatistical form. They explicitly acknowledged that pilot point estimates were part of the random field, which led naturally to a regularization term without any additional assumption. This formulation also allowed demonstrating that a stable inversion was possible, regardless of the number of pilot points, provided that a rigorous statistical formulation was adopted. In fact, Kitanidis and Vomvoris [1983] had already shown, via linearized co-kriging, that the number of parameters is not the real issue, but consistent geostatistics is. An additional advantage of the geostatistical formulation of the PPM is that it leads naturally to conditional simulation techniques, which is advantageous when one is interested in reproducing small scale variability [Lavenue *et al.*, 1995].

A great advantage of the PPM is the ease with which complex problems can be addressed, which coupled with the flexibility of PEST software [Doherty *et al.*, 2010], has allowed an explosion of the method in several directions. The PPM has been used with all kinds of equations and phenomena beyond groundwater flow [e.g., Vesselinov *et al.*, 2001]. It has also been used in moments equation inversion

[Hernandez *et al.*, 2003], for fractured media [Lavenue and De Marsily, 2001], or for multi-point geostatistics [Gravey and Mariethoz, 2020, Ma and Jafarpour, 2018]. In short, the PPM has become the standard for non-linear geostatistical inversion. More details about PPM can be found in White and Lavenue [2022].

Nevertheless, a practical problem with the traditional PPM lies on its reliance on the geostatistical assumptions. It is clear that the PPM can lead to fully consistent and stable solutions. The question is whether the usual geostatistical assumptions provide good representations of geological variability. For example, the traditional multi-Gaussian paradigm was challenged by Gómez-Hernández and Wen [1998], who suggested that high transmissivity regions tend to be connected, which has repeatedly proven right. For example, Pool *et al.* [2015] compared the PPM (assuming stationary multi-Gaussian) to zonation on the basis of a simple geological model (paleo channels connectivity). They found that, while the PPM led to better calibrations (smaller calibration errors), geological zonation was better at predicting seawater intrusion. The implication is that traditional geostatistics is not sufficient and that proper incorporation of geological understanding is needed.

### 3.3. *Parameterization based on geophysics*

The use of geophysical methods such as electrical resistance tomography (ERT), seismic and radar transmission are expected to provide complementary data related to both the hydraulic parameter values in the subsurface and their spatial distribution. Measurements are done from the soil surface and are cheaper than traditional hydrological measurements such as pumping tests, which require “intruding” in the subsurface via wellbores. Furthermore, the acquisition of geophysical signals can be fairly well extended over the whole aquifer, which may prove helpful in tackling two main problems plaguing parameter identification for hydrogeological models:

- The representative volume of the measured transmissivity. This volume ranges from laboratory scale to pumping tests scale in hydrology. It is rarely consistent with the elementary modeling scale at which data should be estimated to document the cell/element scale of the model grid;

- The improvement of parameterization. The geophysical investigations should provide numerous images/patterns/values of geophysical parameters, assumed to be somehow correlated to hydrological parameters.

In short, the additional information from geophysics is expected to reduce the number of possible solutions, or at least permit sorting the “hydraulic” solutions that are in good stands with the subsurface structure evidenced by geophysics.

De Marsily and co-workers nicely highlighted the contribution of electric resistance measurements to the estimation of transmissivity. Ahmed *et al.* [1988] used the co-kriging of measured transmissivity, specific capacity, and electrical resistivity to elaborate transmissivity maps. They underlined that the contribution of electrical resistivity to transmissivity evaluation was important, but the diverse variables should be measured at a significant number of common locations to infer reliable cross-variograms.

The use of geophysical data in hydraulic parameter estimation was also addressed by Rubin *et al.* [1992], who combined head and hydraulic conductivity measurements at wells with a well-known seismic velocity field. Following the same approach, Copty *et al.* [1993] analyzed, in hydraulic parameter estimations, the effects of measurement errors concealed in the seismic velocity values. In their parameter estimation procedure, Dam and Christensen [2003] included the parameters involved in the relationships (state equations) between hydraulic conductivity and geophysical properties. They also used PPM for parameterization of both geophysical and hydrological parameters.

For these three approaches, the relationships between geophysical data and hydraulic conductivity were supposed to be known, although these relationships can be complex, often highly non-linear, and varying over space. Hyndman *et al.* [1994] circumvented this downside by using seismic data to delineate the geometry of lithologic zones. The hydrodynamic parameter values were then estimated for each lithologic zone by minimizing the sum of squared residuals between measured and computed tracer concentrations.

Haber and Oldenburg [1997] developed the joint inversion as a generic approach to invert two data sets when the underlying models are linked by the same structural (geological) heterogeneity. The main

advantage of this approach is that it does not need any assumption about the relationship between the two data sets. Cardiff and Kitanidis [2009] showed the interest of the joint inversion via an adaptive zonation approach based on the level-set method. Finsterle and Kowalsky [2008] performed the joint inversion of ground penetrating radar (GPR) travel times and hydrological data collected during a simulated ponded infiltration experiment. Joint inversion is nowadays widely studied in hydrogeophysics [see Linde and Doetsch, 2016, for a review].

In the field of hydraulic parameter estimation partly relying on geophysical data, the work of de Marsily and co-workers can be considered as seminal in initiating a promising research field leading to a new discipline, hydrogeophysics.

### 3.4. *Parameterization based on lithological models*

Solving the groundwater flow inverse problem has also been applicable in pre-conditioning the inversion by prior guess on the structure of the heterogeneity in the subsurface. When the structure is that of the spatial distribution of hydraulic parameters, for example, a correlated random field, the preconditioning is similar to a regularization technique, prescribing an overall distribution of hydraulic conductivities. Inversion and post-conditioning onto hydrological data can be carried out to add a perturbation to the prior parameter field [e.g., Ramarao *et al.*, 1995].

However, there exist many geological contexts, in which a smooth field of parameters is not likely to represent the subsurface geological heterogeneity. In these contexts, the simulations of diverse “facies” distributions have revealed a better option. Geostatistical techniques contributed to this task by relying upon truncated Gaussian simulations [e.g., Matheron *et al.*, 1987] or indicator sequential simulations [e.g., Schafmeister and de Marsily, 1994]. In short, both techniques come down to calculate the probability of occurrence of a given geological facies at each location of the modeled domain. Nevertheless, the two-point covariance used to simulate the distribution of these probabilities and to avoid building a fully random “salt and pepper” image, still renders relatively continuous representations. These representations would not match the forecasts of

geologists for complex systems such as fracture fields or sedimentary facies distributions in the subsurface of river floodplains. However, it is worth noting that improving the geostatistical methods, especially in developing truncated multi-Gaussian simulations and non-stationary random functions, renders to this day more realistic images [e.g., Beucher and Renard, 2016].

In answer to the question of modeling with less pain the complex heterogeneity of the subsurface compartment in floodplains, G de Marsily and his co-workers elected models of sedimentation, in the form of a “genesis” model. Those are geared towards simulating the mechanisms that sequentially occur over time to build the floodplain. Tetzlaff and Harbaugh [1989] produced, via simplified equations of flow and transport, a model simulating river floodplains and deltas. This model was amended by Kolterman and Gorelick [1996], and employed to produce a reference work simulating the sedimentation over 600,000 years occurring at the east coast of the San Francisco bay. The exercise was conditioned by a detailed history of climate forcing and needed very heavy computations.

The genesis model built by de Marsily and co-workers [de Marsily *et al.*, 1998, Teles *et al.*, 2001, 2004] relies upon empirical rules to move, as random walkers over a grid, elementary parcels of sediments. The rules are inherited from the literature in fluvial geomorphology, and adapted as a function of the successive dynamic episodes that construct the floodplain, mainly: braided systems, meandering, and channel incision. These episodes, even though not well documented for precise hydrodynamic conditions, can be detected and dated along the history of the river by log samples and geomorphological considerations. Parcels of sediments are of regular parallelepiped shape, the size of which depends on the type of sediment conveyed. The variety of sediments encountered in floodplains is simplified into a few classes, e.g., gravels, sands, and loam, to keep some relative continuity in the deposited sedimentary bodies, and avoid complex images that sometimes could resemble patchworks.

The genesis model was used to reconstruct parts of sedimentary floodplains of the Rhône and the Aube rivers (France). In the application to the Aube river, the structure modeled by the genesis model was

compared to a model of facies build via the sequential simulation of indicators. For both models, facies were assigned hydraulic parameters values and forward simulations of flow and solute transport were performed. Regarding flow, both methods render very similar results in terms of hydraulic head distributions. It is worth noting that the flow scenario in the system is strongly conditioned by boundary conditions, and that piezometric heads are “robust” to parameters in the sense: it needs an important variation of hydraulic conductivity (or hydraulic diffusivity) to generate a small variation in heads. As both models, genesis and indicators, generate very different patterns of hydraulic conductivity distributions, the results pose the question of the identification of conductivities on the basis of heads only. Regarding transport results, simulations from both models do not match up at all. Transport (advection), is very sensitive to the flow patterns, both in terms of paths followed by the solute and concentration arrival times. The genesis model with its tortuous channels of high conductivity, guides transport through a few rapid pathways. For its part, the indicator model distributes more evenly the hydraulic conductivity values, with a consequence on transport of more widespread (diffuse) pathways with slower velocities.

In view of the above results, it goes without saying that parameterizations of the inverse problem inheriting from the geological structure of the subsurface are worth a try. Inferring or conjecturing this geological structure can be carried out via near subsurface geophysical investigations, process-imitating models (genesis models) and/or structure-imitating models (geomodels). At least, “geo-modeling” in its broad sense could open the equally probable solutions to an inverse flow problem, to solutions more convincing regarding the structure of the subsurface. To date, applications of geomodels find their way in engineering geology [Fookes *et al.*, 2015]. With regard to hydrogeology, regional aquifers in sedimentary basins are often targeted [e.g., Ross *et al.*, 2005], in almost the same way as for oil reservoirs or subsurface repositories. The geological model is usually aggregated at a scale rendering flow calculations tractable, and in this up-scaling process, the geological facies are directly transformed into a prior guess on hydraulic parameters. Along this line, the procedure looks like a pre-conditioning of inversion exercises on the basis of geological information.

It must be acknowledged that “genesis” models, including one developed by de Marsily and co-workers, did not receive much attention from the hydrological community. Nonetheless, modeling the floodplain construction continued to evolve, mainly for the purpose of simulating and predicting the geomorphological evolution of fluvial corridors [Williams *et al.*, 2016]. Models mimicking either meandering [e.g., Pittaluga *et al.*, 2009] or braided morphodynamics [e.g., Williams *et al.*, 2016] continue to rely upon both a physically-based approach, solving flow and transport [e.g., Sun *et al.*, 2015, Olson, 2021], or on conditioned empirical rules [for example, automata cellular, e.g., Coulthard *et al.*, 2007]. With the easier access to high-performance computing resources, the physically based models tend to overshadow models based on empirical rules.

#### 4. Conclusions

Inverse modeling is widely used today. Regularization, incorporation of prior information, and parameterization in inverse methods are an art required for philosophy and a lot of conjectures, i.e., mathematical statements that are accepted as valid, but whose validity have never been proven or disproven. These conjectures are supported by plausible reasoning [Polya, 1954] based on skills, training, and imitation. The most popular conjecture in inverse methods has often been: “*Simplex sigillum veri*”—simplicity is the sign of truth. This conjecture is now heavily discussed because inverse methods are nowadays not only developed for parameter estimation but also for improving model predictions by quantifying model uncertainty. With the increasing complexity of models, inversion strategies should focus not only on the number of parameters, but also on the smallest possible uncertainty, and in that case, the principle of parsimony may not hold [Gómez-Hernández, 2006].

When applying inverse methods for parameter identification, several problems come up:

- The existing data sets can be complex and include measures of different support volumes, of different nature, with data spanning a broad range of numerical values, and that are more or less “hard” or “soft”, more or less “certain” or “uncertain”. Properly handling these data becomes an important

step in the definition of the model concepts (physical processes taken into account, initial and boundary conditions, sink/source terms, etc.) and linked to the objectives of the model.

- The general design of the inverse procedure can be either deterministic or stochastic (Gaussian, multi-Gaussian) or both [Pool *et al.*, 2015].
- Additional information (also called multiple source information) on the forecasted parameter structure inheriting from geological and geophysical data should be used as model constraints. This structure may be adapted (automatically or not) during joint inversion.
- The parameterization should be consistent with the design and the information on the parameter structure. It should limit the number of parameters (degrees of freedom) to be identified by using discrete locations (nodes, pilot points, master points). These locations should be defined in an “appropriate way” and their number increased during the inversion procedure, while keeping in mind that there exist techniques and indicators limiting useless parameterization [Hassane *et al.*, 2017].
- Parsimony “as simple as possible, but not simplistic” is a good criterion to calibrate the model for specific situations i.e. when the variables lack sensitivity to the parameters (transmissivity estimation for a 2D flow model for example). However, it is not sufficient to address model uncertainty. Addressing model uncertainty, often quantified by parameter distributions and variances, needs the exploration of the ensemble of the plausible solutions [Moore and Doherty, 2005, Ackerer and Delay, 2010, Schöniger *et al.*, 2012].

Research activity tackling the above-mentioned key features has concentrated on how to address regularization and to include prior information as a way to reduce instability of the inverse problem, including: geostatistical or deterministic methods to address spatial variability, and variants of PPM to parameterize this variability. Parameterization is now conditioned by geophysical data and/or lithology

coming from field observations or lithology modeling. All these topics were introduced or addressed in the hydrogeological literature by Ghislain de Marsily. As in many other hydrogeological topics, he did not quite solve them (no one did it!), but, as “Tom Thumb”, he marked the path that many of us still continue to follow.

### Conflicts of interest

Authors have no conflict of interest to declare.

### Acknowledgments

Dr. Ackerer and Prof. Delay gratefully acknowledge the guidance of the young, yet wise, Prof. Carrera. Prof. Carrera thanks the co-authors for their contribution that will pave his future research during at least the next 20 years.

### References

- Ackerer, P. and Delay, F. (2010). Inversion of a set of well-test interferences in a fractured limestone aquifer by using an automatic downscaling parameterization technique. *J. Hydrol.*, 389, 42–56.
- Ahmed, S., de Marsily, G., and Talbot, A. (1988). Combined use of hydraulic and electrical properties of an aquifer in a geostatistical estimation of transmissivity. *Groundwater*, 26(1), 78–86.
- Alcolea, A., Carrera, J., and Medina, A. (2006). Pilot points method incorporating prior information for solving the groundwater flow inverse problem. *Adv. Water Resour.*, 29(11), 1678–1689.
- Ben Ameer, H., Chavent, G., and Jaffré, J. (2002). Refinement and coarsening indicators for adaptive parameterization: Application of the estimation of hydraulic transmissivities. *Inverse Probl.*, 18, 775–794.
- Bennett, R. R. and Meyzer, R. R. (1952). Geology and groundwater resources of the Baltimore area. In *Maryland Geological Survey Bulletin 4*, page 573.
- Berre, I., Lien, M., and Mannseth, T. (2007). A level-set corrector to an adaptive multiscale permeability prediction. *Comput. Geosci.*, 11, 27–42.
- Bertrand, J. (1889). *Calcul des probabilités*. Gauthier-Villars, Paris.
- Beucher, H. and Renard, D. (2016). Truncated Gaussian and derived methods. *C. R. Geosci.*, 348(7), 510–519.
- Capilla, J. E., Gomez-Hernandez, J. J., and Sahuquillo, A. (1998). Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric head data - 3. Application to the Culebra Formation at the Waste Isolation Pilot Plan (WIPP), New Mexico, USA. *J. Hydrol.*, 207(3–4), 254–269.
- Cardiff, M. and Kitanidis, P. K. (2009). Bayesian inversion for facies detection: An extensible level set framework: Level sets for facies detection. *Water Resour. Res.*, 45, article no. W10416.
- Carrera, J. (1988). State of the art of the inverse problem applied to the flow and solute transport equations. In *Groundwater Flow and Quality Modeling*, NATO ASI Series, pages 549–583. Springer, Dordrecht.
- Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., and Slooten, L. (2005). Inverse problem in hydrogeology. *Hydrogeol. J.*, 13, 206–222.
- Carrera, J. and Neuman, S. P. (1986a). Estimation of aquifer parameters under transient and steady-state conditions, 1. Maximum likelihood method incorporating prior information. *Water Resour. Res.*, 22(2), 199–210.
- Carrera, J. and Neuman, S. P. (1986b). Estimation of aquifer parameters under transient and steady-state conditions, 2. Uniqueness, stability and solution algorithms. *Water Resour. Res.*, 22(2), 211–227.
- Coptý, N., Rubin, Y., and Mavko, G. (1993). Geophysical-hydrological identification of field permeabilities through Bayesian updating. *Water Resour. Res.*, 29, 2813–2825.
- Coulthard, T. J., Hicks, D. M., and Van de Wiel, M. J. (2007). Cellular modeling of river catchment and reaches: Advantages, limitations and prospects. *Geomorphology*, 90, 192–207.
- Cui, T., Sreekanth, J., Pickett, T., Rassam, D., Gilfedder, M., and Barrett, D. (2021). Impact of model parameterization on predictive uncertainty of regional groundwater models in the context of environmental impact assessment. *Environ. Impact Assess. Rev.*, 90, article no. 106620.
- Dam, D. and Christensen, S. (2003). Including geophysical data in groundwater model inverse calibration. *Groundwater*, 41(2), 178–189.
- de Marsily, G. (1978). *De l'identification des systèmes géologiques*. Thèse de doctorat d'état ès sciences naturelles, Université de P. et M. Curie, Paris VI.
- de Marsily, G., Combes, P., and Goblet, P. (1992).

- Comment on 'Ground-water models cannot be validated', by L. F. Konikow & J. D. Bredehoeft. *Adv. Water Resour.*, 15, 367–369.
- de Marsily, G., Delay, F., Teles, V., and Schafmeister, M. T. (1998). Some current methods to represent the heterogeneity of natural media in hydrogeology. *Hydrogeol. J.*, 6, 115–130.
- de Marsily, G., Delhomme, J. P., Coudrain-Ribstein, A., and Lavenue, A. M. (2000). Four decades of inverse problems in hydrogeology. In Zhang, D. and Winter, C. L., editors, *Theory, Modeling, and Field Investigation in Hydrogeology: A Special Volume in Honor of Shlomo P. Neumann's 60th Birthday*, pages 1–17. Geological Society of America, Boulder, Colorado. Special paper 348.
- de Marsily, G., Lavedan, G., Boucher, M., and Fasanino, G. (1984). Interpretation of interference tests in a well field using geostatistical techniques to fit the permeability distribution in a reservoir model. In *Geostatistics for Natural Resources Characterization. Part 2*, pages 831–849. D. Reidel Pub. Co., Dordrecht.
- Delay, F., Ackerer, P., and Guadagnini, A. (2011). Theoretical analysis and field evidence of reciprocity gaps during interference pumping tests. *Adv. Water Resour.*, 34(5), 592–606.
- Doherty, J. (2003). Groundwater model calibration using pilot points and regularization. *Ground Water*, 41(2), 170–177.
- Doherty, J. E., Fienen, M. N., and Hunt, R. J. (2010). Approaches to highly parameterized inversion: Pilot-point theory, guidelines, and research directions. U.S. Geological Survey Scientific Investigations Report 2010–5168, 36 p.
- Emsellem, Y. and de Marsily, G. (1971). An automatic solution for the inverse problem. *Water Resour. Res.*, 7(5), 1264–1283.
- Engl, H. W., Hanke, M., and Neubauer, A. (1996). *Regularization of Inverse Problems*. Kluwer, Dordrecht. ISBN 0-7923-4157-0.
- Finsterle, S. and Kowalsky, M. B. (2008). Joint hydrological–geophysical inversion for soil structure identification. *Vadose Zone J.*, 7, 287–293.
- Fookes, P., Pettifer, G., and Waltham, T. (2015). *Geomodels in Engineering Geology: An Introduction*. Whittles Publishing, Dunbeath, Caithness, Scotland, UK.
- Ginn, T. R. and Cushman, J. H. (1990). Inverse methods for subsurface flow: A critical review of stochastic techniques. *Stochastic Hydrol. Hydraul.*, 4, 1–26.
- Gómez-Hernández, J. J. (2006). Complexity. *Ground Water*, 44, 782–785.
- Gómez-Hernández, J. J., Sahuquillo, A., and Capilla, J. E. (1997). Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data. 1. Theory. *J. Hydrol.*, 204(1–4), 162–174.
- Gómez-Hernández, J. J. and Wen, X. H. (1998). To be or not to be multi-Gaussian? A reflection on stochastic hydrogeology. *Adv. Water Resour.*, 21(1), 47–61.
- Gravey, M. and Mariethoz, G. (2020). Quicksampling v1.0: a robust and simplified pixel-based multiple-point simulation approach. *Geosci. Model Dev.*, 13(6), 2611–2630.
- Grimstad, A. A., Mannseth, T., Naevdal, G., and Urkedal, H. (2003). Adaptive multiscale permeability estimation. *Comput. Geosci.*, 7, 1–25.
- Haber, E. and Oldenburg, D. (1997). Joint inversion: a structural approach. *Inverse Probl.*, 13, 63–77.
- Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique. *Bull. Univ. Princeton*, 13, 49–52.
- Hassane, F. M., Delay, F., and Ackerer, P. (2017). Estimating initial conditions for groundwater flow modeling using an adaptive inverse method. *J. Hydrol.*, 552, 52–61.
- Hayek, M. and Ackerer, P. (2006). An adaptive subdivision algorithm for the identification of the diffusion coefficient in two-dimensional elliptic problems. *J. Math. Model. Algorithms*, 6(4), 529–545.
- Hernandez, A. F., Neuman, S. P., Guadagnini, A., and Carrera, J. (2003). Conditioning mean steady state flow on hydraulic head and conductivity through geostatistical inversion. *Stochastic Environ. Res. Risk Assess.*, 17(5), 329–338.
- Hyndman, D. W., Harris, J. M., and Gorelick, S. M. (1994). Coupled seismic and tracer test inversion for aquifer property characterization. *Water Resour. Res.*, 30, 1965–1977.
- Jacquard, P. and Jain, C. (1965). Permeability distribution from field pressure data. *Soc. Pet. Eng. J.*, 5, 281–294.
- Jahns, H. O. (1966). A rapid method for obtaining a two-dimensional reservoir description from well pressure response data. *Soc. Pet. Eng. J.*, 6, 315–327.
- Jaynes, E. T. (1984). Prior information and ambigu-



- ity in inverse problems. In *SIAM-AMS Proceedings*, volume 14. American Mathematical Society, Providence, RI.
- Kitanidis, P. K. (2007). On stochastic inverse modeling. In Hyndman, D. W., Day-Lewis, F. D., and Singha, K., editors, *Subsurface Hydrology: Data Integration for Properties and Processes, Volume 171*, Geophysical Monograph Series, pages 19–30. American Geophysical Union, Washington, DC.
- Kitanidis, P. K. and Vomvoris, E. G. (1983). A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations. *Water Resour. Res.*, 19(3), 677–690.
- Kolterman, C. E. and Gorelick, S. M. (1996). Heterogeneity in sedimentary deposits: a review of structure-imitating, process-imitating, and descriptive approaches. *Water Resour. Res.*, 32, 2617–2658.
- Lavenue, A. M., RamaRao, B. S., de Marsily, G., and Marietta, M. G. (1995). Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields: 2. Application. *Water Resour. Res.*, 31(3), 495–516.
- Lavenue, M. and De Marsily, G. (2001). Three-dimensional interference test interpretation in a fractured aquifer using the pilot point inverse method. *Water Resour. Res.*, 37(11), 2659–2675.
- Le Verrier, U. (1846). *Recherches sur les mouvements de la planète Herschel, dite Uranus*. Bachelier, Paris, <http://hdl.handle.net/1908/3096>.
- Linde, N. and Doetsch, J. (2016). Joint inversion in hydrogeophysics and near-surface geophysics. In Moorkamp, M., Lelièvre, P. G., Linde, N., and Khan, A., editors, *Integrated Imaging of the Earth: Theory and Applications*, Geophysical Monograph Series. American Geophysical Union, Washington, DC.
- Lu, Z. and Robinson, B. A. (2006). Parameter identification using the level set method. *Geophys. Res. Lett.*, 33, article no. L06404.
- Ma, W. and Jafarpour, B. (2018). Pilot points method for conditioning multiple-point statistical facies simulation on flow data. *Adv. Water Resour.*, 115, 219–233.
- Marinoni, M., Delay, F., Ackerer, P., Riva, M., and Guadagnini, A. (2016). Identification of groundwater flow parameters using reciprocal data from hydraulic interference tests. *J. Hydrol.*, 539, 88–101.
- Matheron, G. (1967). *Éléments pour une Théorie des Milieux Poreux*. Masson & Cie, Paris.
- Matheron, G., Beucher, H., de Fouquet, C., Galli, A., Guerillot, D., and Ravenne, C. (1987). Conditional simulation of the geometry of fluvio-deltaic reservoirs. In *SPE Annual Technical Conference and Exhibition, Dallas, Texas, September 1987*. Soc. Petrol. Eng., SPE.
- McLaughlin, D. and Townley, L. R. (1996). A reassessment of the groundwater inverse problem. *Water Resour. Res.*, 32(5), 1131–1161.
- Moore, C. and Doherty, J. (2005). Role of the calibration process in reducing model predictive error. *Water Resour. Res.*, 41, article no. W05020.
- Nelson, R. W. (1960). In place measurement of permeability in heterogeneous media, 1. Theory of a proposed method. *J. Geophys. Res.*, 65(6), 1753–1760.
- Neuman, S. P. (1973). Calibration of distributed parameter groundwater flow models viewed as a multiple-objective decision process under uncertainty. *Water Resour. Res.*, 9(4), 1006–1021.
- Neuman, S. P. (1980). A statistical approach to the inverse problem of aquifer hydrology. 3. Improved solution method and added perspective. *Water Resour. Res.*, 16(2), 331–346.
- Oliver, D. S. and Chen, Y. (2010). Recent progress on reservoir history matching: a review. *Comput. Geosci.*, 15(1), 185–221.
- Olson, N. R. B. (2021). 3-D numerical modeling of braided channel formation. *Geomorphology*, 375, article no. 107528.
- Pittaluga, M. B., Nobile, G., and Seminara, G. (2009). A non-linear model for river meandering. *Water Resour. Res.*, 45(4), article no. W04432.
- Polya, G. (1954). Mathematics and plausible reasoning. In *Induction and Analogy in Mathematics*, volume 1. Princeton University Press, Princeton, NJ.
- Pool, M., Carrera, J., Alcolea, A., and Bocanegra, E. M. (2015). A comparison of deterministic and stochastic approaches for regional scale inverse modeling on the Mar del Plata aquifer. *J. Hydrol.*, 531, 214–229.
- Ramarao, B., LaVenué, M., de Marsily, G., and Marietta, M. G. (1995). Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields. Part 1. Theory and computational experiments. *Water Resour. Res.*, 31, 475–493.
- Ross, M., Parent, M., and Lefebvre, R. (2005). 3-

- D geologic framework models for regional hydrogeology and land use management: A case study from a Quaternary basin of southwestern Quebec, Canada. *Hydrogeol. J.*, 13, 690–707.
- Rubin, Y. and Dagan, G. (1987). Stochastic identification of transmissivity and effective recharge in steady groundwater flow, 1. Theory. *Water Resour. Res.*, 23(7), 1185–1192.
- Rubin, Y., Mavko, G., and Harris, J. (1992). Mapping permeability in heterogeneous aquifers using hydrologic and seismic data. *Water Resour. Res.*, 28, 1809–1816.
- Schafmeister, M.-Th. and de Marsily, G. (1994). The influence of correlation length of highly conductive zones in alluvial media on the transport behavior. In Dracos, T. and Stauffer, S., editors, *Transport and Reactive Processes in Aquifers*, pages 171–176. Balkema, Rotterdam.
- Schöniger, A., Nowak, W., and Hendricks Franssen, H.-J. (2012). Parameter estimation by ensemble Kalman filters with transformed data: Approach and application to hydraulic tomography: parameter estimation by tEnKFs. *Water Resour. Res.*, 48, article no. W04502.
- Stallman, R. W. (1956). Numerical analysis of regional water levels to define aquifer hydrology. *Am. Geophys. Union Trans.*, 37(4), 451–460.
- Sun, J., Lin, B., and Yang, H. (2015). Development and applications of braided river model with non-uniform sediment transport. *Adv. Water Resour.*, 81, 66–74.
- Sun, N.-Z., Yang, S., and Yeh, W. W.-G. (1998). A proposed stepwise regression method for model structure identification. *Water Resour. Res.*, 34, 2561–2572.
- Teles, V., Bravard, J.-P., de Marsily, G., and Perrier, E. (2001). Modeling of the construction of the Rhône alluvial plain since 15,000 years BP. *Sedimentology*, 48, 1209–1224.
- Teles, V., Delay, F., and de Marsily, G. (2004). Comparison of genesis and geostatistical methods for characterizing the heterogeneity of alluvial media: Groundwater flow and simulations. *J. Hydrol.*, 294, 103–121.
- Tetzlaff, D. M. and Harbaugh, J. W. (1989). *Simulating Clastic Sedimentation*. Van Nostrand Reinhold, New York.
- Tihonov, A. N. (1963a). Regularization of incorrectly posed problems. *Sov. Math. Dokl.*, 4, 1624–1627.
- Tihonov, A. N. (1963b). Solution of incorrectly formulated problems and the regularization method. *Sov. Math. Dokl.*, 4, 1035–1038.
- Tung, C.-P. and Tan, C.-C. (2005). An optimal procedure for identifying parameter structure and application to a confined aquifer. *Environ. Geol.*, 47, 1062–1071.
- Vesselinov, V. V., Neuman, S. P., and Illman, W. A. (2001). Three-dimensional numerical inversion of pneumatic cross-hole tests in unsaturated fractured tuff 2. Equivalent parameters, high-resolution stochastic imaging and scale effects. *Water Resour. Res.*, 37(12), 3019–3041.
- White, J. T. and Lavenue, A. M. (2022). Advances in the pilot point inverse method: Où en sommes-nous maintenant? *C. R. Géosci.* Forthcoming.
- Williams, R. D., Brasington, J., and Hicks, D. M. (2016). Numerical modeling of braided river morphodynamics: Review and future challenges. *Geogr. Compass*, 10(3), 102–127.
- Yakowitz, S. and Duckstein, L. (1980). Instability in aquifer identification: theory and case study. *Water Resour. Res.*, 16(6), 1045–1064.
- Yeh, W. W.-G. (1975). Aquifer parameter identification. *J. Hydraul. Div. Am. Soc. Civ. Eng.*, 101(HY9), 1197–1209.
- Yeh, W. W. G. (1986). Review of parameter estimation procedures in groundwater hydrology: The inverse problem. *Water Resour. Res.*, 22, 95–108.
- Yeh, W. W. G. (2015). Review: Optimization methods for groundwater modeling and management. *Hydrogeol. J.*, 23, 1051–1065.
- Zhang, M., Burbey, T. J., Nunes, V. D. S., and Borggaard, J. (2014). A new zonation algorithm with parameter estimation using hydraulic head and subsidence observations. *Groundwater*, 52, 514–524.
- Zhou, H., Gómez-Hernández, J. J., and Li, L. (2014). Inverse methods in hydrogeology: Evolution and recent trends. *Adv. Water Resour.*, 63, 22–37.



---

Research article

Geo-hydrological Data & Models

# Is subsurface geophysics as seismic and acoustic investigations a rescue to groundwater flow inversion?

Frederick Delay<sup>Ⓣ,\*</sup>,<sup>a</sup>, Jean-Luc Mari<sup>b</sup>, Gilles Porel<sup>Ⓣ</sup>,<sup>c</sup>, François Chabaux<sup>a</sup> and Philippe Ackerer<sup>Ⓣ</sup>,<sup>a</sup>

<sup>a</sup> Institut Terre et Environnement de Strasbourg, UMR 7063 CNRS, Université de Strasbourg, Engées, 5 Rue René Descartes, 67000 Strasbourg, France

<sup>b</sup> Sorbonne Université, CNRS, EPHE, UMR 7619 METIS, 4 Place Jussieu, 75005, Paris, France

<sup>c</sup> IC2MP UMR 7285, Department of Earth Sciences, University of Poitiers, CNRS, HydrASA, F-86000 Poitiers, France

*E-mails:* fdelay@unistra.fr (F. Delay), jeanluc90.mari@gmail.com (J.-L. Mari), gilles.porel@univ-poitiers.fr (G. Porel), fchabaux@unistra.fr (F. Chabaux), ackerer@unistra.fr (P. Ackerer)

**Abstract.** Understanding subsurface flow, especially in partly karstified rock formations mainly housing water through a few preferential pathways, is still challenging. This point is the consequence of the poor accessibility of the subsurface and lack of accurate depictions of water bearing bodies and distributions. This notwithstanding, highly-resolved geophysical investigations bring new images of the subsurface.

A 3-D seismic survey with shots and wave monitoring at the surface is carried out over a subsurface karstified reservoir located at the Hydrogeological Experimental Site (HES) of the University of Poitiers (France). Processing the 3-D data, in association with wave velocity calibration from vertical seismic profiles (VSP) recorded via geophones in wells, renders a 3-D velocity block. The velocity block is then converted into pseudo-porosity values revealing three high-porosity, presumably water-productive, layers, at depths of 35–40, 85–87, and 110–115 m.

In addition, full wave acoustic logging (FWAL) can detect, close to wells, porous or open bodies that are too small for being captured by the spatial resolution of 3-D seismic images. A FWAL can also confirm or invalidate data from VSP recorded via hydrophones.

The block of pseudo-porosities is compared to a different representation of the subsurface in the form of hydraulic conductivity distributions (or hydraulic diffusion) obtained by slug tests or by inversion of transient interference testing between wells. The inverted hydraulic conductivity maps do not match up the distribution of porous bodies identified by seismic data. This poses the question of guiding conventional inversions on the basis of a prior guess as the subsurface structure obtained via geophysical investigations.

**Keywords.** 3-D seismic, Vertical seismic profile, Acoustic logging, Karstic bodies, Porosity, Hydraulic conductivity maps.

---

\* Corresponding author.

**Foreword.** This contribution is mainly written as a compilation of more than 15 years spent by the authors studying a hydrogeological site. It is proposed in tribute to Ghislain de Marsily who instilled us the overall philosophy saying that: in research activity, tough is not enough, probably time, patience, and perseverance are better suited. Undoubtedly, something in good stands with the subliminal message expressed by the title “Regards sur 40 ans de problèmes inverses en hydrogéologie”—de Marsily *et al.*, *C. R. Acad. Sci.*—1999.

*Manuscript received 22 July 2022, accepted 6 September 2022.*

## 1. Introduction

Many portions of subsurface aquifers were developed as experimental sites during the past few decades. These sites were mainly designed over small surface areas including several wells for better understanding the behavior of heterogeneous subsurface systems via in-situ measurements and model calibrations regarding flow, transport, reactions, and other features. The University of Poitiers (France) had a Hydrogeological Experimental Site (HES) built near the Campus for the sole purpose of providing facilities to develop long-term monitoring and experiments investigating the water and mass transfer processes within continuous but partly fractured-karstified limestone aquifers [Bernard *et al.*, 2006, Kaczmarsky and Delay, 2007b,a].

This contribution highlights the benefit of combining seismic methods (2-D, 3-D seismic reflection imaging, vertical seismic profile (VSP) and full wave acoustic logging (FWAL)) to characterize a near surface, locally karstified, aquifer. 35 wells were drilled at the HES to perform hydrogeological experiments, such as slug tests, interference testing, tracer tests, etc. In each well, a borehole televiewer (BHTV) was run with the idea of collecting optic images that would detect and visualize cross-cut, sparse karstic bodies within a “compact” host rock. Sparsity should not prevent these bodies from triggering regional flow with sometimes preferential pathways. In 2004, a high frequency band (up to 200 Hz) 3-D seismic survey was recorded at the HES. The experiment resulted, after diverse signal processing and calibration steps, into a 3-D seismic pseudo-porosity block which revealed three high-porosity layers, at depths of 35–40, 85–87 and 110–115 m. The BHTV images confirmed that the three high-porosity layers detected at the seismic scale were karstic layers.

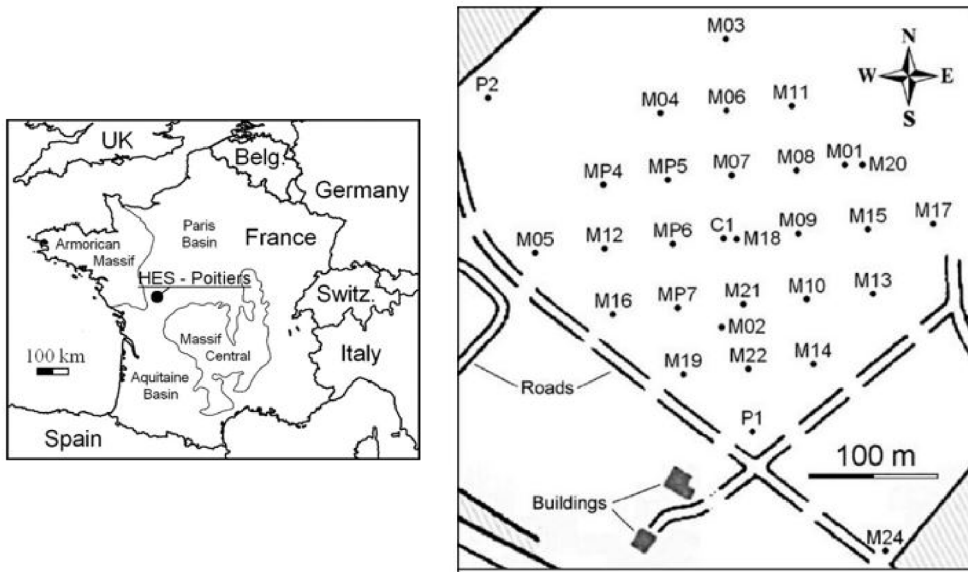
After a rapid description of the hydrogeological context at the HES, this work reviews (but discards the technical recipes) the procedure of 3-D seismic imaging used to map the distributions of karstic bodies in the near subsurface. Raw seismic data are

relative, in the sense they only mark contrasts of wave propagation velocities and wave amplitudes between different portions of a subsurface block. Therefore, it is needed to “tie” and calibrate seismic information onto data collected in wells. These tying and calibration steps of seismic data are carried out by relying upon “absolute” geophysical measurements in boreholes such as VSP and FWAL. It is also worth noting that 3-D seismic images do not have enough resolution, both in the vertical and horizontal directions, to detect karstic bodies of small size (on the order of metric lateral extension and decimeter aperture), like those observed with BHTV in boreholes (for example in well C1 within the 45–60 m depth interval, described later). The paper shows how FWAL and VSP recorded via hydrophones can be used both to confirm the results obtained from 3-D seismic and detect karstic bodies of small size.

Then, the seismic 3-D block evidencing karstic bodies, assumed as bearing the major part of flow, is compared to more conventional images of water conduction in a subsurface system. These images are stemming from local slug tests within single wells but with monitored responses at distant wells. Large-scale interference testing interpreted by inversion of a distributed forward flow model also render maps of hydraulic conductivity values that are compared to the flowing structures revealed by geophysics.

## 2. Hydrogeological context

The aquifer at the HES lies between 20 to 130 m depth and consists of compact karstic carbonates of Middle Jurassic age. It is located at the border, also called the “Poitou threshold”, between the Paris and the Aquitaine sedimentary basins (Figure 1). The carbonates may show sedimentary variability over depth and lateral extension, but in the absence of fracturation and/or dissolution, their hydraulic conductivity remain very low ( $10^{-8}$  m/s) and the porosity less than 2–3%. The HES covers an area of 12 hectares over which 35 wells were drilled to a depth of 120 m



**Figure 1.** Hydrogeological Experimental Site (HES) in Poitiers: site map and well locations.

(Figure 1), these wells showing maximum pumping rates varying from well to well and ranging values between 5 to 150 m<sup>3</sup>/h. Most of the wells are steel cased over the first meter depth (15–25 m) and below, either PVC screened over the whole thickness of the aquifer or left uncased. The top of the reservoir was initially flat and horizontal, 150 million years ago, but was eroded and weathered since, during Cretaceous and Tertiary ages. It is shaped today as a smoothly bumped topographic surface with amplitudes of relative elevations reaching up to 20 m, the whole keeping the underlying limestone aquifer as a hydraulically confined formation. The building phase of the HES started in 2002 and it took approximately 4 years (in two drilling campaigns) to get the set of 35 wells, all bored over the whole thickness of the aquifer. Most wells are documented by drilling records, borehole optic imaging, and logs of various nature among which electrical resistivity, gamma-ray, temperature, and acoustic logs are available. In addition, two wells were entirely cored.

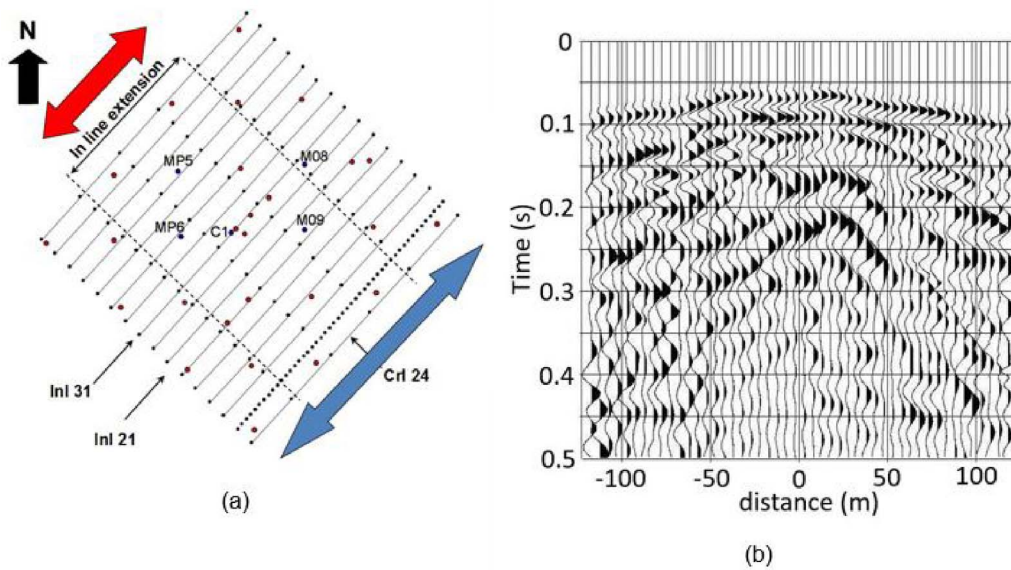
Today, the aquifer responds evenly to the hydraulic stress of a pumped well, with pressure depletion curves (interferences) merged in time and amplitude irrespective of the distance from the pumped well [Ackerer and Delay, 2010, Delay *et al.*, 2011, Trottier *et al.*, 2014]. This merging is likely to be a consequence of a local karstic flow devel-

oped in open conduits that have been unclogged and connected by the drilling and pumping works. The presence of pervasive karstic drains is supported by logs in the wells using optic imaging. Almost all the wells have shown caves and conduits cross-cut by wellbores, with mean apertures of 0.2–0.5 m. These conduits are mostly concealed in three thin horizontal layers at depths of approximately 35, 85, and 110 m. As these open layers are intercepted by almost all vertical wells, this results in a good connection between wells and karstic drains, and eventually between the open layers. The connection would be mainly controlled by the degree to which drains are re-opened in the vicinity of the wells and the connection of drains within a layer.

With the above considerations in mind, it was found reasonable to better image the geometry of the reservoir with a resolution compatible with, on the one hand, the scale of a well and, on the other hand, the scale of the entire experimental disposal. High resolution geophysical tools seem well suited to undertake that kind of investigation.

### 3. Seismic imaging

The seismic reflection method is able to produce a picture of the subsurface in three dimensions (3-D) over a regular grid. Usually, in high-resolution



**Figure 2.** 3D-Seismic acquisition. (a) Seismic line implementation and well location (red dots). In-line extension as the red arrow for cross shots, and in-line extension as the blue arrow for direct and reverse shots. (b) Example of cross-line shot point.

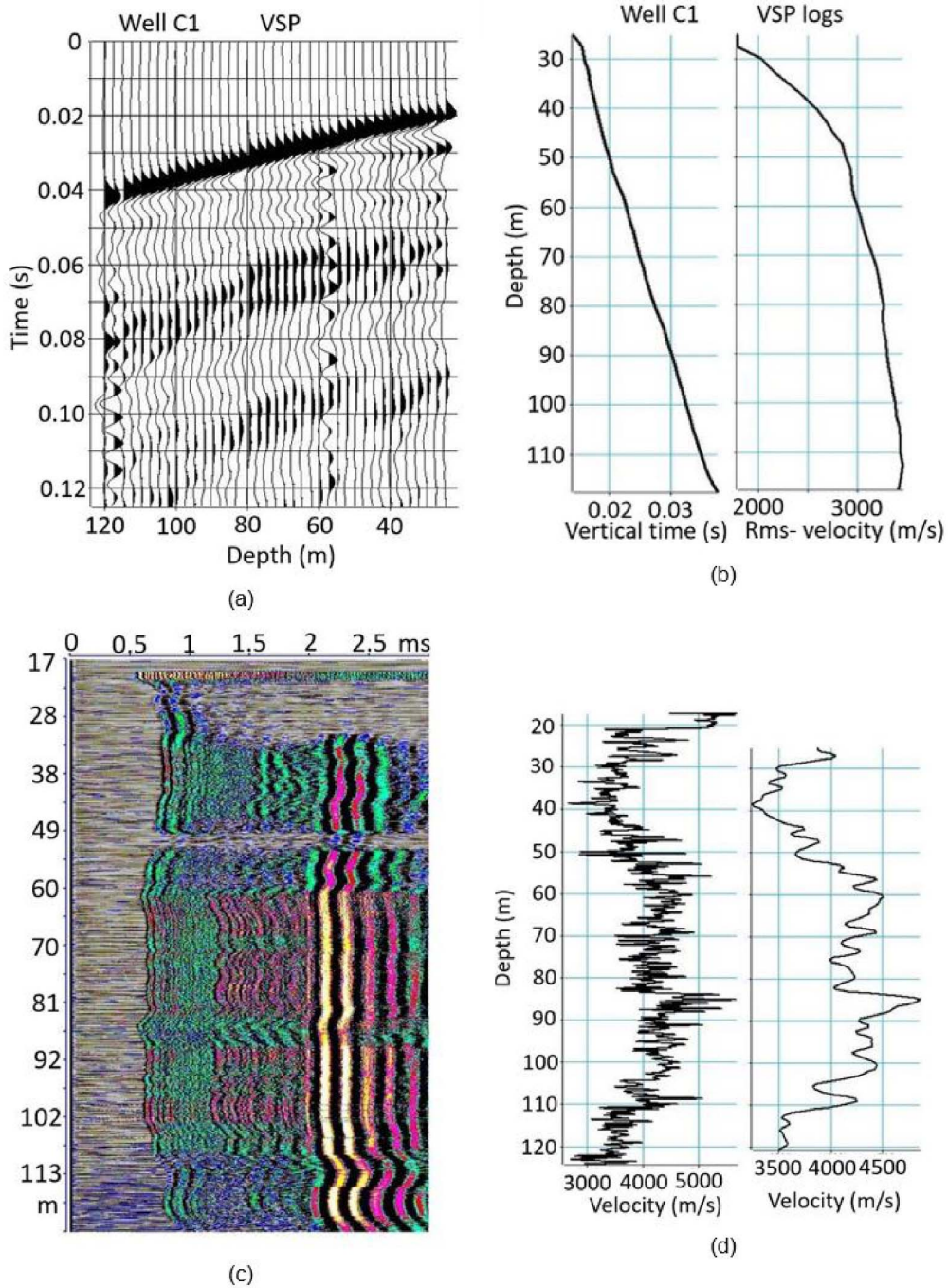
seismic surveys, the size of the elementary parallelepiped grid cell is of the order of a few tens of meters along the horizontal directions, and of a few (3–5) m along the vertical direction, with an even precision up to 200 m depth. This level of spatial resolution requires dealing with large amounts of raw data and heavy computations for data processing.

At the HES, which is not very large (approximately  $300 \times 300 \times 150 \text{ m}^3$ ), the 3-D seismic survey was refined but designed to obtain a 3-D block recorded in low amounts of data [Mari and Porel, 2008]. The complete survey is composed of 20 receiver lines (the in-line direction) with a 15 m lag distance between adjacent lines (Figure 2a), each line being composed of 48 single geophones with 5 m spacing between adjacent geophones. For each receiver line investigated, direct and reverse shots with sources in line with the receiver line were recorded. Cross-line shots fired at distances of 40, 50, and 60 m perpendicular to the receiver line were also recorded (example of cross-line shot in Figure 2b). Processing the data from in-line direct and reverse shots gathers the results in a vertical section of 240 m in-line extension (the blue arrow, in Figure 2a), while a cross-line shot gathers results in a vertical section of only 120 m extension along the in-line direction (the red arrow on

the location map of seismic lines, in Figure 2a).

The whole set of single-fold vertical sections of arrival times were assembled to build a 3-D block of 240 and 300 m extension along the in-line and cross-line directions, respectively. For the in-line direction profile, the reference zero in the horizontal direction is taken as the location of the sources points, in the middle of the receiver lines. This renders reflecting point locations varying between  $-120$  and  $+120$  m and a lag distance between two neighbor reflecting points of 2.5 m (half the distance between two adjacent receivers). For the cross-line shots, the reference zero is still taken at the middle of the receiver lines, but with a lag distance between neighbor reflecting points of 5 m.

Seismic data processing was carried out to transform raw arrival times into a 3-D block of wave amplitudes reflected at diverse depth and horizontal locations of the aquifer system. A VSP recorded via geophones placed in well C1 (Figure 3a), was also processed to obtain a law (a conversion) of reflected time versus depth and a velocity model ( $V_{\text{rms}}$ ) of reflected wave propagation. A 3-D refraction seismic tomography [Mari and Mendes, 2012] was finally carried out, with the idea of mapping the irregular shape of the clay cover at the surface (up to 25–30 m depth)



**Figure 3.** Vertical seismic profile (VSP) and acoustic logging at well C1. (a) VSP section. (b) Time vs depth law and velocity ( $V_{rms}$ ) log. (c) Acoustic section; vertical axis: depth in m, horizontal axis: time in ms. (d) Raw acoustic velocity logs, and after smoothing.

overlying the limestone aquifer. This experiment is assumed to render static corrections for the seismic reflected block and a propagation velocity model in

the clay cover of the aquifer (also for corrections of the seismic 3-D block).

The processing sequence [Mari and Delay, 2011,

Mari and Porel, 2008] includes: amplitude recovery, deconvolution, wave separation, static corrections, and normal move-out corrections (using the VSP velocity model, Figure 3b). Times of the reflected sections were also converted into depths via the VSP time versus depth law inferred in well C1 (Figure 3b).

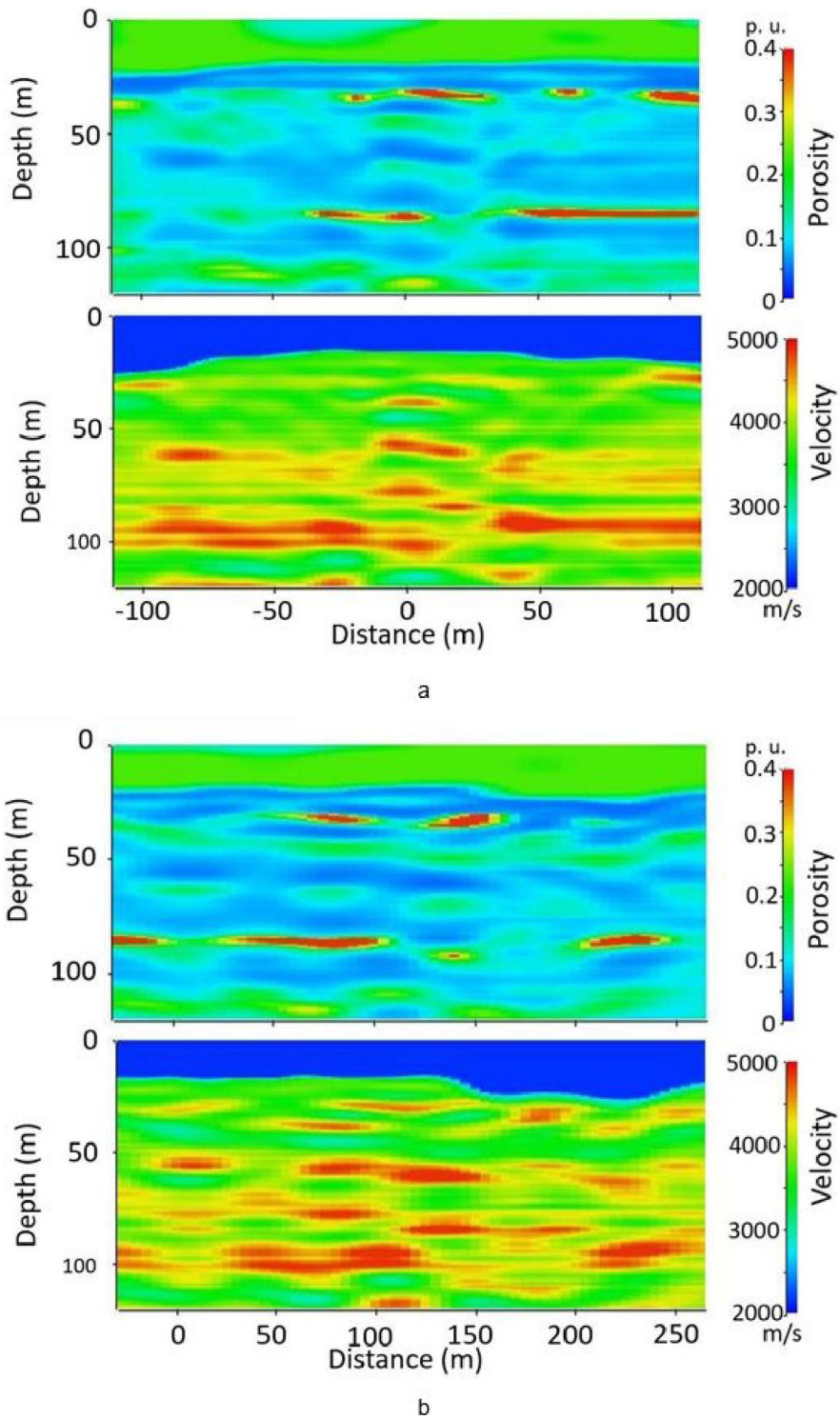
Acoustic data were also recorded in the boreholes of the HES and that of the well C1 is exemplified hereafter. The acoustic probe is a flexible monopole tool holding a source as a magnetostrictive transducer and two pairs of receivers beneath: a pair of near receivers (1 and 1.25 m lag-distances (offsets) beneath the source) and a pair of far receivers (3 and 3.25 m offsets). The acoustic data were recorded in the 1–20 kHz frequency band. It is worth noting that compared with seismic data, the vertical resolution of the acoustic logging is very high, on the order of 10 cm, but the lateral investigation is limited to a few tens of centimeters around the borehole. Figure 3c reports on the 3 m constant offset section (receiver R1) with the vertical axis as the depth of receiver R1 from ground level, and the horizontal axis the recorded travel times (ms) of acoustic waves from source to receiver. In the acoustic section, the refracted *P*-waves appear in the 0.5–1.0 ms time window. Sampling the arrival times of the refracted *P*-wave recorded by two adjacent receivers (R1: 3 m offset, and R2: 3.25 m) allows for obtaining by difference the acoustic velocity within the formation at different depths (Figure 3d). This velocity log is usually highly variable over very short distances, but can be filtered in wavenumber and resampled over larger intervals (here 0.5 m) than the sampling step of the probe (Figure 3d-right). These larger resampling intervals are assumed to be compatible with the vertical resolution of the seismic block, with the aim to constrain the block via acoustic logs. More precisely, the filtered velocity functions computed from acoustic data recorded at wells C1, MP5, MP6, M08, and M09 (locations in Figure 1) were used to convert the 3-D block of seismic wave amplitudes into a 3-D pseudo velocity block [Mari and Delay, 2011, Mari and Porel, 2008].

It is reiterated that we are interested in evaluating the porosity variations within the aquifer at the HES, under the expectation that high porosity values could mark the locations of karstic features. To this end, the seismic velocities were first converted into electrical resistivity values [Mari et al., 2009], using

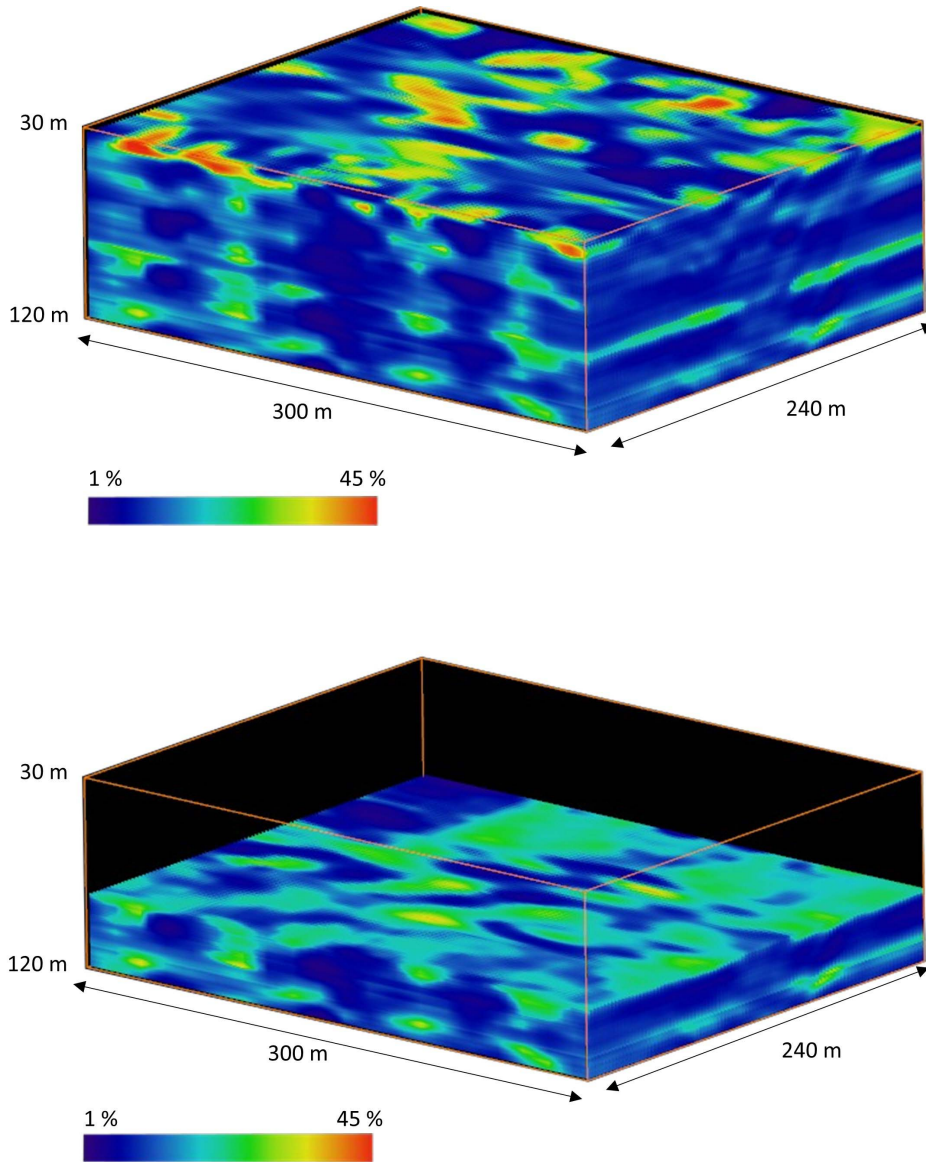
borehole resistivity logs as calibration and the empirical relationship between seismic velocity and actual formation resistivity proposed by Faust [1953]. This transformation is justified by the fact that we do not have redundant information in terms of velocity propagation because only a few shots were performed over the seismic line implementation at the surface. The velocities are therefore unbounded and may vary from one shot to the other. In the absence of petrophysical data informing on the actual values of velocity propagation over diverse samples of host rocks, the raw velocity data should be calibrated. This is why we relied upon a transformation of velocity into electrical resistivity, according to the empirical non-linear Faust's law and the 35 logs of resistivity monitored in the wells of the HES. The parameters of the Faust's law at the wells were interpolated over the whole surface area of the HES and velocities were transformed accordingly into a single 3-D block of electrical resistivity values. Those were then converted into pseudo-porosity values, by using the non-linear Archie's law [Archie, 1942] under the assumption of uniform water resistivity over the whole 3-D block. Figure 4 exemplifies the pseudo velocity and porosity spatial distributions via the in-line #21 section (Figure 4a) and the cross-line #24 section (Figure 4b). A few discrepancies can be observed between the patterns revealed by velocities and those of the transformation into porosities. This is mainly the consequence of the non-linearity of both Faust's and Archie's laws in the transformation velocity  $\rightarrow$  resistivity  $\rightarrow$  porosity, and to some extent, an artifact due to color scales in Figure 4 with porosity spanning a larger range of relative values than the velocities. Nevertheless, the images are consistent from one another with areas of weak porosities corresponding to high velocities and high porosity values located in low velocity areas.

After transformation of the whole set of vertical sections and reassembly, the resulting 3-D seismic pseudo-porosity block (Figure 5) revealed three high-porosity, presumably water-productive layers, at depths of 35–40, 85–87 and 110–115 m [Mari et al., 2020]. The 85–87 m depth layer appears the most porous one, with wide patches of porosity values higher than 30%, that would represent, for this layer, the karstic part of the reservoir (see bottom draw in Figure 5, upper layer).





**Figure 4.** Velocity and porosity sections. (a) In-line #21 section. (b) Cross-line #24 section.



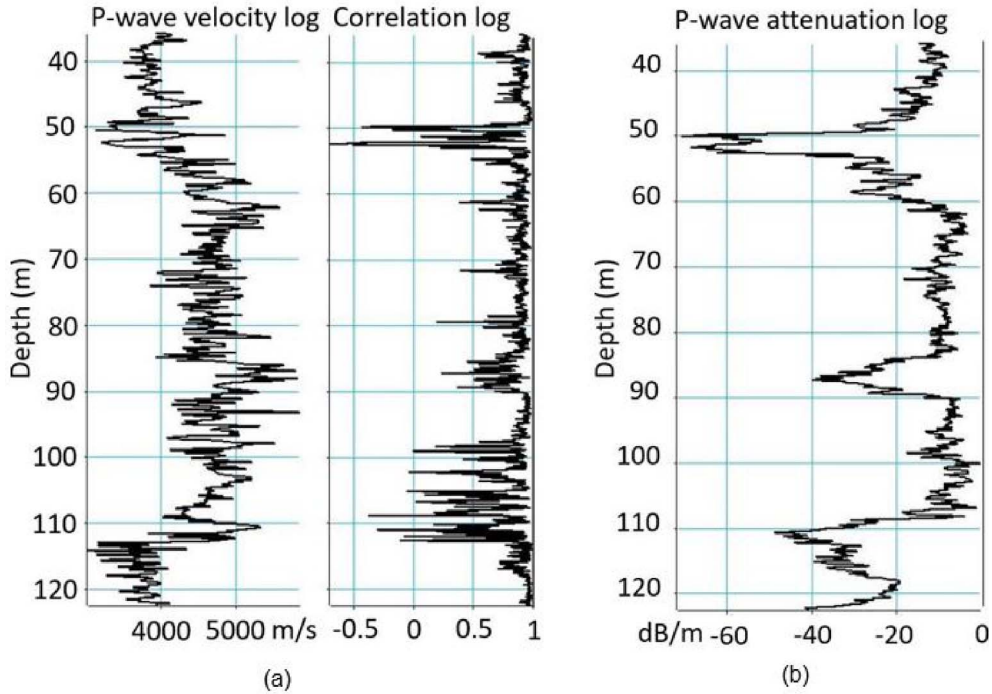
**Figure 5.** HES site–pseudo porosity block in the 30–120 m depth interval (top) and in the 85–120 m depth interval (bottom). 240 m of inline extension, NE–SW direction; 300 m of cross-line extension, NW–SE direction.

#### 4. Additional information from full waveform acoustic logging

As told earlier, acoustic logging is of very high resolution over depth but investigates a radius of a few tens of centimeters around the well. Full waveform acoustic logging (FWAL) not only analyzes the refracted *P*-waves propagating in the rock formation crosscut by the well, but also exploits the other acoustic wave

lengths monitored by the receivers. After data processing, a FWAL can either confirm or question locally the results from the 3-D seismic block, but also detect small open bodies in the host rock that the seismic resolution is unable to capture. As an example, the acoustic logging in well C1 (Figure 3c) is re-handled for a complete treatment.

The well C1 is equipped with a steel casing from the surface up to 22 m depth, below which, it



**Figure 6.** Full wave acoustic logging (FWAL) at well C1: refracted  $P$ -waves. (a) Velocity log and its associated correlation coefficient log. (b) Attenuation log.

becomes an open-hole wellbore in which the water table was detected at 20 m depth during the logging. On the acoustic section (Figure 3c), the refracted  $P$ -waves appear in the 0.5–1.0 ms time interval, the converted refracted  $S$ -waves and pseudo-Rayleigh waves in the 1.2–2 ms time interval, and the Stoneley wave in the 2–3 ms time interval. The logging distinguishes:

- The water table at 20 m depth (no acoustic signal monitored above 20 m), and between 20 and 22 m, some resonances due to a loose cementation of the casing.
- The 22–33 m depth interval with a very poor signal to noise ratio; only refracted  $P$ -wave are captured.
- The 33–60 m depth interval with visible refracted  $P$ -waves, converted refracted  $S$ -waves, and Stoneley waves; in the 49–54 m depth interval, all these waves are strongly attenuated.
- The 60–108 m depth interval of homogeneous profile with a very good signal to noise ratio; all the acoustic waves are visible and of strong amplitude.
- Finally, the 108–124 m depth interval, also relatively homogeneous with refracted  $P$ -waves and Stoneley waves of good amplitude; converted refracted  $S$ -waves are visible, but noisy.

The acoustic data have been processed in the 35–124 m depth interval, by first picking the arrival times of a wave train to obtain the velocity and the attenuation of the considered wave train [Mari and Vergnault, 2018]. Figure 6 shows the results obtained from the refracted  $P$ -waves. A log of correlation between signals monitored by the receivers R1 and R2 of the acoustic probe at diverse depths is associated with the velocity log. A high value of the correlation coefficient indicates that the shape of the acoustic signal is the same on the two receivers (R1 and R2). It also indicates that the picked times are accurate. A decrease in the correlation coefficient can sign a poor picking due to either a poor signal to noise ratio, or a change in the signal shape. The correlation coefficient log can be used to “edit” the velocity log, or at least being confident (or not) in some local velocity values (Figure 6a). The wave attenuation log (Figure 6b), calculated for a vertical wave

propagation over 0.25 m, is simply defined as the ratio of energy logs at equivalent locations between the log at 3 m offset (distance emitter–receiver R1) and the log at 3.25 m offset (receiver R2). This attenuation expressed in dB/m, is significant at 3 levels located at 50–54 m, 85–90 m, and 110–120 m depth. The two deepest levels are that of the karstic layers revealed by the 3-D seismic imaging. In the 50–54 m depth interval, a low energy, and a very strong attenuation of more than 60 dB/m are observed, even though no karstic (highly porous) layer is detected by the seismic approach.

The same procedure has been applied to the Stoneley waves (Figure 7), with energy and attenuation logs of the Stoneley wave underlining the same three levels identified by the analysis of the refracted *P*-wave. Interestingly, a very strong attenuation is measured in the 50–54 m depth interval, when the attenuation measured in the two deepest levels (85–90 and 110–120 m) is weaker. The energies of the different wave packets depend on the permeability of the ways through which the waves propagate, that are mainly the host rock for *P*-waves, and the interface between casing and wellbore for Stoneley waves. The sensitivity to changes in permeability is partly evidenced by a loss of total energy in the acoustic signal. The karstic level at 55 m depth crosscut by C1 is confirmed by VSP data recorded using hydrophones (Figure 7c). The VSP is highly corrupted by Stoneley waves. A conversion of downgoing *P*-wave (blue line in Figure 7c) in down (red line) and up going Stoneley waves is often observed in highly permeable formations [Hardage, 1984, Mari and Vergniault, 2018, Mari *et al.*, 2020]. In the present case, the first arrivals in the form of downgoing *P*-waves are strongly attenuated at 55 m, the waves being converted into downgoing Stoneley waves then reflected at the bottom of the well.

The analysis of acoustic data leads to the following observations:

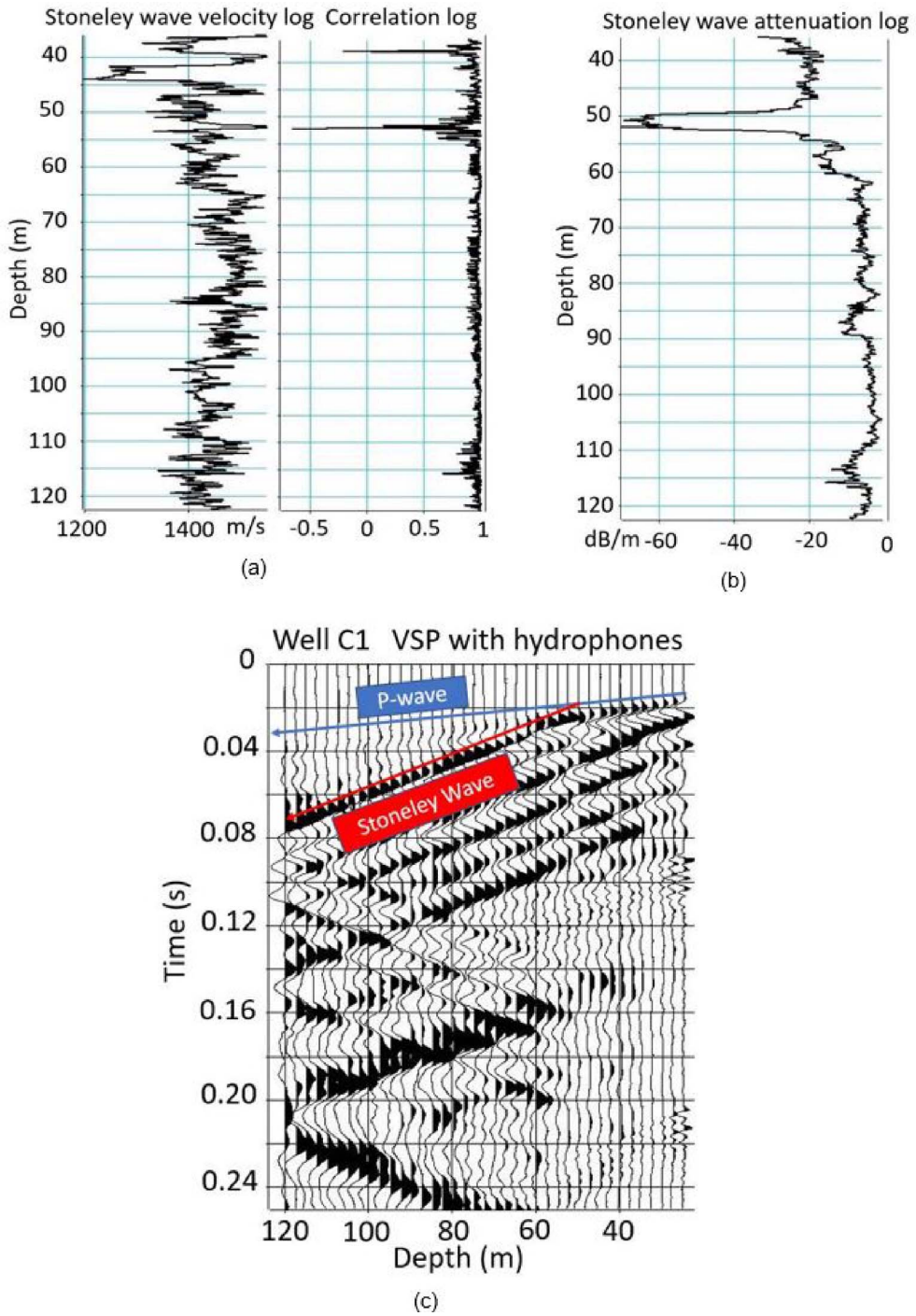
- The attenuation of the different wave packets indicates that the two deep levels (85–90 m and 110–115 m) revealed by seismic data have a high permeability. This notwithstanding, the continuity of the acoustic images suggests that no geological discontinuity, such as a karstic body, occurring at these depths in the close vicinity of well C1. This interpretation is confirmed by
- the 3-D seismic block showing that C1 does not crosscut karstic bodies in the 60–120 m depth interval.
- The loss of continuity of acoustic images between 50 and 54 m depth, and the very strong attenuation of the different waves, mainly the Stoneley waves, suggest that this level could be a karstic body with a very small lateral extension at C1 not visible by seismic data.

## 5. Hydraulic investigations

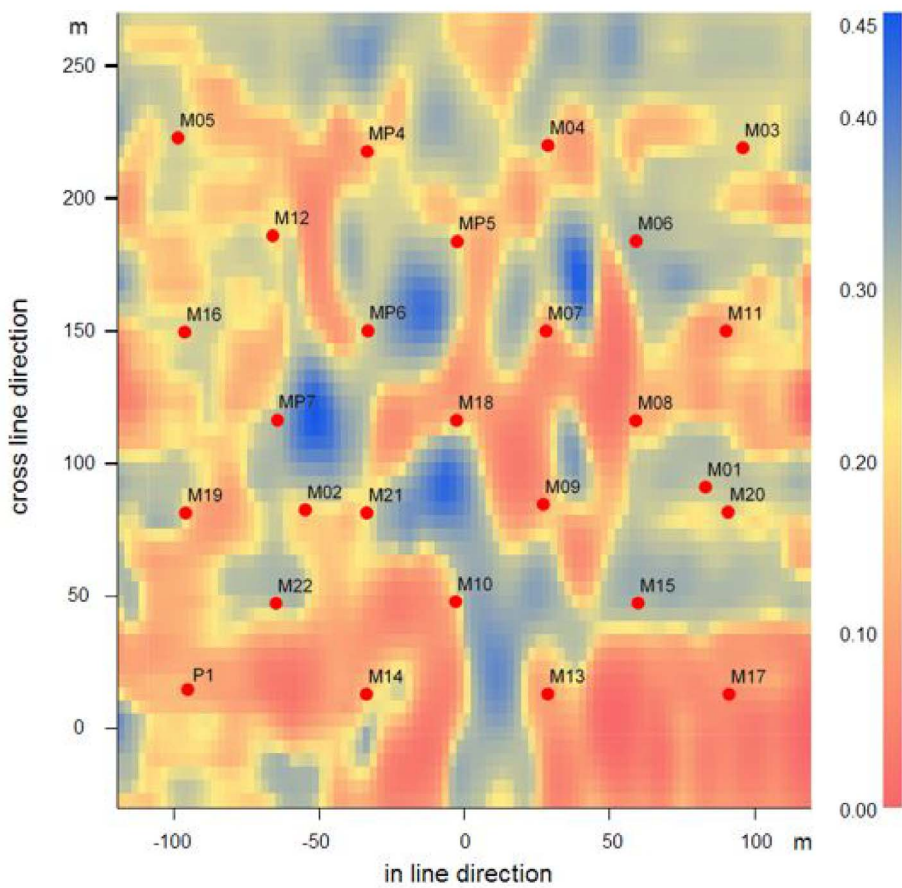
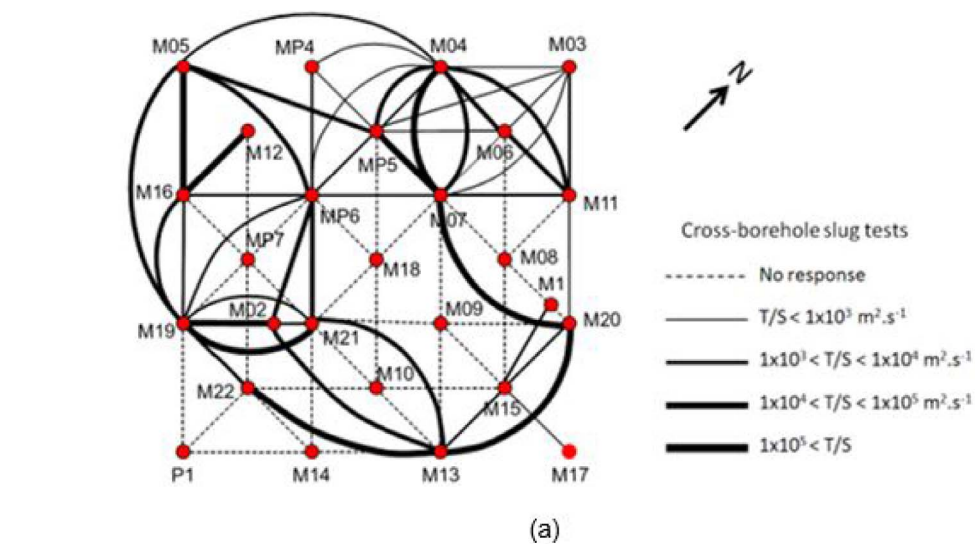
### 5.1. Slug tests

In view of the spatial distribution revealed by the seismic 3-D block regarding supposedly highly permeable bodies (Figure 5), the 35 wells regularly disposed at the HES should show variable connections between pairs of wells. It was first proceeded with a series of slugs testing individually each well, but also monitoring the response at distant wells [Audouin and Bodin, 2007, 2008]. In its classical form, a slug test investigates a very little portion of the aquifer by injecting instantaneously a small quantity of water (on the order of 100 liters) in the casing of a well [e.g., Butler Jr, 2019, Butler Jr *et al.*, 2005, Chen, 2006]. Head in the well suddenly increases and its level is monitored over time in the injected well until returning back to the initial state. Usually, the hydraulic stress generated by the injection is not strong enough to observe head responses at distant wells. In the case of highly permeable systems such as the HES, responses to slugs were also monitored at distant wells, which indicated non-negligible head variations (more than 10 cm) up to distances of 100 m. These responses allow for calculating an equivalent (homogeneous) hydraulic diffusion coefficient for 2-D radial horizontal flow between the pairs of injected and monitored wells [Audouin and Bodin, 2008]. This coefficient, as the ratio of transmissivity [ $L^2 \cdot T^{-1}$ ] to storage capacity [–], roughly measures the level of connection between wells for pressure head propagation, but not necessarily water transfer.

Results in terms of hydraulic diffusion between wells are reported in Figure 8. There can exist high contrasts of diffusivity between different close pairs of wells, or between a single well and its neighbors. The size of an elementary five spot in Figure 8a (four wells at the corners of a square, one well at the center)



**Figure 7.** Full wave acoustic logging (FWAL) at well C1: Stoneley waves. (a) Velocity log and its associated correlation coefficient log. (b) Attenuation log. (c) VSP with hydrophones.



**Figure 8.** (a) Diffusivity map evaluated from distant responses to slug tests. (b) Seismic porosity map at a depth of 85–87 m.

is 75 m, with a few wells 50 m apart from the tested well not responding to the slug tests (e.g., wells MP6, M18, M09, M15, along a W–E line in the middle of the HES) occasionally, or conversely highly connected wells with diffusion coefficients above  $10^5 \text{ m}^2/\text{s}$  over distances up to 150 m (e.g., pairs of wells M22–M13, M13–M20, M07–M20).

It is not easy to compare the map of connections (diffusivity) between wells and a seismic porosity map, for example the layer at 85–87 m depth extracted from the 3-D block (Figure 8b). At least, the distribution of hydraulic diffusivities between wells is compatible with the overall spreading of high versus low seismic porosity values, showing widespread patches of high-porosity bodies but discontinuous within the layer 85–87 m. This distribution may render pairs of wells located in the same highly porous body (e.g., wells M16, M12), resulting in high hydraulic diffusion for the pairs. Conversely, wells, even close, can be located in non-porous areas (e.g., wells M07, M09), or in separate and non-connected porous areas (e.g., M02, M21), resulting in poor to fair diffusivity values.

It is worth noting that almost all the 35 wells in the HES are open or screened boreholes, fully penetrating the whole thickness of the aquifer, thus providing “artificial” but active connections between the 3 main porous layers evidenced by seismic data. In addition, the limestone host rock is also riddled with vertical fractures of large extension, also prone to connect naturally the 3 porous layers. Both, natural and artificial connections between the layers cannot be evidenced by seismic data, their spatial resolution being not fine enough. Connections between porous layers add complexity to the interpretation of high versus low diffusivity values, because a pair of wells can be not connected to the same porous body within a given layer, but connected within another layer or via boreholes or fractures bridging two different layers (e.g., wells M04, M07 both in separate non porous bodies at 85–87 m but connected elsewhere).

## 5.2. Interference testing

Two series of interference testing were also carried out at the HES. They classically consisted of pumping a well at constant flow rate and monitoring over time the head drawdown responses at distant wells. With the number of sequentially pumped wells,

it can be considered that the series of tests allowed us for having a good image of the hydraulic response of the whole HES under forced flow conditions. Interference testing is interesting when solving the inverse groundwater flow problem, as forced flow generating important head gradients is usually more sensitive to hydraulic parameters, than, for example, a regional sweeping flow with smooth head gradients.

Several inversion exercises were undertaken to interpret interference testing. The simplest ones relied upon homogeneous or fractal cylindrical radial flow in either single-porosity or dual-porosity systems to identify from pairs of pumped-observed wells the specific storage capacity and the hydraulic conductivity of the pairs [Bernard *et al.*, 2006, Delay *et al.*, 2007, Kaczmaryk and Delay, 2007b]. Paradoxically, these studies did not reveal large variations in the sets of identified parameters. Irrespective of the single versus dual porosity assumptions employed, the hydraulic conductivities were in the range  $1\text{--}5 \times 10^{-5} \text{ m/s}$ , and (confined) specific storage capacity in  $2 \times 10^{-7}\text{--}4 \times 10^{-6} \text{ m}^{-1}$ . Nevertheless, a few values of hydraulic diffusion (the ratio of conductivity to specific storage) reaching  $10^5 \text{ m}^2/\text{s}$  suggested that non-fully-Darcian flow could occur between pairs of wells, probably conducive to inertial effects, or at least a wave-type propagation of head variations along a few flow paths [Kaczmaryk and Delay, 2007a]. It is worth noting that these high values of hydraulic diffusion were also evidenced by slug tests with distant responses (see above and Figure 8).

Incidentally, further studies also underlined lack of Lorentz reciprocity between pairs of wells: a stress in well A resulting in a response in well B, which would not correspond to the response in well A when stressing well B. With diffusive flow in a single confined porosity system, responses should always be reciprocal, irrespective of the heterogeneity of the system and its boundary conditions [Delay *et al.*, 2011, 2012]. Gaps in reciprocity at the HES are still non fully explained but could come from inertial effects in Darcian flow, the participation of both fractures (drains) and matrix in the overall flow, etc. This occurrence of reciprocity gaps remains however compatible with the idea stemming from geophysical investigations (see above) that the subsurface structural heterogeneity (the distribution of seismic porosity) is highly non-uniform and might conduct to contrasted flow patterns, locally departing from

classical Darcian flow.

### 5.3. *Inversion of spatially distributed subsurface flow*

As told earlier, with a set of interference testing covering the whole HES, and resulting in evenly distributed information (transient drawdowns), an inversion of flow at the scale of the HES can be attempted, with the aim to identify the distribution of hydraulic parameters resulting in flow simulations similar to drawdown observations. Head or drawdown observations in open wellbores usually do not see the vertical components of flow, because head values are homogenized in the wellbore and are almost uniform over its whole depth. This feature does not prohibit carrying out an inversion of flow over a three-dimensional system. But with lack of conditioning data, the distribution of hydraulic parameters, especially along the vertical direction should be highly conjectured.

Two main attempts of inverting interference testing assuming two-dimensional flow and aimed at retrieving the spatial distribution of hydraulic parameters were carried out. One attempt relied upon a forward flow model in a dual-porosity medium. By assuming that both a fractured medium and a matrix medium coexist at the same location everywhere in the system, the approach is prone to homogenize the seismic structural heterogeneity seemingly separating diverse subareas between highly-porous (flowing) patterns versus weakly-porous (non-flowing) entities. The results from the inversion of a dual-porosity aquifer (not reported here) rendered bimodal (fracture) hydraulic conductivities at either  $10^{-3}$  or  $10^{-5}$  m/s, but distributed over the whole HES as wide patches of almost uniform values [Trottier *et al.*, 2014].

The other attempt employed a forward two-dimensional flow model in a single-porosity medium. It was assumed that a single medium with contrasted hydraulic properties could render valuable flow simulations of interference testing. The resulting parameter fields should also mark contrasted regions of hydraulic behavior, as revealed by the structural seismic images. Flow in a single porosity medium is probably a slightly flawed hypothesis for grounding in the forward problem, because we already mentioned that flow was probably

not fully diffusive in a confined system (see above). This notwithstanding, by selecting a large number of data (5597 transient head values spread over the whole HES) and a parsimonious parameterization, eventual slight misconceptions in the forward model are expected not to render results worse than those associated with measurement errors on heads.

With the idea of regrouping all available flow information in a single package conditioning inversion, six interference testing were regrouped into a fictitious scenario adding sequentially over time the drawdown responses at distant wells. A relaxation period was added between the periods pumping in a single well, allowing for the water table to come back to its initial value. Raw data of drawdowns were resampled with a logarithmic time step to avoid giving more weight to short pumping times. The whole history of resampled drawdowns was introduced in a simple least-square objective function adding the squared errors on heads between simulations and observations. No regularization terms were added to the objective function, the aim being of letting the inverse problem and flow data to seek the appropriate parameter values and spatial distributions without any prior guess on them.

Parameterization for both specific storage capacity and hydraulic conductivity employed the so-called adaptive multiscale parameterization [AMT, Ackerer and Delay, 2010, Ackerer *et al.*, 2014]. The principle is to superimpose over the (forward problem) computation grid, a coarse triangular mesh, also referred to as the parameter grid. The nodes of the parameter grid hold a few local parameter values serving as seeds in any type of smooth interpolation calculating the effective parameter value in each cell (element) of the computation grid. The seed parameters are those manipulated by the optimization technique seeking the best inverse solution. During the optimization, the parameter grid can be refined locally, which also states that increasing locally the number of seeds may help to improve the depiction of parameter heterogeneity for better inverse solutions [Ackerer *et al.*, 2014]. Local refinements obey two main criteria on either the local value of errors between model outputs and data (i.e. the local value of the objective function), or the gradient components (with respect to parameters) of the objective function. With almost zero value local errors between model outputs and data, needless to refine



locally the parameter grid. Conversely, important errors would incline to refine the grid. Regarding the local components of the objective function, a stiff local gradient component indicates that the grid can be refined, when a flat component tells that the inverse problem is locally not sensitive and does not deserve further parameter grid refinement. It is worth noting that this overall parametrization procedure is left free of retrieving any spatial distribution of parameters, notably that witnessed by flow data only.

With a few refinements of the parameter grid, the number of seeds, that is, the number of master parameters sought by the inversion, may rapidly reach a few hundreds. In the present context, with almost 5600 data values, the problem is not over-parameterized. The parameterization is also easily manipulated by the optimization technique relying upon a Quasi-Newton method supplemented by the calculation of a discrete adjoint state to flow [Ackerer *et al.*, 2014] to infer the gradient components of the objective function.

As an example, Figure 9a reports on the observed and simulated drawdown curves at a few monitored wells in the sequential flowing scenario. Simulations fit fairly well the observations. Over 50 equiprobable inverse solutions were calculated (by randomizing the AMT parameterization procedure), any error on head never exceeded a few centimeters for drawdowns values between 1 and 5 m. Figure 9b is an example of inverse solution as the map of hydraulic conductivities retrieved at the scale of the computation grid. The scale of the map is not that of the seismic block and encloses a wider area surrounding the zone of interest with all the wells of the HES. This surrounding shows that the AMT parameterization technique rapidly stops its grid refinements over areas with lack of data, thus rendering almost uniform values of parameters. Here, conductivity values around the area of interest are relatively small, on the order of  $10^{-7}$ – $10^{-8}$  m/s, as a good way to “naturally” avoid noise propagated by boundary conditions (even if those are far from the HES).

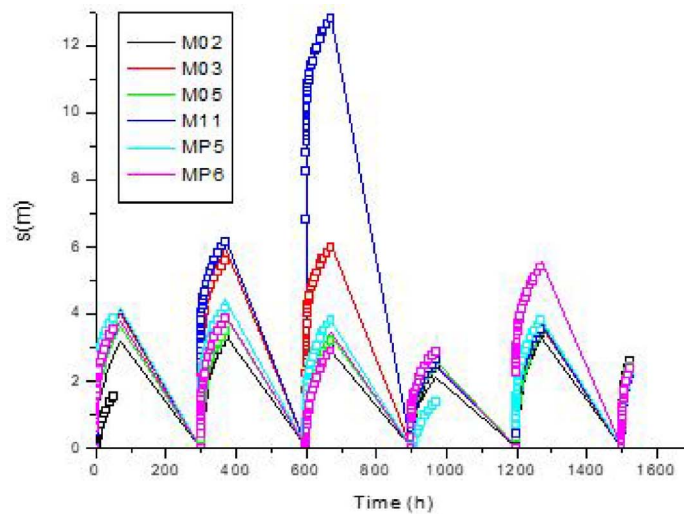
#### 5.4. Hydraulic conductivities compared with structural heterogeneity

Glancing at the hydraulic conductivity map in Figure 9b shows that the hydraulic heterogeneity revealed by inverting flow is not the structural hetero-

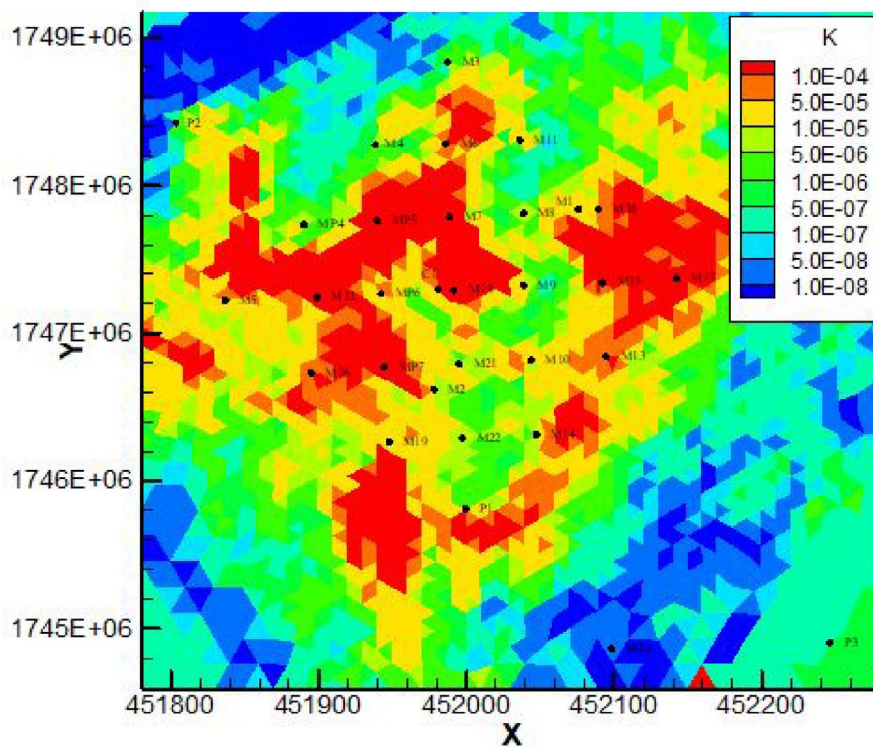
geneity revealed by seismic data. Straight at the HES, the conductivities vary between  $10^{-4}$  and  $10^{-6}$  m/s. They are distributed as large patches of convoluted contours with high conductivity to the NW and the SE of the site separated by a central area (N–S line between wells M11 and M22) of low values. For the purpose of easier comparison with the structural heterogeneity of the seismic block, a 2-D map of seismic porosity is drawn as the averaged porosity values along the vertical direction of the three porous layers encountered at 35–40, 85–87 and 110–115 m depth (Figure 10). This map is that of a high porosity layer (values between 0.15 and 0.3) uniformly bombed with low porosity (0.05) patches of 20–50 m extension. The high porosity layer, presumably highly conductive, is still continuous, and connected over the HES but populated with many small islands of weakly-conductive formations. In any case, the image rendered by the seismic porosity is not that from flow inversions as a system highly conductive over its largest part with a single weakly-conductive band in its central part.

Therefore, it must be concluded that seismic data, even collected at a scale compatible with that of flow computations (and hydraulic parameters), evidence a structural information which is not the hydraulic heterogeneity stemming from flow pattern inversion. Several reasons to this embarrassing question can be mentioned.

- Flow is usually considered as a process robust to its parameters with the meaning that hydraulic parameters must be varied over wide ranges to observe significant modifications in flow patterns. This often yields in inversion exercises to relatively smooth parameter fields with large correlation lengths [e.g., Ackerer and Delay, 2010, Trottier *et al.*, 2014]. For their part, seismic data, and especially wave velocities in porous formations, are highly sensitive to bulk rock densities variations over short distances.
- As mentioned earlier, seismic data are not sensitive, given their spatial resolution, to smaller features, as the loose network of vertical fractures riddling the subsurface at the HES, or the network of 35 open wellbores enhancing hydraulic connection between the diverse layers of the HES. Fractures and open wellbores are witnesses by flow as rapid path-

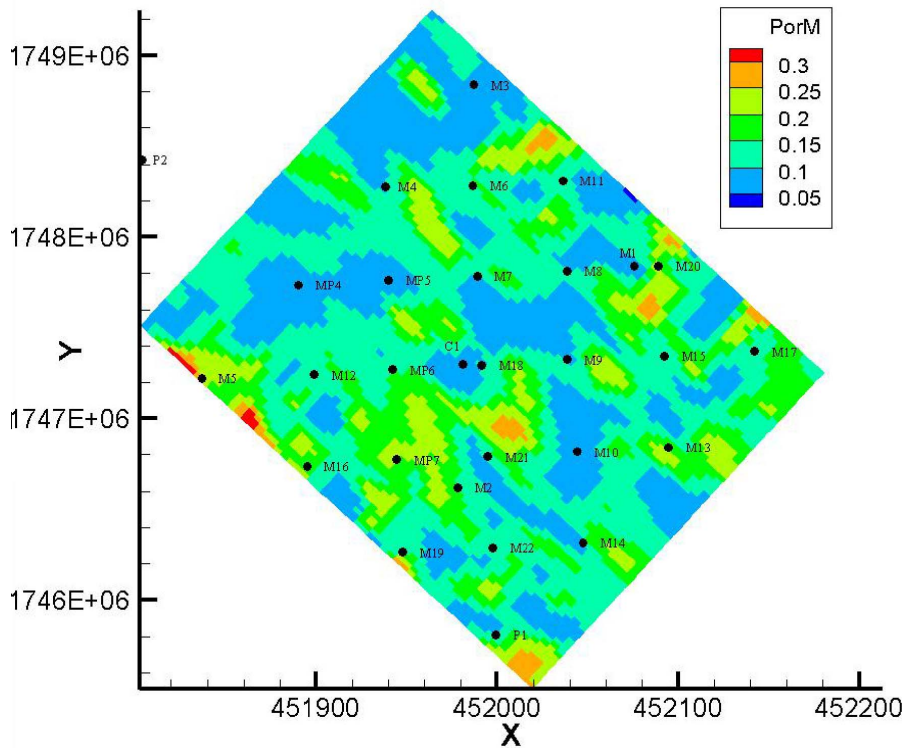


(a)



(b)

**Figure 9.** (a) Fitting of the fictitious scenario adding sequentially over time the drawdown responses from different pumping tests. Lines = simulations, dots = observations. The straight lines not fitting data mark the relaxation periods between each pumping test allowing the system to come back to its initial state. (b) Example of hydraulic conductivity field sought by inversion of the fictitious pumping scenario.



**Figure 10.** Map of porosity of the 3-D seismic block averaging over the vertical direction the values the three porous layers of the 3-D block at 35–40, 85–87, and 110 m depth.

ways, if these features are hydraulically active (not clogged).

- Seismic data, and also acoustic logs, have been geared towards detecting karstic bodies. But their detection does not predict if those bodies are fully open to flow or partly clogged by infilling materials such as sandy clay. Several works at the HES did not deny the existence of karstic bodies in a few layers of the HES. But it was mentioned that these bodies could have been initially clogged and partly reopened, especially close to the wellbores by drilling works and hydraulic investigations [e.g., Kaczmaryk and Delay, 2007a, Audouin and Bodin, 2008, Ackerer and Delay, 2010]. The partial re-opening changed the overall hydraulic behavior of the HES, increasing connectivity between wells and ensuring rapid propagation of head variations within the subsurface. With seismic data hardly making the difference be-

tween clogged and open karstic bodies, it is not illogical that the structural heterogeneity evidenced by seismics is not the effective hydraulic heterogeneity.

- As an aside, which also goes with the preceding observation, regarding connections; the HES today probably has a too dense network of wells. Within its building phase, it was ignored that the wells would crosscut layers of very contrasted hydraulic behavior. Thus, the site was designed for highly resolved hydraulic experiments and mass transfer experiments at the field scale, but carried out over reasonable durations. With 35 wells connecting highly conductive layers, it is very likely that interference testing and other hydraulic tests are partly corrupted by a rapid propagation of pressure head variations via pipes. This also applies to the 2-D modeling of flow as that used for inversions. Some prospective simulations of 3-D flow, explicitly ac-

counting for wells as conductive pipes, could probably tell us to which extent, the excess of connection by vertical wells corrupts the assumption of a large-scale two-dimensional flow.

However, a question remains unanswered. The structural heterogeneity from seismics has not been inverted regarding groundwater flow. It is well documented in the ongoing literature that the inverse problem is often ill-posed, a consequence being here that multiple solutions exist with sometimes very different patterns. It can be wondered on how the structural heterogeneity could reproduce interference testing at the HES. The seismic block could be segmented into a few classes of porosity values; those classes being each assigned a specific “facies” with its own hydraulic parameters. This would result into a small number of parameters distributed over a 3-D block, then inverted by modifying the parameter values without changing their spatial distribution. This task is still to be undertaken but takes a form out of the scope of the present work. Several options could be selected for the purpose of comparisons. For example: (1) a 3-D approach to Darcian flow in a single porosity system with at least an explicit representation of preferential flow along the wells, and inverting flow data only, or flow data and a pre-conditioning of parameters onto the seismic structure; (2) the same approach in terms of data and pre-conditioning of inversion but with dual-porosity systems or local inertial effects on Darcian flow (to account for preferential flow and eventual lack of Lorentz reciprocity).

## 6. Conclusions

We exemplified how near subsurface geophysics can be employed to image the structural heterogeneity of a karstified limestone aquifer. This aquifer is not a karst in the typical sense mainly composed of caves, open conduits etc. It is a standard porous and fractured limestone, concealing a continuous confined ground water system, but with the sparse occurrence of open karstic bodies that mainly appear along three partly dissolved layers of 2–5 m thickness at 35, 85 and 110 m depth. This type of structure is not straightforward to detect only via vertical wells. The reason is that karstic bodies are of small aperture

(0.5 m) and not continuous over each layer. Therefore, a wellbore can either crosscut an open feature or miss it, but in any case, the well does not inform on the lateral extension of the open body and its eventual connection with others.

High resolution seismic data reveal an interesting technique to image open bodies in the near subsurface, as the local wave propagation velocities are sensitive to the density of rocks, in short, their small versus large porosity. The technique is also non-intrusive, with all the investigations carried out from the surface, even if a few wells are welcome to obtain a reference (a model) of wave velocity over depth. Seismic data can be complemented by acoustic logs and vertical seismic profiles along wells; they mainly detect very small open bodies close to wells and not identified by seismic data.

3-D seismic data, eventually coupled with logs, is probably not a suitable combination to image karstic aquifers at the regional scale, to fix the ideas, over territories of several km of horizontal extension. But as shown in this study, the resulting images are of sufficient resolution for the accurate delineation of fractured or karstic bodies at a scale compatible with the security perimeter usually surrounding subsurface catchments for water supply. In view of the short transit times associated with preferential flow paths in karstified aquifers, and the subsequent risks of rapid contamination of catchments, it makes sense to envision 3-D seismic as a valuable tool for assessing the effectiveness of such perimeters.

Unfortunately, in the specific case of the HES, the structural heterogeneity revealed by 3-D seismic data is not that revealed by inversion of flow on the basis of high resolution hydraulic data from interference testing. Hydraulic heterogeneity here is of much wider correlation length separating large zones of high conduction neighboring large zones weakly connected. For its part, the 3-D seismic proposes a highly conductive field simply populated with small clusters of non-conductive areas. With the idea suggested above that karstified systems need for the assessment of mass transfer characteristics, both types of heterogeneity would not render the same flow patterns and, by the way, very different transport scenarios.

It is not completely clear why these divergences between hydraulic and structural images exist. But it is worth noting that the HES has been completely

remodeled by important drilling works and hydraulic tests under forced flow conditions. Those could have changed the diverse connections between layers in the aquifer or within the same layer, things that are not captured by seismic data.

A duplication of high resolution geophysics imaging to other hydrological contexts would be welcome. As mentioned earlier, the HES is probably plagued by an excess of connections due to the number of wells and the reopening of porous bodies by drilling works and hydraulic investigations. For example, if the delineation of an experimental site was foreseen, a prior analysis via geophysics imaging, which could also include electrical resistivity tomography (ERT) and magnetic resonance sounding (MRS) could be attempted. A second campaign of the same measurements after finalization of the diverse facilities at the site could tell us how geophysical imaging and incidentally the hydraulic behavior are sensitive to our “invasive” fingerprint in the system.

This study posed the question of how near subsurface geophysics could help groundwater flow inversion. The answer is that geophysics at the very least reveals new ideas on the subsurface heterogeneity and should motivate hydrogeologists to widen their conceptual models (families) of inverse solutions. In 1999, G. de Marsily *et al.* published a paper on “Forty years of inverse problems in Hydrogeology”. More than twenty years later, sixty years of inverse problems let the question open.

## Conflicts of interest

Authors have no conflict of interest to declare.

## References

- Ackerer, P. and Delay, F. (2010). Inversion of a set of well-test interferences in a fractured limestone aquifer by using an automatic downscaling parameterization technique. *J. Hydrol.*, 389, 42–56.
- Ackerer, P., Trottier, N., and Delay, F. (2014). Parameter estimation using an adaptive multiscale method. *Adv. Water Resour.*, 73, 108–122.
- Archie, G. E. (1942). The electrical resistivity log as an aid in determining some reservoir characteristics. *Pet. Technol.*, 146, 54–62.
- Audouin, O. and Bodin, J. (2007). Analysis of slug tests with high-frequency oscillations. *J. Hydrol.*, 334(1–2), 282–289.
- Audouin, O. and Bodin, J. (2008). Cross-borehole slug test analysis in a fractured limestone aquifer. *J. Hydrol.*, 348, 510–523.
- Bernard, S., Delay, F., and Porel, G. (2006). A new method of data inversion for the identification of fractal characteristics and homogenization scale from hydraulic pumping tests in fractured aquifers. *J. Hydrol.*, 328, 647–658.
- Butler Jr, J. J. (2019). In *The Design, Performance, and Analysis of Slug Tests*, page 280. CRC Press, Boca Raton, 2nd edition.
- Butler Jr, J. J., Garnett, E. J., and Healey, J. M. (2005). Analysis of slug tests in formations of high hydraulic conductivity. *Ground Water*, 41(5), 620–631.
- Chen, C. S. (2006). An analytical data analysis method for oscillatory slug tests. *Ground Water*, 44(4), 604–608.
- de Marsily, G., Delhomme, J.-P., Delay, F., and Buoro, A. (1999). Regards sur 40 ans de problèmes inverses en hydrogéologie. *C. R. Acad. Sci. Earth Planet. Sci.*, 329, 73–87.
- Delay, F., Ackerer, P., Belfort, B., and Guadagnini, A. (2012). On the emergence of reciprocity gaps during interference testing in unconfined aquifers. *Adv. Water Resour.*, 46, 11–19.
- Delay, F., Ackerer, P., and Guadagnini, A. (2011). Theoretical analysis and field evidence of reciprocity gaps during interference pumping tests. *Adv. Water Res.*, 34, 592–606.
- Delay, F., Kaczmaryk, A., and Ackerer, P. (2007). Inversion of interference hydraulic pumping tests in both homogeneous and fractal dual media. *Adv. Water Res.*, 30, 314–334.
- Faust, L. Y. (1953). A velocity function including lithologic variation. *Geophysics*, 18, 271–288.
- Hardage, B. A. (1984). In *Vertical Seismic Profiling, Part A: Principles*, page 419. Geophysical Press, London. ISBN-10: 094663100X.
- Kaczmaryk, A. and Delay, F. (2007a). Improving dual-porosity-medium approaches to account for karstic flow in a fractured limestone. Application to the automatic inversion of hydraulic interference tests. *J. Hydrol.*, 347, 391–403.
- Kaczmaryk, A. and Delay, F. (2007b). Interpretation of interference pumping tests in fractured limestone by means of dual-medium approaches. *J. Hydrol.*, 2007(337), 133–146.
- Mari, J.-L. and Delay, F. (2011). Contribution of seismic and acoustic methods to reservoir

- model building. In *Hydraulic Conductivity/Book 1*. InTech-Open Access Publisher. ISBN: 978-953-307-288-3.
- Mari, J.-L. and Mendes, M. (2012). High resolution near surface imaging of fracture corridors and cavities by combining Plus Minus method and refraction tomography. *Near Surf. Geophys.*, 10, 185–195.
- Mari, J.-L. and Porel, G. (2008). 3D seismic imaging of a near—surface heterogeneous aquifer: a case study. *Oil Gas Sci. Technol.*, 63, 179–201.
- Mari, J.-L., Porel, G., and Bourbiaux, B. (2009). From 3D seismic to 3D reservoir deterministic model thanks to logging data: The case study of a near surface heterogeneous aquifer. *Oil Gas Sci. Technol.*, 64, 119–131.
- Mari, J.-L., Porel, G., and Delay, F. (2020). Contribution of full wave acoustic logging to the detection and prediction of karstic bodies. *Water*, 12(4), article no. 948.
- Mari, J.-L. and Vergniault, C. (2018). In *Well Seismic Surveying and Acoustic Logging*, page 135. EDP Sciences Publisher.
- Trottier, N., Delay, F., Bildstein, O., and Ackerer, P. (2014). Inversion of a dual-continuum approach to flow in a karstified limestone: Insights to aquifer heterogeneity revealed by well test interferences. *J. Hydrol.*, 508, 157–169.

# **Surface Hydrology**







Research article

Geo-hydrological Data & Models

# On the computation of the Basal Envelope Surface of Talwegs using the Analytic Element Method

Nicolas Le Moine <sup>a</sup>

<sup>a</sup> UMR 7619 Metis, Sorbonne Université/CNRS/EPHE, 4 Place Jussieu, 75252 Paris cedex 05, France

E-mail: [nicolas.le\\_moine@sorbonne-universite.fr](mailto:nicolas.le_moine@sorbonne-universite.fr)

**Abstract.** The stream network is a major feature of a landscape, conveying water, sediment, and solute from hillslopes to the ocean. Noticeably, from a large-scale point of view, the elevation of the talwegs of perennial streams is an important head boundary condition for both surface and groundwater flow originating from hillslopes. Assuming a wireframe (1D) representation of talweg lines, the problem of interpolating elevation between talwegs has received attention for applications such as flood mapping using Height Above Nearest Drainage (HAND, Nobre et al. [2011]), or groundwater level interpolation in low-conductivity aquifer systems. In this study we propose an alternate definition of this large-scale base level concept introduced by Wyns et al. [2004], namely the Basal Envelope Surface of Talwegs (BEST) and the associated Height Above the Basal Envelope Surface of Talwegs (HABEST), along with a procedure to compute it using the Analytic Element Method (AEM). It can be defined as the head distribution satisfying Laplace equation (Darcy flow with vanishing divergence), with stream segments set as Dirichlet boundary conditions. The BEST is thus the real part of a complex analytic (holomorphic) function which can be modeled using analytic slit elements, with very low computational requirements and without the need for kriging, as it is often seen in the literature. This analytic model is extended to the case of a non-zero, uniform divergence flow (head distribution satisfying a Poisson equation) which can be useful to analyse groundwater levels at catchment scale.

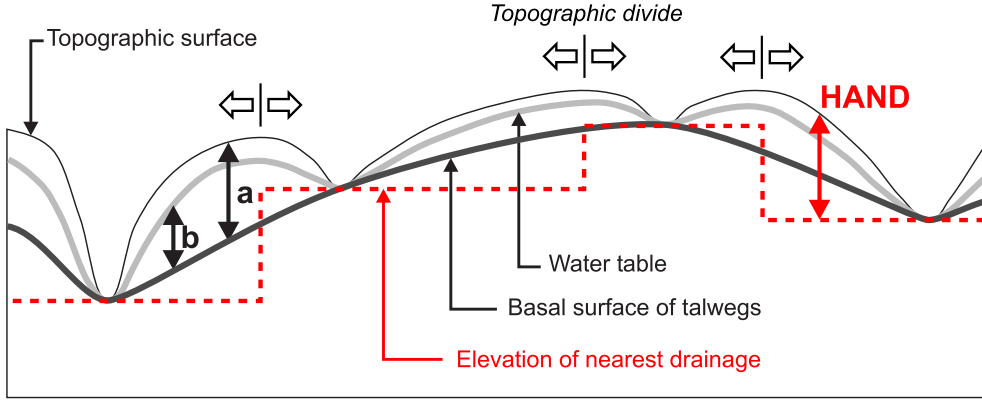
**Keywords.** Analytic Element Method, Envelope Surface, Talwegs, Geomorphology, Groundwater, Hillslope, Diffusive processes.

Manuscript received 1 April 2022, revised 1 August 2022 and 13 September 2022, accepted 14 September 2022.

## 1. Introduction

It is well known since the pioneering work of Dunne [1969, 1980] and Dunne and Black [1970] that saturated groundwater flow (both subsurface and deep) plays an important role in forming and sustaining channelized flow, especially in humid environments. From this perspective, neither can streams be considered as merely fixed (Dirichlet) boundary conditions for the groundwater system, nor can the groundwater flow feeding the streams be considered a mere “base-flow contribution”: the energy level in the aquifer system and in the streams are strongly related for all flow conditions.

Wyns et al. [2004] showed that in low conductivity, shallow aquifer systems (such as weathered crystalline basement settings), there is a strong relation between the groundwater level at a given location, and the drop in topographic elevation from this location to the nearest wet talweg. More precisely, there is a correlation between the height above a *basal envelope surface of talwegs* (height labeled *a* on Figure 1), and the height of the groundwater level above this same envelope surface (height labeled *b*). In such systems, the low hydraulic conductivity implies a steepening of groundwater head gradients in order to convey groundwater from hillslopes to channels.



**Figure 1.** Illustration of different potentiometric levels on a topographic/groundwater transect, adapted from Wyns et al. [2004] in order to show both the *Basal Surface of Talwegs* and the *Elevation of Nearest Drainage* of Nobre et al. [2011].

Obviously, the steepest gradients are achieved when the groundwater potentiometric surface is close to the topographic surface, i.e.

$$\frac{b}{a} \lesssim 1.$$

Interestingly, a parallel concept has been introduced in the field of flood mapping, namely the *Height Above Nearest Drainage* (HAND) of Nobre et al. [2011]. There are a couple of differences between the two relative elevations defined respectively by the *Height Above the Basal Envelope Surface of Talwegs* (HABEST) and the HAND. These differences are illustrated in Figure 1:

- the HAND computation procedure relies on a hydrologically-conditioned Digital Elevation Model (DEM) in which a reference drainage network is identified using a minimum support area. It then requires a downstream drainage allocation of each pixel, so that it creates a jump at the crossing of a water divide when two pixels on either side of the divide are “snapped” to two distinct drainages. Put differently, the reference potentiometric surface from which the HAND is computed is not continuous (we should call it the *Elevation of Nearest Drainage*, though it is not explicitly introduced by Nobre et al. [2011]).
- In contrast, the surface constructed following Wyns et al. [2004] is continuous. It is produced by interpolating point-scale elevation data in the talweg network.

From the point of view of free surface (channel) flow, a surface of minimum potential energy joining all the talwegs is of course, an important reference level: the transient rise of water during a flood event corresponds to a relative elevation above the “normal” drainage level, i.e. a local HAND value [Nobre et al., 2016].

Since the computation of such a base level has several applications, there is a need to give a formal definition and an unambiguous procedure to compute it.

## 2. A formal definition of the Basal Envelope Surface of Talwegs (BEST)

Let us consider the vertically-integrated, transient diffusivity equation for groundwater flow with the Dupuit–Forchheimer assumption [e.g., De Marsily, 1986]:

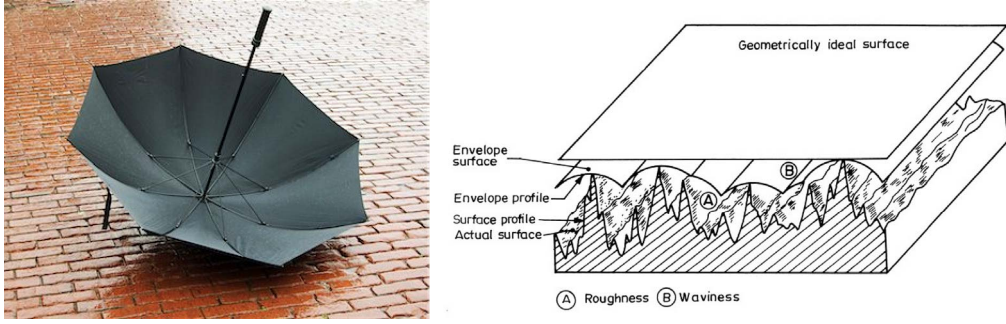
$$S \frac{\partial h}{\partial t} + \vec{\nabla} \cdot \vec{q} = r \quad \text{with} \quad \vec{q} = -T(h) \vec{\nabla} h \quad (1)$$

where  $h$  is the groundwater head,  $S$  the storage coefficient,  $\vec{q}$  the 2D (vertically-integrated) Darcy flux density,  $T$  the transmissivity ( $\text{m}^2 \cdot \text{s}^{-1}$ ), and  $r$  the (positively defined) recharge rate ( $\text{m} \cdot \text{s}^{-1}$ ). Hence,

$$\vec{\nabla} \cdot (T(h) \vec{\nabla} h) = S \frac{\partial h}{\partial t} - r(x, y, t). \quad (2)$$

Assuming a uniform or slowly-varying transmissivity across the domain, we can take  $T$  out of the divergence operator. In the following, we will assume a strictly uniform transmissivity  $T_0$ , hence:

$$T_0 \Delta h = S \frac{\partial h}{\partial t} - r(x, y, t). \quad (3)$$



**Figure 2.** (Left) Example of an envelope surface, in the form of the canopy of an umbrella; (right) definition by Häsing [1964].

We can consider a hypothetical situation where the groundwater level reaches a minimum ( $\partial h / \partial t \approx 0$ ) in conjunction with a vanishing recharge rate ( $r(x, y, t) \approx 0$ ), with the perennial streams still connected with the shallow aquifer. The head distribution would then satisfy:

$$\begin{cases} \Delta h = \frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} = 0 \\ h(x, y) = z_{\text{topo}}(x, y) \quad \text{if } (x, y) \text{ is in the stream network} \end{cases} \quad (4)$$

where  $z_{\text{topo}}$  denotes the elevation of the topographic surface. Petroff et al. [2012] as well as Cohen et al. [2015] have studied in great detail the interaction of a stream network with a Poisson diffusive field in the more general case of a nonlinear transmissivity  $T(h) = K \cdot (h - B)$  for an unconfined aquifer, where  $B(x, y)$  denotes the elevation of the underlying impermeable layer and  $K$  the saturated hydraulic conductivity of the aquifer. In the following, we will restrict ourselves to the linear case, as it allows simple analytic developments based on the superposition principle. Here, the condition  $\Delta h = 0$  means that the flux density  $\vec{q}$  is assumed to have a vanishing divergence under hillslopes (it is a solenoidal vector field). The resulting piezometric surface is a minimum-curvature surface joining the 3-dimensional talweg lines, in a way similar to the gores of an umbrella's canopy stretched over the ribs (Figure 2). Such a surface is an envelope surface in the sense of Häsing [1964], which can be defined in several ways.

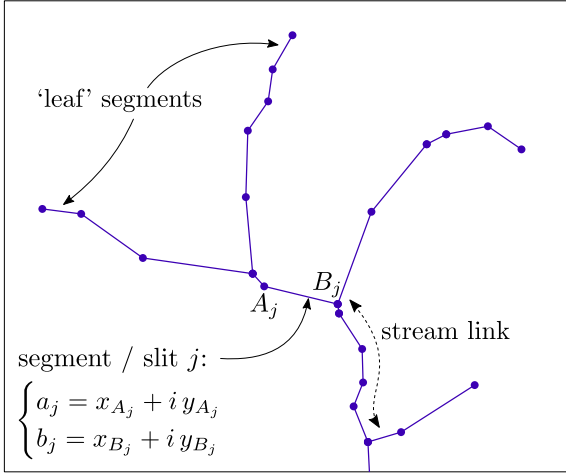
Usually [see e.g., Deffontaines, 1990, Wyns et al., 2004, Yao et al., 2017], this surface is constructed by kriging of scattered talweg elevation data: a more straightforward way of constructing it is to

solve (4). In fact, under the assumption of uniform transmissivity we can directly model the analytic function using the Analytic Element Method [AEM, see e.g. Strack, 1989, Steward, 2020], instead of solving the steady-state equilibrium equation numerically with stream elevations as Dirichlet boundary conditions. As will be shown further, the procedure boils down to a simple linear least-square optimization, without the need for variography.

### 3. Mathematical modeling with analytic slit elements

#### 3.1. Vector representation of the talweg graph

The approach proposed in this study relies on a vector representation of the talweg or stream network. Such representations are usually provided by national geographic agencies, as it is the case for the BD Topage in France [Système d'Information sur l'Eau/Sandre, 2020], but can also be derived from the analysis of Digital Elevation Models (DEM). The network is a set of *links*, i.e., sections of a stream channel connecting two successive junctions, a junction and the outlet, or a junction and the drainage divide [ESRI, 2022]. We assume that the topology of the links is also available in the form of an oriented graph. This general setting is illustrated in Figure 3. Each link is a polyline that can be decomposed into segments, each defined by a start point  $A_j$  and an end point  $B_j$ .



**Figure 3.** Vector representation of the talweg or stream network. It is a collection of elementary segments/slits, each defined by two endpoints which can be positioned in the complex  $\mathbb{C}$ -plane with their complex affix.

### 3.2. General form of the potential

In AEM, the head distribution is related to a real-valued potential  $\Phi$  [e.g., De Marsily, 1986], such that:

$$\vec{q} = -\vec{\nabla}\Phi.$$

For unconfined aquifers such as the shallow groundwater systems developed in weathered bedrock, our focus in this study, the general form of the flux is:

$$\vec{q} = -K \cdot (h - B) \vec{\nabla} h.$$

In the case of a piecewise uniform elevation  $B$  of the substratum (i.e.  $\vec{\nabla} B = \vec{0}$  by subdomains), we have the integral form for the potential [e.g., Steward, 2020]:

$$\begin{aligned} \vec{\nabla}\Phi &= -\vec{q} = +K \cdot (h - B) \vec{\nabla} h \\ \vec{\nabla}\Phi &= K \cdot (h - B) \vec{\nabla} (h - B) = \frac{K}{2} \vec{\nabla} (h - B)^2 \\ \Rightarrow \Phi &= \frac{K}{2} (h - B)^2. \end{aligned}$$

This expression hence leads to a nonlinear relation between potential  $\Phi$  (in  $\text{m}^3/\text{s}$ ) and head  $h$  (in m). In the following, we will linearize the problem and consider a uniform transmissivity  $T_0$  of the groundwater system, independent of  $h$ . This linearization is motivated by several considerations:

- first, as weathering develops from the surface down, it is no stronger an assumption to consider a rather uniform *depth* for the lower limit of the transmissive layer, than to consider a piecewise uniform *elevation* for this limit; this amounts to assume a rather uniform saturated thickness at first order, if the groundwater table is always close to the surface;
- second, our objective is to use observed groundwater levels in boreholes as well as topographic elevation of streams in order to infer large-scale hydrodynamic properties of the system such as, precisely, an “effective”, large-scale estimate of transmissivity. This is much easier to do when using groundwater head  $h$  directly as a proxy, linearly related to potential  $\Phi$ : it will be clear when we present the degrees of freedom used to compute the value of potential  $\Phi$  using AEM in the next paragraphs.

With the assumption of a uniform transmissivity  $T_0$ , the potential  $\Phi$  is indeed linearly related to groundwater head  $h$ :

$$\vec{\nabla}\Phi \stackrel{\text{def}}{=} T_0 \vec{\nabla} h \stackrel{\text{def}}{=} +T_0 \vec{\nabla} h \stackrel{\text{def}}{=} \text{constant} \cdot \vec{\nabla} h \quad (5)$$

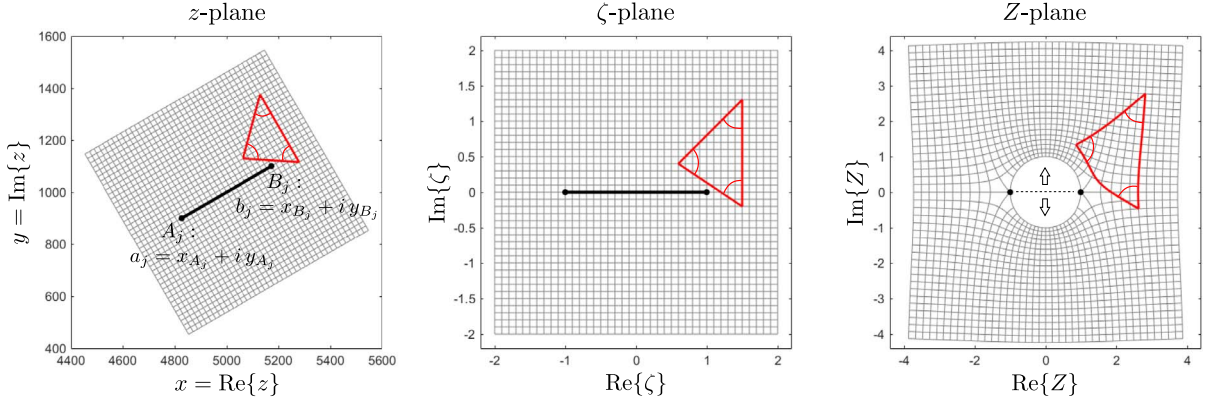
In order to model the envelope surface, we will work in the complex plane  $\mathbb{C}$ . The start- and end-point of each segment  $j$  in the vector talweg network,  $A_j$  and  $B_j$ , are first given a complex affix in the  $\mathbb{C}$ -plane (see Figure 3):

$$\begin{cases} a_j = x_{A_j} + i y_{A_j} & \text{affix of start point } A_j \text{ of} \\ & \text{element } j \text{ in the complex plane} \\ b_j = x_{B_j} + i y_{B_j} & \text{affix of end point } B_j \text{ of} \\ & \text{element } j \text{ in the complex plane} \end{cases} \quad (6)$$

with  $i$ , a number such that  $i^2 = -1$ . Hereafter, the letter  $z$  will denote an affix  $z = x + i y$  in the complex plane (with the exception of  $z_{\text{topo}}$  which will denote the topographic elevation). In the case of a divergence-free flux density  $\vec{q}$ , the real-valued potential  $\Phi$  is the real part of a complex-valued potential  $\Omega$ :

$$\Omega(z) = \Phi(x, y) + i\Psi(x, y)$$

where  $\Psi$  is the associated, real-valued *stream function*. The complex potential  $\Omega(z)$ , which is defined in the complex plane  $\mathbb{C}$  and has values also in  $\mathbb{C}$ , has a



**Figure 4.** Illustration of the transformation  $Z_j(z)$  associated to slit element  $j$ , with endpoints  $A_j$  and  $B_j$ . It is obtained by the composition of a similarity  $\zeta_j(z)$ , and the Joukowski transformation  $Z_j(z) = \zeta_j(z) + \sqrt{\zeta_j(z) + 1} \sqrt{\zeta_j(z) - 1}$ . Both transformations are 2D *conformal mappings* or *holomorphisms*: they preserve angles locally (see red triangle), hence orthogonality (see grid).

distinct property: it is *analytic* or *holomorphic*, which means that the derivative of  $\Omega$  with respect to the

complex variable  $z$  is defined (this is not true in general for a function from  $\mathbb{C}$  to  $\mathbb{C}$ ):

$$\left. \frac{d\Omega}{dz} \right|_{z_0 \in \mathbb{C}} = \lim_{z \rightarrow z_0} \frac{\Omega(z) - \Omega(z_0)}{z - z_0} \quad \text{exists: does not vary depending on the direction from which } z_0 \text{ is approached.}$$

Synonymously, it means that the real potential  $\Phi$  and the stream function  $\Psi$  satisfy the Cauchy–Riemann conditions:

$$\left\{ \begin{array}{ll} \frac{\partial \Phi}{\partial x} = \frac{\partial \Psi}{\partial y} & \text{(i)} \\ \frac{\partial \Phi}{\partial y} = -\frac{\partial \Psi}{\partial x} & \text{(ii)} \\ \frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} = \Delta \Phi = 0 & \text{(iii), consequence of (i) and (ii); the actual property we want for } \Phi \\ \frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} = \Delta \Psi = 0 & \text{(iv), consequence of (i) and (ii).} \end{array} \right.$$

Contours of the stream function define streamlines; the area between two streamlines defines a stream tube, in which the flux is conserved.

Following the approach of Steward [2015, 2020], we will then consider each segment  $j$  of the talweg network as a *slit* element creating a complex potential  $\Omega_j(z)$  at location  $z$ , with the form:

$$\Omega_j(z) = Q_j \times \left( \frac{\ln Z_j(z)}{2\pi} - 1 \right) + \sum_{k=1}^{n_{\text{coef}}} c_{j,k} Z_j^{-k}(z). \quad (7)$$

$Q_j$  is the net algebraic discharge (in  $\text{m}^3/\text{s}$ ) entering the slit, and the  $c_{j,k}$  are the complex coefficients of a Laurent series that can be adjusted to match various boundary conditions. If  $Q_j > 0$ , the slit is a sink to the system (it subtracts water from the aquifer) while if  $Q_j < 0$ , it is a source (it leaks water to the aquifer). The mapping function  $Z_j(z)$  for each slit element is given by:

$$\zeta_j(z; a_j, b_j) = \frac{z - \frac{1}{2}(a_j + b_j)}{\frac{1}{2}(b_j - a_j)} \quad (8)$$

$$Z_j(z) = \zeta_j(z) + \sqrt{\zeta_j(z) + 1} \sqrt{\zeta_j(z) - 1}. \quad (9)$$

The function  $Z(\zeta) = \zeta + \sqrt{\zeta + 1} \sqrt{\zeta - 1}$  is the Joukowski transformation [Joukowski, 1910]. Figure 4 illustrates the two steps of the transformation  $Z_j(z) = \zeta_j(z) + \sqrt{\zeta_j(z) + 1} \sqrt{\zeta_j(z) - 1}$ . The first step (mapping from  $z$ -plane to  $\zeta$ -plane, Equation (8)) is just a similarity, i.e. a combination of translation, rotation and scaling that sends the two endpoints  $A_j$  and  $B_j$  of slit  $j$  to affixes  $-1$  and  $+1$  respectively. Then, the Joukowski transformation (from  $\zeta$ -plane to  $Z$ -plane, Equation (9)) maps the  $\mathbb{R} \times \mathbb{R}$  plane to the *outside* of the unit circle: it amounts to “pop-opening” the slit segment  $[A_j B_j]$  *conformally*, i.e. with a transformation that conserves angles.

By superposition, the total (complex) potential at location  $z$  in the complex plane is simply the sum of the contributions of all slit elements in the domain:

$$\Omega_{\text{tot}}(z) = \Phi_{\text{tot}}(z) + i\Psi_{\text{tot}}(z) = \Phi_0 - \overline{q_0} z + \sum_{j=1}^{n_{\text{slit}}} \Omega_j(z) \quad (10)$$

where  $q_0 = q_{0,x} + i q_{0,y}$  is a complex number giving the background, regional direction of the flow (“tilt” in the groundwater table; since it is uniform, it does not add any divergence to the flux).  $\overline{q_0} = q_{0,x} - i q_{0,y}$  is its complex conjugate, and  $\Phi_0$  an offset value for the real potential  $\Phi = \text{Re}\{\Omega\}$ .

For each element  $j$ , we have  $(2n_c + 1)$  real degrees of freedom: the sink term  $Q_j$ , the  $n_c$  real parts  $\text{Re}\{c_{j,k}\}$  and  $n_c$  imaginary parts  $\text{Im}\{c_{j,k}\}$  with  $c_{j,k} = \text{Re}\{c_{j,k}\} + i\text{Im}\{c_{j,k}\}$ . Together with  $\Phi_0$ ,  $v_{0,x}$ , and  $v_{0,y}$ , the total number of real degrees of freedom is thus  $N = 3 + n_s \times (2n_c + 1)$ .

In Figure 5 we show the general form of the potential for three cases:

- a pure sink element with  $Q \neq 0$  and all terms set to zero in the Laurent series; far from the slit, the real potential created by such an element is identical to the well known Thiem

solution  $h = \Phi/T_0 = (Q/2\pi T_0) \ln r + \text{constant}$  [e.g., De Marsily, 1986],

- the potential  $\Omega(z) = c_1 Z^{-1}(z)$  with a real-valued  $c_1$  coefficient. In this case, the real potential  $\Phi$  (gray levels) is continuous,
- the potential  $\Omega(z) = c_1 Z^{-1}(z)$  with a pure imaginary  $c_1$  coefficient. In this case, the real potential  $\Phi$  is discontinuous across the slit.

The two last situations do not produce any net inflow nor outflow.

Of course, in the general case we have no idea of the transmissivity of the porous medium; in order to work directly with head levels/drops we will normalize the expression of the complex potential by the (unknown) uniform transmissivity  $T_0$ :

$$\begin{aligned} \Omega'_{\text{tot}}(z) &= \frac{1}{T_0} \Phi_{\text{tot}}(z) + i \frac{1}{T_0} \Psi_{\text{tot}}(z) = \Phi'_0 - \overline{q'_0} z \\ &+ \sum_{j=1}^{n_{\text{slit}}} \left[ Q'_j \times \left( \frac{\ln Z_j(z)}{2\pi} - 1 \right) + \sum_{k=1}^{n_{\text{coef}}} c'_{j,k} Z_j^{-k}(z) \right] \end{aligned} \quad (11)$$

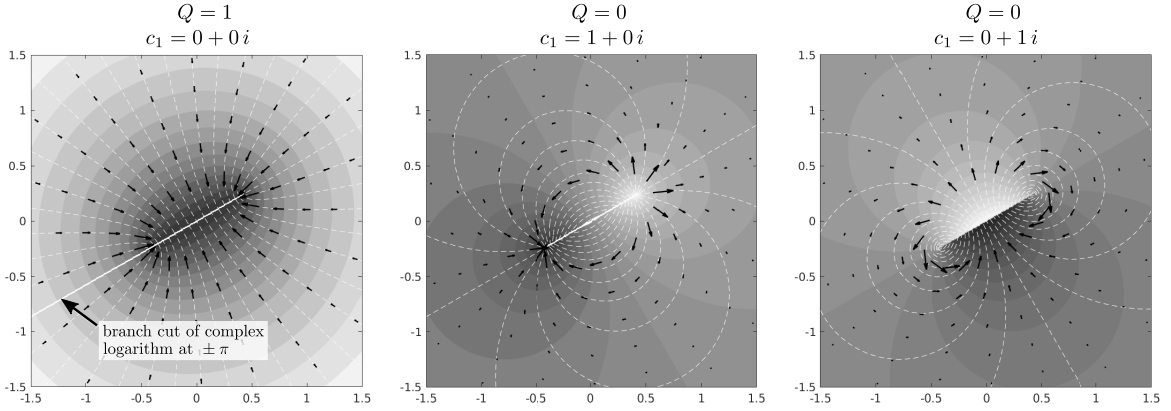
with the normalized quantities:

$$\begin{cases} Q'_j = \frac{Q_j}{T_0} & \text{normalized discharge extracted by} \\ & \text{slit } j \text{ (depth or drawdown in m)} \\ \Phi'_0 = \frac{\Phi_0}{T_0} \\ q'_0 = \frac{1}{T_0} q_0 \\ c'_{j,k} = \frac{1}{T_0} c_{j,k}. \end{cases}$$

### 3.3. Optimization procedure

Given a set of control points  $\{z_p\}_{1 \leq p \leq M}$  where we want to specify a value  $h_p = h_{\text{tot}}(z_p) = \text{Re}\{\Omega'_{\text{tot}}(z_p)\}$  of the groundwater head, we can build a linear system to estimate the optimum value of all degrees of freedom. With the assumption of a uniform transmissivity,  $h$  is linear with respect to all degrees of freedom (11) so the regression matrix is simply the  $[M \times N]$  Jacobian matrix:

$$J_{p,l} = \text{Re} \left\{ \frac{\partial \Omega'_{\text{tot}}(z_p)}{\partial \beta_l} \right\}$$



**Figure 5.** General form of the potential created by a slit element. In each figure, the gray levels show the value of  $\Phi = \text{Re}\{\Omega(z)\}$  while the dashed white lines are contours of the stream function  $\Psi = \text{Im}\{\Omega(z)\}$ . Black arrows show the flux vector  $q = -\overline{d\Omega/dz}$ . (Left) Potential created by a pure sink. (Center) Potential created by a single, real-valued coefficient in the Laurent series; note that the real potential  $\Phi$  is continuous between the “top” and “bottom” of the slit, while the stream function  $\Psi$  is discontinuous. (Right) Potential created by a single, pure imaginary coefficient in the Laurent series; there is now a jump in the real potential  $\Phi$  between the two sides of the slit while the stream function, in turn, is continuous.

where  $\beta_l$  is the  $l$ -th real degree of freedom. The entry of the matrix on row  $p$  and column  $l$  reads as:

$$\begin{cases} J_{p,l} = 1 & \text{if } \beta_l = \Phi'_0 \\ J_{p,l} = -\text{Re}\{z_p\} & \text{if } \beta_l = q'_{0,x} \\ J_{p,l} = +\text{Im}\{z_p\} & \text{if } \beta_l = q'_{0,y} \\ J_{p,l} = \text{Re}\left\{\frac{\ln Z_j(z_p)}{2\pi}\right\} - 1 & \text{if } \beta_l = Q'_j \\ J_{p,l} = +\text{Re}\{Z_j^{-k}(z_p)\} & \text{if } \beta_l = \text{Re}\{c'_{j,k}\} \\ J_{p,l} = -\text{Im}\{Z_j^{-k}(z_p)\} & \text{if } \beta_l = \text{Im}\{c'_{j,k}\}. \end{cases} \quad (12)$$

The control points are positioned along the segments of the stream network. Since there are  $2n_c + 1$  degrees of freedom for each slit, the system is overdetermined if we position more control points than this value along each segment. The vector of optimum values for the parameters (the  $\beta_l$ 's) is the least squares solution to:

$$\mathbf{J} [\beta_1, \beta_2, \dots, \beta_N]^T = [\Phi'_1, \Phi'_2, \dots, \Phi'_M]^T.$$

Instead of Dirichlet conditions at control points, we can also specify Neumann conditions on the normalized velocity vector

$$q = q_x + iq_y = -\left[\frac{\partial\Phi}{\partial x} + i\frac{\partial\Phi}{\partial y}\right] = -\overline{\frac{d\Omega}{dz}}.$$

Using the chain rule

$$\frac{d\Omega_j}{dz} = \frac{d\Omega_j}{dZ_j} \frac{dZ_j}{d\zeta_j} \frac{d\zeta_j}{dz},$$

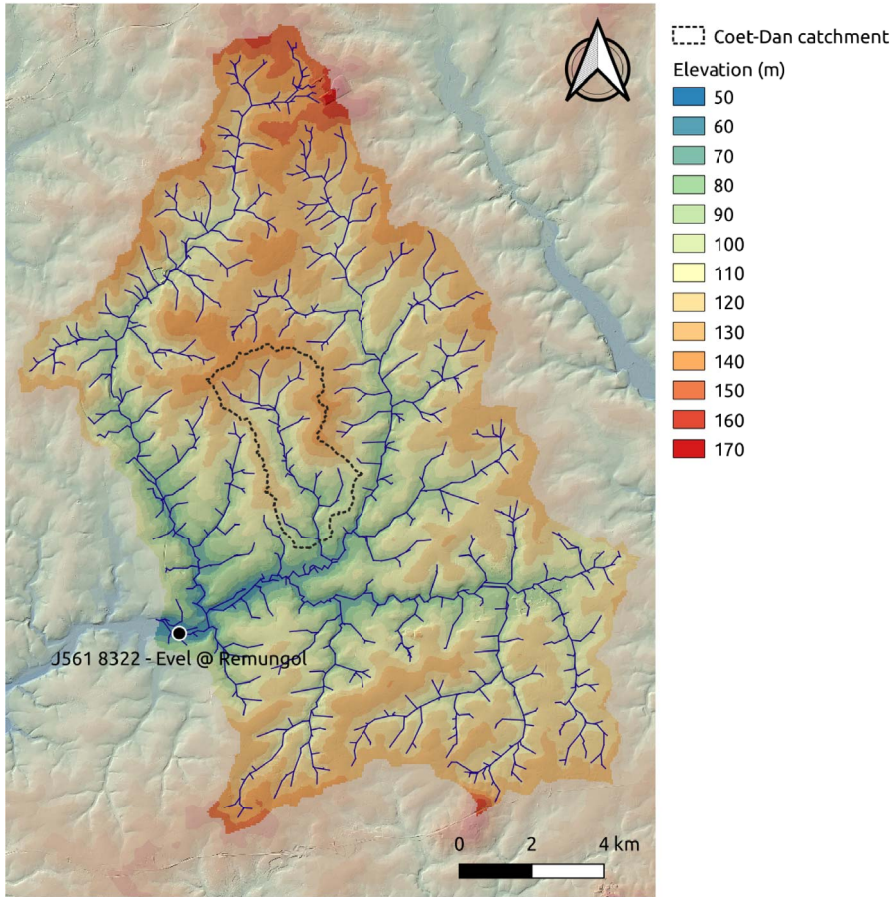
the contribution of each slit to the total flux hence reads:

$$\begin{aligned} q_j(z) &= -\overline{\frac{d\Omega_j}{dz}} \\ &= \frac{2}{(b_j - a_j)\sqrt{\zeta_j(z) + 1}\sqrt{\zeta_j(z) - 1}} \left[ -\frac{Q_j}{2\pi} + \sum_{k=1}^{n_{\text{coef}}} c_{j,k} k Z_j^{-k}(z) \right] \\ q_j(z) &= \frac{2}{(b_j - a_j)\sqrt{\zeta_j(z) + 1}\sqrt{\zeta_j(z) - 1}} \\ &\quad \times \left[ -\frac{Q_j}{2\pi} + \sum_{k=1}^{n_{\text{coef}}} (\text{Re}\{c_{j,k}\} - i\text{Im}\{c_{j,k}\}) k \overline{Z_j^{-k}(z)} \right] \end{aligned}$$

or, in the normalized form:

$$\begin{aligned} q'_j(z) &= \frac{2}{(b_j - a_j)\sqrt{\zeta_j(z) + 1}\sqrt{\zeta_j(z) - 1}} \\ &\quad \times \left[ -\frac{Q'_j}{2\pi} + \sum_{k=1}^{n_{\text{coef}}} (\text{Re}\{c'_{j,k}\} - i\text{Im}\{c'_{j,k}\}) k \overline{Z_j^{-k}(z)} \right]. \end{aligned}$$

This expression is again linear with respect to all real degrees of freedom.



**Figure 6.** Map of the Upper Evel catchment and the nested Coët-Dan subcatchment. The vector stream network is extracted with a support area of  $0.1 \text{ km}^2$ .

**Table 1.** List of the degrees of freedom in the least squares problem

Type of elements	Number of elements	Number of sink terms $Q'_j$	Number of $\text{Re}\{c'_{j,k}\}$	Number of $\text{Im}\{c'_{j,k}\}$	Number of degrees of freedom
Stream segm.	1562	0	$2 \times 1562$	0	3124
Boundary segm.	176	0	$2 \times 176$	$2 \times 176$	704
				Total incl. $\Phi'_0$	3829

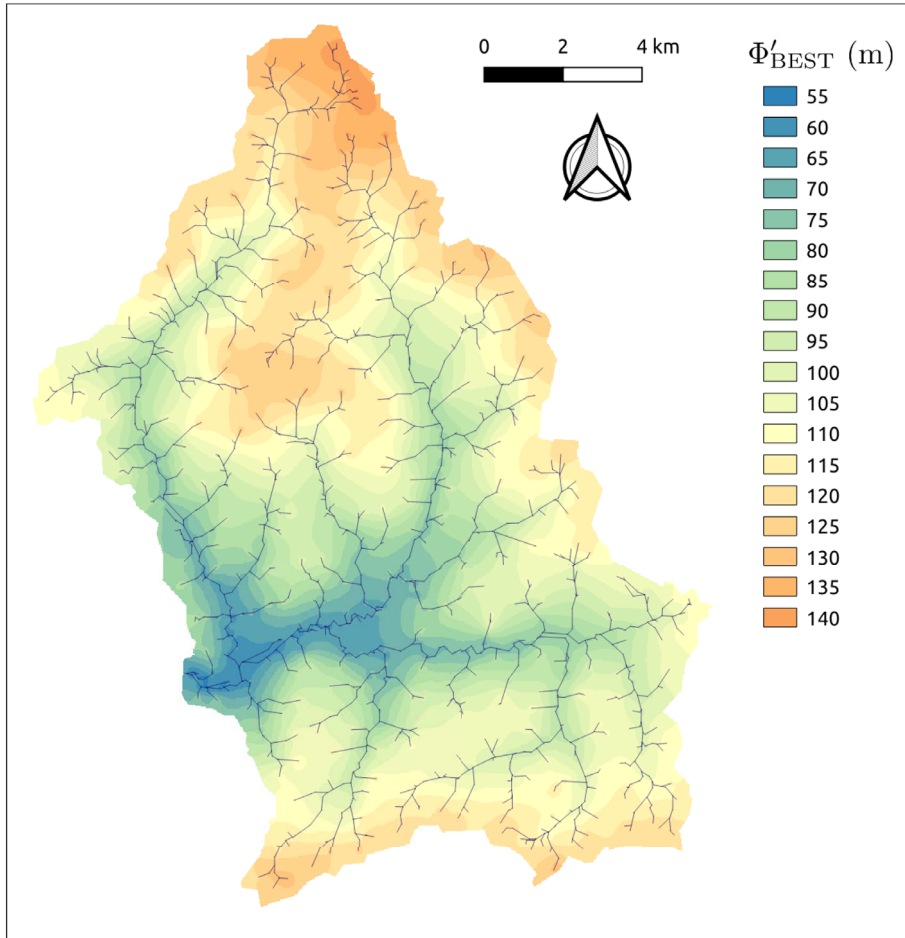
#### 4. Two applications of the Basal Envelope Surface of Talwegs

##### 4.1. Study site: the Evel catchment and Coët-Dan subcatchment

We illustrate the concepts of the previous section on a small catchment located in Brittany (Western France), the Coët-Dan catchment, where sur-

veys have been conducted since the early 1970s [Fovet et al., 2018]. This catchment is developed on brioverian shale with low hydraulic conductivity. In order to avoid edge effect, we consider the larger catchment of the Upper Evel River in which the Coët-Dan catchment is completely nested (gauging site J5618322—Evel at Remungol).





**Figure 7.** Elevation map of the Basal Envelope Surface of Talwegs (BEST) resulting from the least squares regression against talweg elevations, with a Neumann condition on the boundary.

#### 4.1.1. Digital elevation model and derived vector datasets

Figure 6 shows the map of topographic elevation. The original data is the 1-m resolution RGE ALTI [Institut Géographique National, 2018]; in order to speed up the extraction of the hydrographic network without altering channel elevations, we downscale this dataset to a resolution of 5 m by keeping the minimum elevation on each  $5 \times 5$  pixels block. The resulting DEM is then hydrologically conditioned, and the flow direction and flow accumulation maps are computed. The stream network is extracted in vector form from the DEM with a support area of  $0.1 \text{ km}^2$  [ESRI, 2022]. The boundary of the Evel catchment is

also extracted in vector form, as the segments making up this boundary will be assigned a zero normal velocity condition in the least squares regression. In order to reduce the geometrical complexity, the stream links and the boundary polyline are simplified using the Douglas–Peucker algorithm [Douglas and Peucker, 1973] with a planar tolerance of 40 m.

#### 4.1.2. Estimation of the Basal Envelope Surface of Talwegs

The total number of segments in the geometric setting is given in Table 1, along with the associated degrees of freedom.

We use  $n_{\text{coef}} = 2$  complex coefficients for the Laurent series on each slit element, that is, up to 4 real

degrees of freedom per slit. However, we recall that in order for the real potential  $\Phi$  to be continuous between the two “sides” of a slit, the imaginary parts of these coefficients must be zero: hence there are only 2 real degrees of freedom for stream slit elements. This condition of continuity is not required for boundary slit elements: here we only want the normal velocity component to be zero on the “inner” side of each slit. We completely disregard the value of the real potential outside of the boundary, so we use both a real part  $\text{Re}\{c_{j,k}\}$  and an imaginary  $\text{Im}\{c_{j,k}\}$  part for the 2 complex coefficients of each boundary slit element. We set the background velocity to zero ( $v_{0,x} = v_{0,y} = 0$ ), but the constant  $\Phi_0$  is needed in any case and yields an additional degree of freedom, rising this number to 3829.

In order to get a least squares solution for all degrees of freedom, we use the overspecification principle of Janković and Barnes [1999] with  $M = 8$  control points on each slit, whose positions are given by:

$$\zeta_m = \cos\left(\pi \frac{m - \frac{1}{2}}{M}\right) \quad (m = 1, 2, \dots, M).$$

The total number of control points is then  $M \times n_{\text{slit}} = 8 \times (1562 + 176) = 13,904$ , much larger than the number of degrees of freedom (3829). This resulting system is  $13,904 \times 3829$ , which is rather large, but still very easily solved by classical linear algebra packages. If we were to solve an even larger system, we could use the iterative solution introduced by Janković and Barnes [1999], sequentially solving for the coefficients of each slit to match its boundary conditions while holding all other coefficients constant for other slits, until convergence is reached [see also, e.g., Steward, 2020].

Figure 7 shows the resulting envelope surface, that is, the real part  $\Phi'_{\text{BEST}}$  of the normalized complex potential  $\Omega'$ . Elevation ranges from about 50 m a.s.l. at the outlet to about 140 m in the North of the catchment.

#### 4.2. HABEST as a continuous substitute of HAND for flood mapping

Having built a continuous energy level joining all talweg lines, we can now compute the difference between the Basal Envelope Surface of Talwegs (BEST) and the topographic surface. This difference can be coined *Height Above the Basal Envelope Surface of*

*Talwegs* (HABEST). It is conceptually the same thing as the *Height Above Nearest Drainage'* (HAND), with two advantages:

- It is continuous in space. The HAND computation procedure requires a downstream drainage allocation of each pixel, so it creates a jump at the crossing of a water divide when two pixels on either side of the divide are “snapped” to two distinct drainages;
- if a reference hydrographic network is available in vector form, there is no need for hydrological preconditioning of the DEM (cf. Section 4.1.1).

Figure 8 shows the resulting index  $\Phi'_{\text{HABEST}} = (\Phi'_{\text{topo}} - \Phi'_{\text{BEST}}) = (z_{\text{topo}} - \Phi'_{\text{BEST}})$ . Contours of this surface can be used for flood mapping in the same way as HAND contours.

#### 4.3. Groundwater level interpolation

##### 4.3.1. Regression using HABEST as predictor

In this section, we use the resulting HABEST index as a predictor for groundwater level, following the idea of Wyns et al. [2004]. Figure 9 is a zoom on the map of Figure 7 on the Coët-Dan subcatchment; it also shows the location of 74 boreholes in which the groundwater level was surveyed during the month of February, 1996 [Pauwels et al., 1996]. The model reads:

$$\Phi'_{\text{GW}} - \Phi'_{\text{BEST}} = a(\Phi'_{\text{topo}} - \Phi'_{\text{BEST}}) + b \quad (13)$$

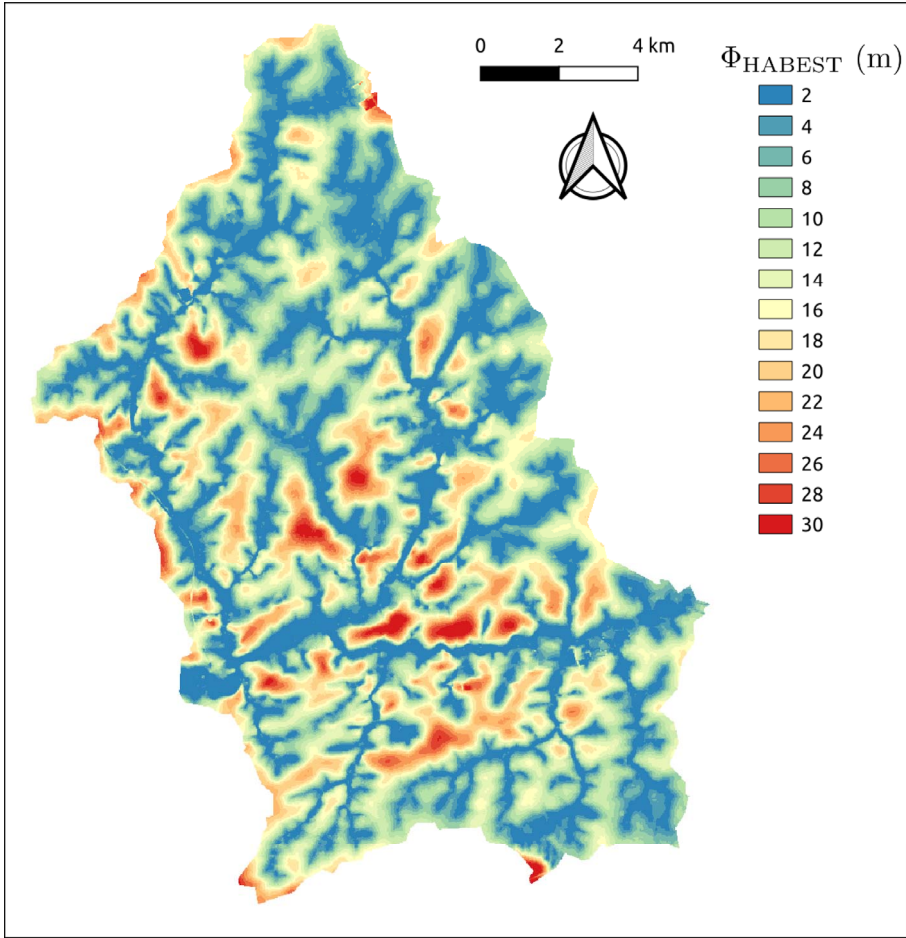
with  $\Phi'_{\text{GW}} = h$  the groundwater head (or real potential normalized by  $T_0$ ), and  $\Phi'_{\text{topo}} = z_{\text{topo}}$  the topographic elevation.

The result of the regression is shown in Figure 10. The parameters are:

$$\begin{cases} a = 0.912 \\ b = -1.74 \text{ m} \end{cases}$$

with  $R^2 = 0.962$  and a rooted mean squared error (RMSE) of 1.67 m.

It is worth noting that the estimate of groundwater level using (13) is not completely insensitive to the value of the support area chosen for extracting the channel network. The RMSE obtained for different values of the support area is shown in Figure 11; we see that the optimal value is about  $A_c = 0.4 \text{ km}^2$ .



**Figure 8.** Map of the Height Above the Basal Envelope Surface of Talwegs (HABEST) index, defined as the difference between topographic elevation  $z_{\text{topo}} = \Phi'_{\text{topo}}$  and the elevation  $\Phi'_{\text{BEST}}$ ; this index is closely related to the HAND index.

#### 4.3.2. A non-zero divergence solution with area-sink

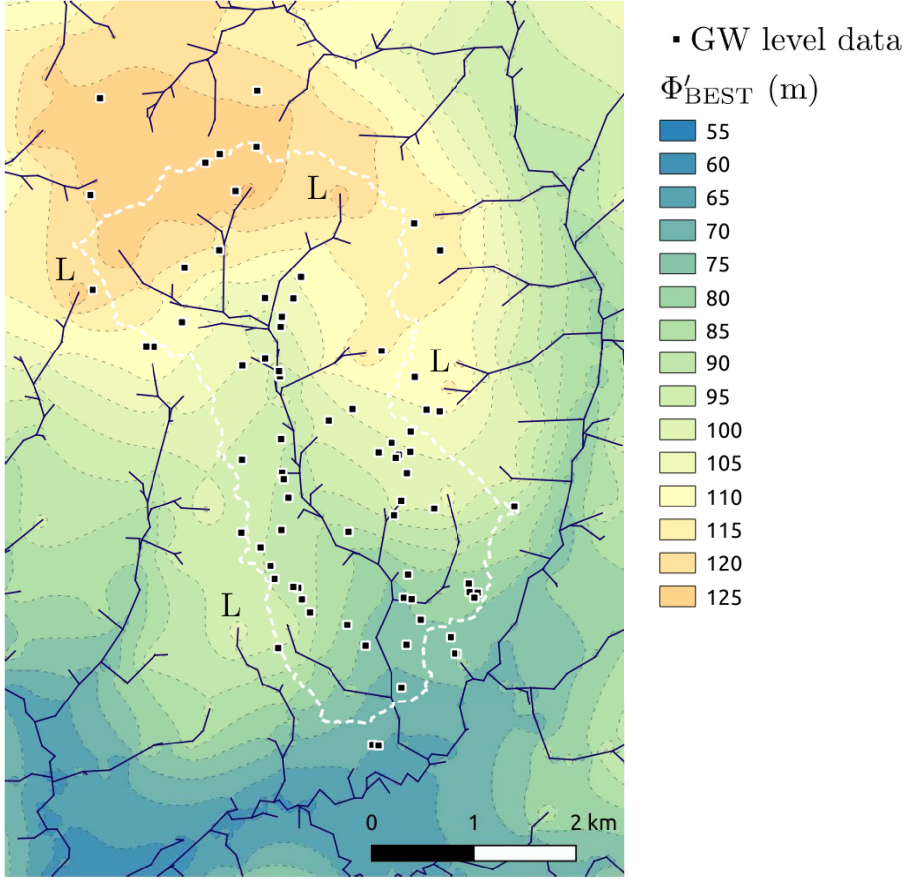
The definition of the envelope surface as a harmonic potential (corresponding to a divergence-free velocity field) is straightforward, and the result shown in Figure 7 exactly corresponds to the specifications so far. However, if we look closer at Figure 9, we see that slits corresponding to “leaves” in the network graph often correspond to a maximum of the envelope surface, just as if the surface was locally “hanging” from those elements, with the flux oriented from the slit to the domain (see points denoted with a “L”). It is no big issue if we do not want to give any particular physical meaning to this envelope surface; however, this is a spurious behaviour if we want to interpret it, for instance, as a (dry-season)

groundwater level.

We can achieve a slightly more satisfactory interpolating surface by computing a non-zero divergence velocity field, still using AEM. For this we will set a non-zero sink term  $Q_j$  for each stream slit, in conjunction with an area sink [Strack, 2017] with uniform infiltration rate. The Laplacian of the potential now satisfies the following steady-state equation away from the slits:

$$T_0 \Delta h = T_0 \Delta \Phi' = -r_0$$

where  $r_0 > 0$  is a recharge depth per unit time (positively-defined, source term). As the transmissivity is unknown, it is again more convenient to rewrite



**Figure 9.** Close-up on the elevation map of the Basal Envelope Surface of Talwegs in the Coët Dan subcatchment (dashed white line). Black squares indicate the location of the 74 boreholes in which the groundwater level was surveyed during the month of February, 1996.

this equation as:

$$\Delta\Phi' = -\frac{r_0}{T_0} = +\gamma_0$$

where the sign of  $\gamma_0$  is set according to the convention used in Strack [2017]. Just as we normalized  $Q'_j = Q_j/T_0$  for a slit sink term,  $\gamma_0 = -r_0/T_0$  is a normalized areal extraction rate term which has the dimension of the inverse of a depth (e.g.,  $\text{m}^{-1}$ ).

Since we add a source term in the system, we need to add one or several sink term(s) somewhere else: obviously, it has to be at stream slits. The idea is to make the sink term  $Q_j$  for each slit element  $j$  a function of  $\gamma_0$ . Let consider the portion of the stream slit graph shown in Figure 12: if we assume that we have only very local flow systems, we can consider that each slit cannot extract more than the amount of recharge on the subcatchment, it drains. Denoting

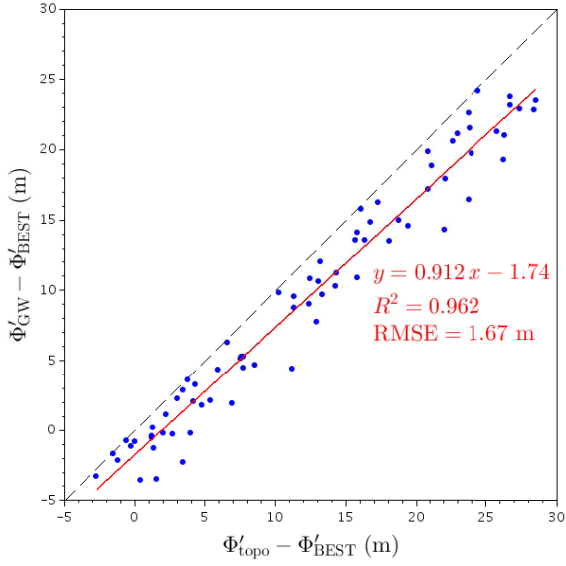
$dA_j$  the area of this subcatchment, which is a purely geometric quantity that can be easily computed in a Geographic Information System (GIS), we will even consider that the two quantities exactly match:

$$Q_j = r_0 dA_j$$

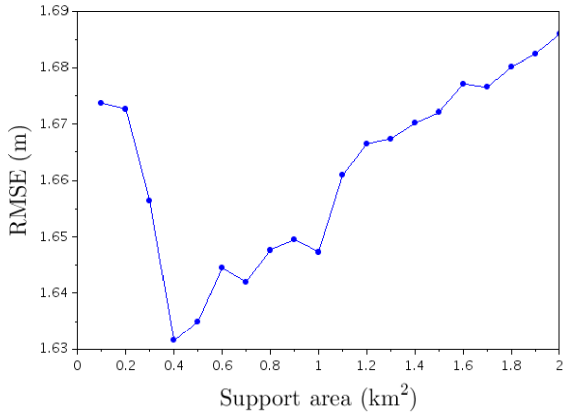
or, equivalently,

$$Q'_j = \frac{r_0}{T_0} dA_j = -\gamma_0 dA_j.$$

It is worth reiterating the fact that even though  $Q_j$  represents a flux extracted from the groundwater system, it does not necessarily represent a contribution to streamflow. A fraction of this flux might as well represent riparian or wetland evapotranspiration concentrated in the vicinity of the talweg, especially for the smallest streams that often dry up during summertime.



**Figure 10.** Result of the regression of relative groundwater above the Basal Envelope Surface of Talwegs,  $(\Phi'_{\text{GW}} - \Phi'_{\text{BEST}})$ , against the Height Above the Basal Envelope Surface of Talwegs  $(\Phi'_{\text{topo}} - \Phi'_{\text{BEST}})$ .



**Figure 11.** Rooted-Mean-Squared-Error (RMSE) of the regression of  $(\Phi'_{\text{GW}} - \Phi'_{\text{BEST}})$  against  $(\Phi'_{\text{topo}} - \Phi'_{\text{BEST}})$ , as a function of the support area used for extracting the talweg network from the DEM. We see a slight dependence, with an optimal value of the support area about  $A_c = 0.4 \text{ km}^2$ .

For the sake of brevity, the reader is referred to Strack [2017] for the expression of  $\Phi_{\text{areal}}(z)$  and the associated velocity  $\nu_{\text{areal}}(z)$ . Both are linear with respect to the new degree of freedom  $\gamma_0$ . Of course,

there is no stream function  $\Psi_{\text{areal}}$  associated with  $\Phi_{\text{areal}}$ : because of the non-zero divergence there is no holomorphic complex potential. Figure 13 shows the contours of  $\Phi_{\text{areal}}$  for a unit extraction rate  $\gamma_0 = 1$ .

By superposition, the real potential produced by all elements in the domain finally reads:

$$\begin{aligned} \Phi'_{\text{GW}}(z) = & \Phi'_0 + \text{Re}\{\bar{\nu}'_0 z\} + \gamma_0 \Phi'_{\text{areal}}(z; \gamma_0 = 1) \\ & + \sum_{j=1}^{n_{\text{sink}}} \left[ -\gamma_0 dA_j \text{Re} \left\{ \frac{\ln Z_j(z)}{2\pi} - 1 \right\} + \sum_{k=1}^{n_{\text{coef}}} c_{j,k} \text{Re}\{Z_j^{-k}(z)\} \right] \\ & \text{(stream slit elements)} \\ & + \sum_{j'=1}^{n_{\text{bound}}} \sum_{k=1}^{n_{\text{coef}}} \text{Re}\{c_{j',k} Z_{j'}^{-k}(z)\} \\ & \text{(boundary slit elements)}. \end{aligned}$$

Just as before, the coefficients of the Laurent series  $c_{j,k}$  for stream slits are real-valued in order to ensure the continuity of  $\Phi'$ , while the coefficients  $c_{j',k}$  for boundary slits can remain complex-valued as long as we are not interested in computing a potential outside of the boundary.

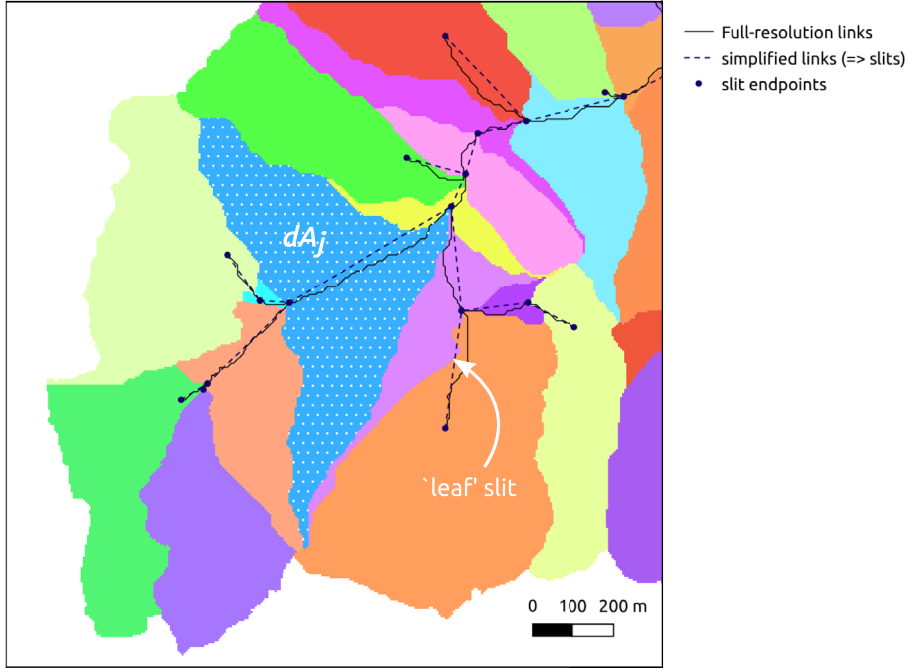
The degrees of freedom are better identified with ancillary potential data at groundwater observation boreholes. Figure 14 shows the final map; the regression yields  $\gamma_0 = -1.19 \times 10^{-4} \text{ m}^{-1}$ .

In Figure 15, we compare the observed versus estimated groundwater level at the 74 observation boreholes. In order to narrow down the range of values we plot the estimated depth,  $\Phi'_{\text{topo}} - \hat{\Phi}'_{\text{GW}}$  as a function of observed depth  $\Phi'_{\text{topo}} - \Phi'_{\text{GW}}$ . The RMSE is substantially improved compared to the previous method (0.70 m versus 1.63 m).

Interestingly, if we have an independent estimate of recharge rate  $r_0$ , we can get an estimate of large-scale, single-layer equivalent transmissivity. Various estimates of annual effective rainfall can be found in the literature for the region, ranging from 180 mm [Mougin et al., 2004] to 393 mm [BRGM, 2019]. Taking a mean estimate and considering a 50%–50% partition of this amount between surface runoff and recharge, this leads to a rough estimate of  $r_0 \approx 4.5 \times 10^{-9} \text{ m}\cdot\text{s}^{-1}$ . We can then estimate the equivalent, uniform transmissivity  $\hat{T}_0$  with:

$$\hat{T}_0 = \frac{r_0}{|\gamma_0|} \sim 3.8 \times 10^{-5} \text{ m}^2\cdot\text{s}^{-1}.$$

We can compare this conceptual estimate with values given by aquifer tests: Martelat and Lachassagne [1995] give the ranges  $0.8\text{--}1.5 \times 10^{-3} \text{ m}^2\cdot\text{s}^{-1}$  for the superficial schist alterite, and  $1.5\text{--}3.0 \times 10^{-5} \text{ m}^2\cdot\text{s}^{-1}$  for the deeper, fissured



**Figure 12.** For each stream slit  $j$ , we compute  $dA_j$  the area of the subcatchment drained by the slit, that is, the difference between the value of the flow accumulation map at the downstream endpoint of the slit minus the value at the upstream endpoint. For “leaf” slits,  $dA_j$  is simply set to the drained areal at the downstream end of the slit. Each slit is then allowed to extract a normalized discharge  $Q'_j = (r_0/T_0) dA_j = -\gamma_0 dA_j$ .

schist. We see that even if it remains conceptual, the value obtained by regression using the AEM method could be an interesting proxy for characterizing the subsurface at large scale.

## 5. Conclusion and perspectives

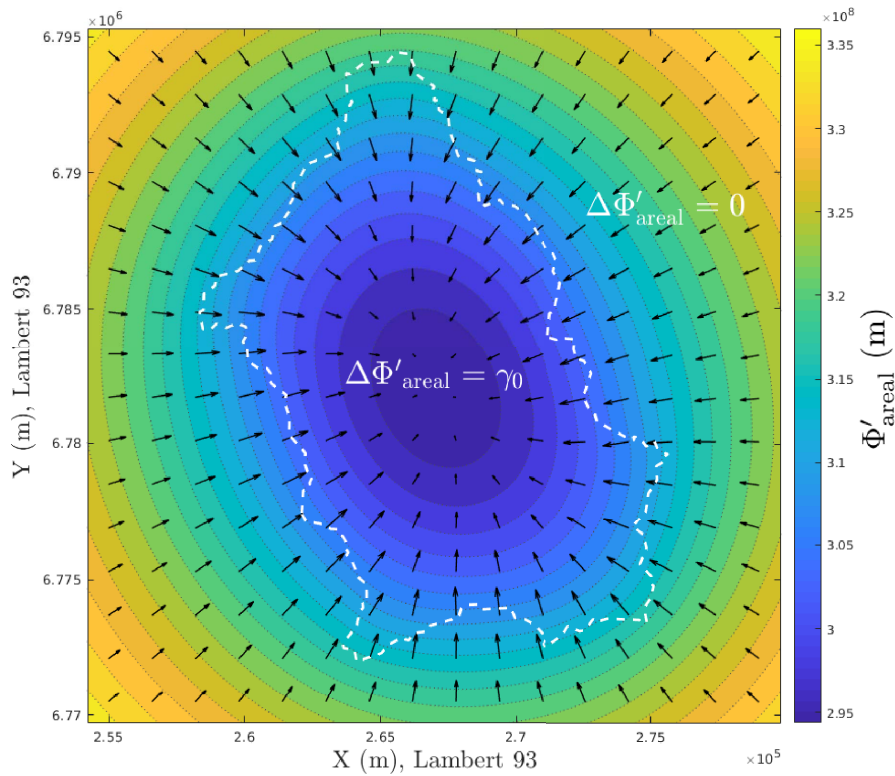
In this concluding section we sum up our findings by making a parallel between the decomposition of the topographic surface and that of the groundwater level. Figure 16 illustrates the following decomposition:

$$\begin{cases} z_{\text{topo}} = \Phi'_{\text{topo}} = \Phi'_{\text{BEST}} + \underbrace{(\Phi'_{\text{topo}} - \Phi'_{\text{BEST}})}_{\text{HABEST}} \\ h = \Phi'_{\text{GW}} = \Phi'_{\text{BEST}} + \underbrace{(\Phi'_{\text{GW}} - \Phi'_{\text{BEST}})}_{\text{GW level relative to BEST}} \end{cases}$$

As can be seen on the figure, the topographic height above the envelope surface of talwegs ( $\Phi'_{\text{topo}} - \Phi'_{\text{BEST}} = \text{HABEST}$ ) and the relative groundwater level

above this envelope surface ( $\Phi'_{\text{GW}} - \Phi'_{\text{BEST}}$ ) have a very similar behaviour. It comes as no surprise for geomorphologists as, in the hillslope domain, both sediment transfer at the surface and groundwater transfer in the porous medium are ruled by diffusive processes [Anderson and Anderson, 2010]. It confirms that the Basal Envelope Surface of Talwegs is an interesting energy level for hydrological and geomorphological analysis.

From the point of view of catchment hydrological modeling, the kind of parsimonious and analytical reconstruction of the groundwater surface proposed in this study, as well as the identification of a conceptual, large-scale estimate of transmissivity  $\hat{T}_0$ , could provide useful ancillary information for constraining the calibration of lumped models, with a rationale similar to that of, e.g., TOPMODEL [Beven and Kirkby, 1979]. Only a handful of observation boreholes are needed: if the observations are available during both dry and wet seasons, the approach



**Figure 13.** Contour map of real potential created by an areal sink of unit normalized extraction rate  $\gamma_0 = -r_0/T_0 = 1$ . The Laplacian of the potential uniformly equals  $\gamma_0$  inside the polygon (white dashed line), and it is harmonic outside ( $\Delta\Phi'_{\text{areal}} = 0$ ). Note the very large magnitude of the values (in the order of  $10^8$ ): of course, it has to be scaled by a much smaller factor  $|\gamma_0| \ll 1$  in order to model topographic or groundwater elevation.

could also yield a large-scale estimate of storage coefficient,  $\hat{S}$ . The two parameters  $\hat{T}_0$  and  $\hat{S}$  can yield first-order estimates of lumped groundwater storage and fluxes which can easily be used as state variables in the model.

Codes used in this study are available in the author's Basilisk sandbox at <http://basilisk.fr/sandbox/nlemoine/AEM/enveloppe/enveloppe.c>.

### Conflicts of interest

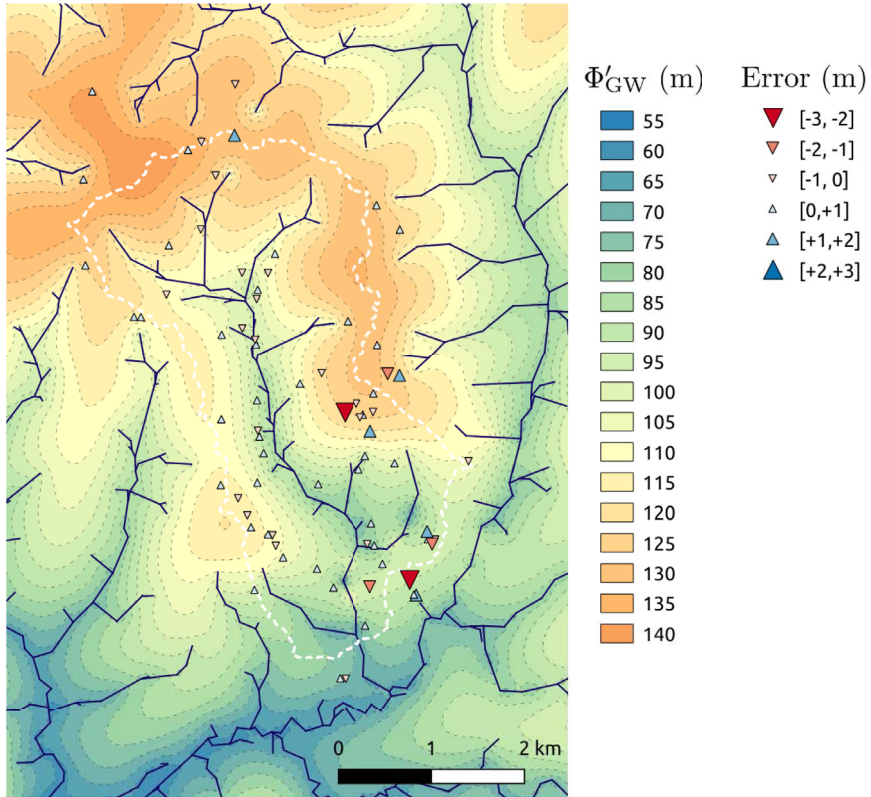
The author has no conflict of interest to declare.

### Acknowledgments

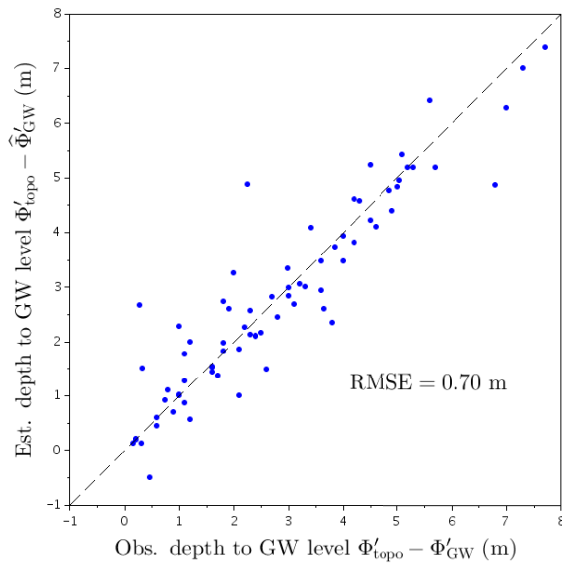
I would like to take the opportunity of this special issue to thank Ghislain de Marsily for all his incredible teaching, especially for passing on to students a taste for analytic approaches.

### Appendix A. List of symbols

See Table A1.



**Figure 14.** Simulated map of the groundwater potential  $\Phi'_{GW}$  in the Coët-Dan subcatchment, with the additional degree of freedom  $\gamma_0$  (normalized areal exfiltration rate). The regression yields  $\gamma_0 = -1.19 \times 10^{-4} \text{ m}^{-1}$  (net infiltration). The error at the 74 observation borehole is represented with graduated triangles.



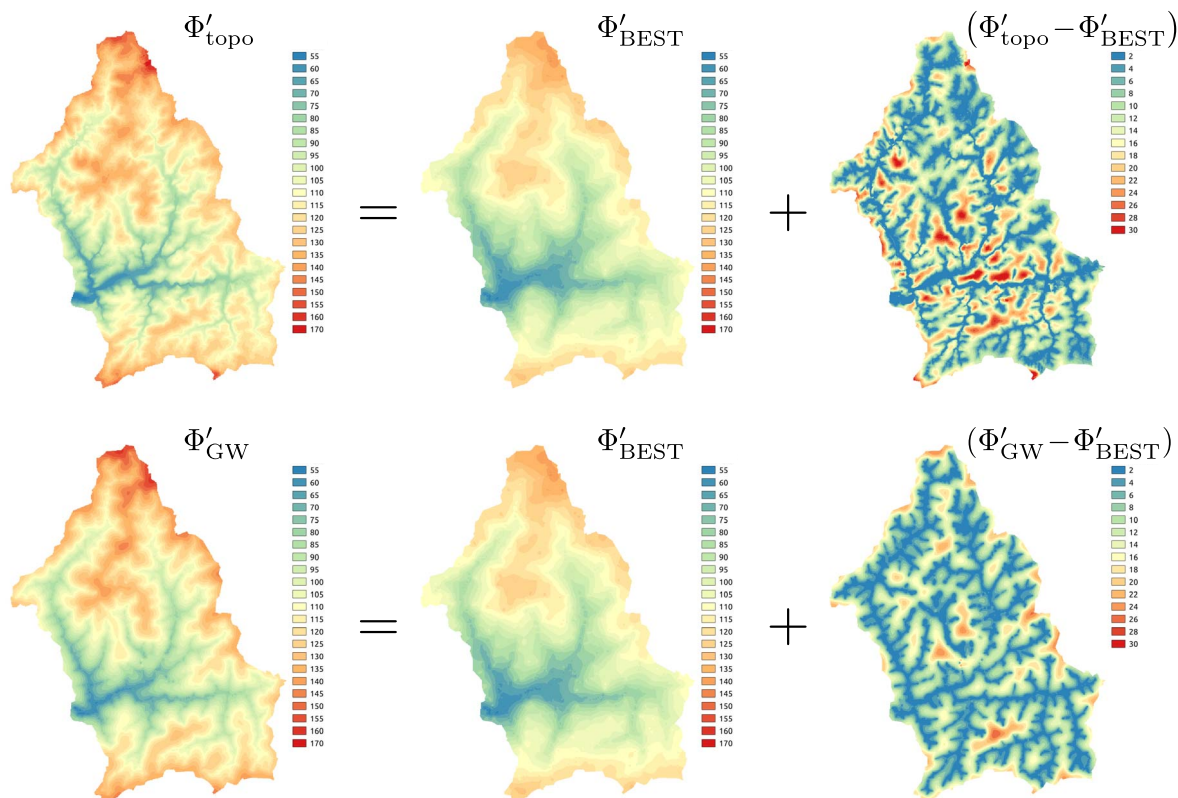
**Figure 15.** Scatter plot of observed versus estimated depth to groundwater level with the additional degree of freedom  $\gamma_0$  (normalized areal exfiltration rate).



**Table A1.** List of symbols used throughout the paper

Symbol	Definition	Values in	Unit
$h$	Groundwater level w.r.t. reference elevation (e.g., sea level)	$\mathbb{R}$	m
$\vec{\nabla} h$	2D gradient vector of $h$ , $\vec{\nabla} h = \begin{bmatrix} \partial h / \partial x \\ \partial h / \partial y \end{bmatrix}$	$\mathbb{R} \times \mathbb{R}$	Norm $\ \vec{\nabla} h\ $ is a slope in $\text{m} \cdot \text{m}^{-1}$ ( $\emptyset$ )
$z_{\text{topo}}$	Topographic elevation w.r.t. reference elevation	$\mathbb{R}$	m
$K$	Hydraulic conductivity	$\mathbb{R}$	$\text{m} \cdot \text{s}^{-1}$
$T(h)$	Transmissivity	$\mathbb{R}$	$\text{m}^2 \cdot \text{s}^{-1}$
$T_0$	Uniform value of transmissivity	$\mathbb{R}$	$\text{m}^2 \cdot \text{s}^{-1}$
$S$	Storage coefficient	$\mathbb{R}$	$\emptyset$
$r(x, y, t)$	Recharge rate	$\mathbb{R}$	$\text{m} \cdot \text{s}^{-1}$
$\vec{q}$	Vertically-integrated Darcy flux density, $\vec{q} = \begin{bmatrix} q_x \\ q_y \end{bmatrix} = -T(h) \vec{\nabla} h$	$\mathbb{R} \times \mathbb{R}$	Norm $\ \vec{q}\ $ in $\text{m}^2 \cdot \text{s}^{-1}$
$q$	Same vector but represented in a complex plane, $q = q_x + i q_y$	$\mathbb{C}$	Modulus $ q  = \ \vec{q}\ $ in $\text{m}^2 \cdot \text{s}^{-1}$
$\Phi$	Real-valued potential such that $\vec{q} = -\vec{\nabla} \Phi$	$\mathbb{R}$	$\text{m}^3 \cdot \text{s}^{-1}$
$i$	Complex number such that $i^2 = -1$	$\mathbb{C}$	—
$z = x + i y$	Affix of a point in the complex plane $\mathbb{C}$ associated with a given set $(x, y)$ of projected coordinates (e.g., Lambert 93)	$\mathbb{C}$	$ z $ in m
$\Omega$	Complex-valued potential such that $\Phi = \text{Re}\{\Omega\}$	$\mathbb{C}$	Modulus $ \Omega $ in $\text{m}^3 \cdot \text{s}^{-1}$
$\Psi$	Real-valued stream function, $\Psi = \text{Im}\{\Omega\}$	$\mathbb{R}$	$\text{m}^3 \cdot \text{s}^{-1}$
$a_j$	Affix of the start point of slit element $j$ : $a_j = x_{A_j} + i y_{A_j}$	$\mathbb{C}$	$ a_j $ in m
$b_j$	Affix of the end point of slit element $j$ : $b_j = x_{B_j} + i y_{B_j}$	$\mathbb{C}$	$ b_j $ in m
$\Omega_j(z)$	Contribution of slit element $j$ to total potential $\Omega_{\text{tot}}(z)$ at location $z$ in the complex plane	$\mathbb{C}$	$ \Omega_j(z) $ in $\text{m}^3 \cdot \text{s}^{-1}$
$Z_j(z)$	Local transformed coordinate (complex affix) w.r.t. to slit $j$	$\mathbb{C}$	Modulus $ Z_j(z) $ is $\emptyset$
$Q_j$	Sink term for slit element $j$	$\mathbb{R}$	$\text{m}^3 \cdot \text{s}^{-1}$
$c_{j,k}$	Coefficient of power term $Z_j^{-k}(z)$ in $\Omega_j(z)$	$\mathbb{C}$	$ c_{j,k} $ in $\text{m}^3 \cdot \text{s}^{-1}$ (same as $ \Omega $ )
$\Omega' = \Omega / T_0$	Normalized complex potential	$\mathbb{C}$	$ \Omega' $ in m (head)
$\Phi' = \Phi / T_0$	Normalized real potential. $\Phi' = \text{Re}\{\Omega'\}$	$\mathbb{R}$	m (head)
$q' = q / T_0$	Normalized flux density	$\mathbb{C}$	Modulus $ q' $ is $\emptyset$
$Q'_j = Q_j / T_0$	Normalized sink term for slit element $j$	$\mathbb{R}$	m (head loss)
$c'_{j,k}$	Normalized coefficient, for term $Z_j^{-k}(z)$ in $\Omega'_j(z)$	$\mathbb{C}$	$ c'_{j,k} $ in m (same as $ \Omega' $ )
$r_0$	Uniform recharge rate	$\mathbb{R}$	$\text{m} \cdot \text{s}^{-1}$
$\gamma_0 = -r_0 / T_0$	Normalized uniform recharge rate	$\mathbb{R}$	$\text{m}^{-1}$
$dA_j$	Area of topographic subcatchment drained by slit element $j$	$\mathbb{R}$	$\text{m}^2$
$\Phi'_{\text{topo}}$	Normalized real potential corresponding to topographic surface. $\Phi'_{\text{topo}} = z_{\text{topo}}$	$\mathbb{R}$	m (head)
$\Phi'_{\text{GW}}$	Normalized real potential corresponding to groundwater level. $\Phi'_{\text{GW}} = h$	$\mathbb{R}$	m (head)
$\Phi'_{\text{BEST}}$	Normalized real potential corresponding to the elevation of the envelope surface of talwegs	$\mathbb{R}$	m (head)
$\hat{\Phi}'_{\text{GW}}$	Estimated groundwater level	$\mathbb{R}$	m (head)

Symbols are (approximately) listed by order of appearance in the text. The type of each variable is indicated: real number (in  $\mathbb{R}$ ), 2D vector (in  $\mathbb{R} \times \mathbb{R}$ ) or complex number (in  $\mathbb{C}$ ).



**Figure 16.** Decomposition of topographic elevation ( $\Phi'_{\text{topo}}$ , top row) and groundwater level elevation ( $\Phi'_{\text{GW}}$ , bottom row) using the ancillary level  $\Phi'_{\text{BEST}}$ . We see that in both cases this level plays the role of a base level for diffusive processes in the hillslope domain (diffusive sediment transport on the topographic surface, and diffusive Darcy flow for groundwater).

## References

- Anderson, R. and Anderson, S. (2010). *Geomorphology: The Mechanics and Chemistry of Landscapes*. Cambridge University Press, Cambridge.
- Beven, K. J. and Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrol. Sci. Bull.*, 24(1), 43–69.
- BRGM (2019). 195AC01 – Socle métamorphique dans le bassin versant de l’Evel de sa source au Blavet (nc). [https://sigebre.brgm.fr/files/fiches/BDLISA/LISA\\_Bretagne\\_195AC01.pdf](https://sigebre.brgm.fr/files/fiches/BDLISA/LISA_Bretagne_195AC01.pdf). BD LISA database, accessed: 2022-03-30.
- Cohen, Y., Devauchelle, O., Seybold, H. F., Yi, R. S., Szymczak, P., and Rothman, D. H. (2015). Path selection in the growth of rivers. *Proc. Natl. Acad. Sci. USA*, 112(46), 14132–14137. <https://www.pnas.org/content/112/46/14132>.
- De Marsily, G. (1986). *Quantitative Hydrogeology: Groundwater Hydrology for Engineers*. Elsevier Science, Amsterdam.
- Deffontaines, B. (1990). *Développement d’une méthodologie d’analyse morphostructurale et morphonéotectonique. Analyse des surfaces enveloppes du réseau hydrographique et des modèles numériques de terrain; application au nord-est de la France*. PhD thesis, Université Pierre et Marie Curie.
- Douglas, D. H. and Peucker, T. K. (1973). Algorithms for the reduction of number of points required to represent a digitized line or its caricature. *Cartogr. Int. J. Geogr. Inform. Geovisual.*, 10, 112–122.
- Dunne, T. (1969). *Runoff production in a humid area*. PhD thesis, Johns Hopkins University, Baltimore, MD.
- Dunne, T. (1980). Formation and controls of channel networks. *Prog. Phys. Geogr.*, 4, 211–239.
- Dunne, T. and Black, R. (1970). Partial area contribu-

- tion to storm runoff in a small New England watershed. *Water Resour. Res.*, 6, 1296–1311.
- ESRI (2022). ArcGIS Pro geoprocessing tool reference. <https://pro.arcgis.com/en/pro-app/2.8/tool-reference/spatial-analyst/stream-link.htm>. Accessed: 2022-02-21.
- Fovet, O., Ruiz, L., Gruau, G., Akkal, N., Aquilina, L., Busnot, S., Dupas, R., Durand, P., Faucheux, M., Fauvel, Y., Fléchar, C., Gilliet, N., Grimaldi, C., Hamon, Y., Jaffrezic, A., Jeanneau, L., Labasque, T., Le Henaff, G., Mérot, P., Molénat, J., Petitjean, P., Pierson-Wickmann, A.-C., Squidant, H., Viaud, V., Walter, C., and Gascuel-Oudou, C. (2018). Agrhys: An observatory of response times in agrohydro systems. *Vadose Zone J.*, 17(1), article no. 180066.
- Häsing, J. (1964). Bestimmung der Glättungstiefe rauher Flächen. *PTB-Mitt.*, 4, 339–340.
- Institut Géographique National (2018). RGE ALTI<sup>®</sup> version 2.0 – Descriptif de contenu. [https://geoservices.ign.fr/sites/default/files/2021-07/DC\\_RGEALTI\\_2-0.pdf](https://geoservices.ign.fr/sites/default/files/2021-07/DC_RGEALTI_2-0.pdf).
- Janković, I. and Barnes, R. (1999). High-order line elements in modeling two-dimensional groundwater flow. *J. Hydrol.*, 226(3), 211–223. <https://www.sciencedirect.com/science/article/pii/S0022169499001407>.
- Joukowsky, N. (1910). Über die Konturen der Tragflächen der Drachenflieger. *Z. Flugtechnik Motorluftschiffahrt*, 1(22), 281–285.
- Martelat, A. and Lachassagne, P. (1995). Bassin Versant Représentatif Expérimental du Coët Dan (Naizin, Morbihan). Hydrogéologie : détermination des caractéristiques hydrodynamiques du système aquifère au lieu-dit Le Stimoës. Technical Report BRGM/RR-38474-FR.
- Mougin, B., Thomas, E., Wyns, R., Blanchin, R., and Mathieu, F. (2004). Qualité des eaux en Bretagne – Ruissellement – Infiltration – Temps de réponse – Bassin versants du Yar (22), de l’Horn (29), et du Coët Dan (56) – Rapport final. Technical Report BRGM/RP-52731-FR.
- Nobre, A., Cuartas, L., Hodnett, M., Rennó, C., Rodrigo, G., Silveira, A., Waterloo, M., and Saleska, S. (2011). Height above the nearest drainage – a hydrologically relevant new terrain model. *J. Hydrol.*, 404(1), 13–29. <https://www.sciencedirect.com/science/article/pii/S0022169411002599>.
- Nobre, A. D., Cuartas, L. A., Momo, M. R., Severo, D. L., Pinheiro, A., and Nobre, C. A. (2016). HAND contour: a new proxy predictor of inundation extent. *Hydrol. Process.*, 30(2), 320–333.
- Pauwels, H., Martelat, A., Foucher, J. C., and Lachassagne, P. (1996). Dénitrification dans les eaux souterraines du bassin du Coët-Dan : Suivi géochimique et hydrogéologique du processus. Technical Report BRGM/RR-39055-FR.
- Petroff, A. P., Devauchelle, O., Kudrolli, A., and Rothman, D. H. (2012). Four remarks on the growth of channel networks. *C. R. Géosci.*, 344(1), 33–40. <https://www.sciencedirect.com/science/article/pii/S1631071312000028>.
- Steward, D. R. (2015). Analysis of discontinuities across thin inhomogeneities, groundwater/surface water interactions in river networks, and circulation about slender bodies using slit elements in the Analytic Element Method. *Water Resour. Res.*, 51(11), 8684–8703.
- Steward, D. R. (2020). *Analytic Element Method: Complex Interactions of Boundaries and Interfaces*. Oxford University Press, Oxford.
- Strack, O. D. (1989). *Groundwater Mechanics*. Prentice Hall, Hoboken, NJ.
- Strack, O. D. (2017). *Analytical Groundwater Mechanics*. Cambridge University Press, Cambridge.
- Système d’Information sur l’Eau/Sandre (2020). Document de présentation – Description du référentiel hydrographique (BD TOPAGE<sup>®</sup>) version 1. [http://www.sandre.eaufrance.fr/ftp/documents/fr/pre/eth/1/sandre\\_pres\\_eth\\_1.pdf](http://www.sandre.eaufrance.fr/ftp/documents/fr/pre/eth/1/sandre_pres_eth_1.pdf).
- Wyns, R., Baltassat, J.-M., Lachassagne, P., Legchenko, A., Vairon, J., and Mathieu, F. (2004). Application of proton magnetic resonance soundings to groundwater reserve mapping in weathered basement rocks (Brittany, France). *Bull. Soc. Géol. Fr.*, 175(1), 21–34.
- Yao, T. K., Fouché, O., Kouadio, K. E., and Oga, M.-S. (2017). Discontinuous nature of phreatic aquifers in granitic rocks at watershed scale: A stratiform model from perennial streams and well data. *Asian Rev. Environ. Earth Sci.*, 4(1), 20–27.





Research article

Geo-hydrological Data & Models

# Pitfalls of space-time trading when parametrizing a land use dependent hydrological model

Ludovic Oudin<sup>①,\*,a</sup> and Morgane Lalonde<sup>①,a</sup>

<sup>a</sup> Sorbonne Université, CNRS, EPHE, UMR METIS, Case 105, 4 place Jussieu, F-75005 Paris, France

*E-mails:* ludovic.oudin@upmc.fr, Ludovic.oudin@sorbonne-universite.fr (L. Oudin), morgane.lalonde-le\_pajolec@sorbonne-universite.fr (M. Lalonde)

**Abstract.** Evaporation is a function of both climate conditions and other environmental conditions, including land use. In the context of large-scale environmental changes, understanding the relative impacts of each driver of past and future evaporation changes is necessary for land and water planning. While climate change impacts on evaporation can be estimated straightforwardly by original Budyko formulations, including the role of land use within these formulations remains an open question. In this paper, we collected an extensive set of 5026 worldwide catchments to parametrize a land use dependent Budyko-type formulation. By trading space for time, we then assess the potentialities of the proposed formulation in predicting the impacts of land use changes on the evaporation changes. Results show a clear modulation of land use on evapotranspiration, suggesting larger and lower evaporation rates over croplands and urban areas respectively. The proposed formulation was able to reasonably predict the magnitude of the decrease of the evaporative ratio on urbanizing catchments, but fails to efficiently predict the hydrological impacts of vegetated land use conversions, both in terms of direction and magnitude of changes. This suggests either the proposed formulation is too crude, or the underlying hypotheses of space-time trading are not valid.

**Keywords.** Rainfall-runoff modeling, Land use change, Budyko, Hydrology, Evaporation, Urbanization, Afforestation.

*Manuscript received 31 March 2022, revised 21 June 2022, accepted 18 July 2022.*

## 1. Introduction

Most, if not all, of hydrological models are conceptual models requiring validation tests specific to their desired objective [de Marsily, 1994]. When these models are applied under non-stationary conditions e.g. for climate and land use change impact studies, it is necessary to ensure that their structure/parameters are transferable in time [Thirel et al., 2015], by adopting more rigorous cross-validation tests, e.g. the differential split-sample test proposed by Klemesš

[1986]. This strategy is rather limited since models are often developed and used for extrapolation. Consequently, pure cross-validation for specific environmental changes is seldom possible, because of a lack of historical environmental conditions similar to projected environmental conditions. An appealing way to circumvent this problem is trading time for space [Peel and Blöschl, 2011], the rationale being that if a model can deal with different environments in different locations, it will deal with different environments for different periods at a given location. Despite the widespread use among hydrologists of the “trading time for space” approach [Singh et al., 2011], very few experimental studies were pro-

\* Corresponding author.

posed to (in)validate it.

In this paper, we develop a modeling framework designed to test the “trading time for space” method. We start from the simplest expression of rainfall-runoff transformation at the pluriannual time scale using Budyko-type formulations that relate the ratio of long-term average evaporation ratio ( $E/P$ ) to the ratio of long-term average potential evaporation  $PET$  to long-term average rainfall  $P$  [Budyko, 1974, Ol’Dekop, 1911, Schreiber, 1904, Turc, 1954]. These formulations had perceived a renewed interest over the last two decades since they may provide simple assessment tools to the attribution problem, i.e. they may be employed to separate the effects of climate to other environmental changes on long-term average evapotranspiration and consequently long-term average runoff [Jaramillo et al., 2018, Roderick and Farquhar, 2011, Wang, 2014]. Adaptations of the classical models flourished intending to include more and more drivers in addition to long-term climate settings [Donohue et al., 2007, Li et al., 2013, Zhang et al., 2004]. Among these attempts, including land use and vegetation characteristics into the model formulation is the most popular since vegetation likely exerts a significant influence on regional water balance and feedback to the atmosphere.

A common feature of the frameworks aiming at developing vegetation-dependent water balance models is to collect hydroclimatic data from climatologically diverse “steady-state” catchments, in the idea to use then these parametrizations for land use changes studies, i.e. trading time for space. The diversity of the catchment set used to derive the parametrized Budyko-type formulations is often put forward as a necessary condition, as advocated by Large-Sample Hydrology groups [Gupta et al., 2014].

Two objectives form the main sections of this article: (i) develop a land use dependent Budyko formulation based on an extensive set of 5026 worldwide catchments to encompass a large variety of climatic and land use conditions, (ii) assess the potentialities of the derived land use dependent Budyko formulation to detect and quantify hydrological impacts of land uses modifications at the catchment scale. As stated by Singh et al. [2011] on climate, the hypothesis is that the spatial relationship between land use and streamflow characteristics is similar to the one observed between land use and streamflow over long periods at a single location.

## 2. Material

### 2.1. Data

We selected a large sample of catchments, from several international and national databases: the Global Streamflow Indices and Metadata Archive (GSIM) [Do et al., 2018] consisting of a collection of streamflow time series for more than 35,000 catchments worldwide, the 9322 GAGE-II US stream gauges maintained by the U.S. Geological Survey (USGS) [Falcone, 2011], the HydroPortail database (<http://www.hydro.eaufrance.fr>), where flow measurements are available for more than 4000 stations across France [Leleu et al., 2014] and the UK National River Flow Archive for which streamflow data for more than 1500 stations are available. Since the GSIM database contains catchments from the national databases (France, UK, US) we consider extracting data from these national databases preferable and excluding catchments for these countries from the GSIM database, to avoid any duplicated catchments. The selection was made following two criteria:

- (1) Availability of at least five years of annual streamflow data between 1992 and 2015. This criterion stems from the use of pluriennial water balance models and the availability of land use historical data.
- (2) Catchment area between 10 km<sup>2</sup> and 10,000 km<sup>2</sup>. The motivation for the lower limit was due to the spatial resolution of the climate reanalysis data (1/24°), and the upper limit was set so that reduce mixing effects of both land use and climate over a given catchment.

Applying the above criteria led to a selection of 6002 catchments. As many of these catchments are nested within each other, this poses the problem of redundant information that may bias model evaluation using classical cross-validation experiments. To remove redundancy, we subset the database to reach a final set of 5026 non-nested catchments (Figure 1). The catchment represents a large variety of climatic conditions but with a lack of representation of drylands, originating from a lack of river network in these areas. The mean number of available years over the catchment set is 18 and the mean catchment area is 1800 km<sup>2</sup>. According to the Koppen classification,

majority of catchments (64%) are under mild temperate climate, 19% are under snow environment, 12% are under tropical climate and only 3% and 2% are under dry and polar climates respectively.

Climatic data were extracted from the TerraClimate gridded data [Abatzoglou et al., 2018], with a spatial resolution of  $1/24^\circ$ . Annual precipitation ( $P$ ), Penman–Monteith potential evaporation ( $PET$ ), and net radiation ( $Rn$ ) were retrieved for each catchment over the period 1992–2018 and temporally aggregated at an annual time scale. TerraClimate uses climatically aided interpolation, combining high-spatial-resolution climatological normals from the WorldClim dataset, with monthly data from other sources namely CRU Ts4.0 and JRA-55.

Land use data were extracted from the ESA CCI Land Cover time-series [ESA, 2017], for which historical land use maps are available at the annual time-scale over the period 1992–2018 at a spatial resolution of 300 m. This dataset was produced in two steps: first, the Medium Resolution Imaging Spectrometer (MERIS) archive from 2003 to 2012 was used to construct a 10-year baseline land cover map, second, the annual time series are obtained by combining the baseline map with land cover changes detected from (i) Advanced Very-High-Resolution Radiometer (AVHRR) time series from 1992 to 1999, (ii) SPOT-Vegetation (SPOT-VGT) time series from 1998 to 2012 and (iii) PROBA-Vegetation (PROBA-V) and Sentinel-3 OLCI (S3 OLCI) time series from 2013 to 2019. The original classification of ESA CCI Land Cover contains 21 classes (Table 1). A 6-level aggregation product was derived using the original classes into IPCC classes that are conventionally used to detect land use change at the global scale. Though highly simplified, this 6-level land use classification appears as a reasonable compromise between the expected role of land use on evaporation and the representativeness of each land use over the catchment set (Figure 2).

## 2.2. Tixeront–Fu formulation and benchmark water balance models

Several equations/models were developed under the Budyko framework. In this study, we used the so-called Tixeront–Fu parametrized equation [Fu, 1981, Tixeront, 1964], which was used in many previous studies [Li et al., 2013, Teuling et al., 2019, Zhang

et al., 2004]. This formulation includes one free parameter and we hypothesized in this study that this parameter may depend on land use. Three calibration-free models, namely the Schreiber [1904], the Ol’Dekop [1911] and the Budyko [1974] equations are also used as benchmarks. The equations of these models are presented in Table 2.

## 2.3. Tixeront–Fu parameter determination

Li et al. [2013] reviewed numerous studies aiming at determining the parameter  $\omega$  of the Tixeront–Fu equation. As an integrator of many physical processes in water and energy budgets,  $\omega$  can potentially be related to many physiographic patterns of a catchment: land surface characteristics, including vegetation, soil types, and topography, as well as climate intra-annual variability (e.g. seasonality of  $P$  and  $PET$ ). Given that land use reflects integrated landscape properties, we hypothesized in this study that prescribing a single value of  $\omega$  for each land use class might improve the performance of the Tixeront–Fu model. This is done by spatially averaging the contribution of  $E/P$  for each land use class present within each catchment, following (5), and illustrated in Figure 3:

$$\frac{E}{P} = \sum_{i=1}^{n_{LU}} \frac{S_i}{S} f\left(\frac{PET}{P}, \omega_i\right) \quad (5)$$

where  $f((PET/P), \omega_i)$  is the Tixeront–Fu equation (4) that requires estimation of the aridity index ( $PET/P$ ) and a value for the parameter  $\omega_i$  specific for the land use class  $i$ ,  $S_i/S$  represent the fractional area of land use class  $i$  over the total catchment area and  $n_{LU}$  is the total number of land use classes considered. As the proposed water balance models are designed for pluriennial time step, climatic variables and the fractional area of each land use class are averaged over the the record period by arithmetic mean of annual values.

Given (5), the number of free parameters is equal to the number of land use classes considered, i.e. six classes. The calibration of these parameters was dealt with a single objective function based on the root mean square error over the catchment set (6) over a range of acceptable values of  $\omega$  from unity to five. Below this lower limit, the physical basis of the equation is lost and above the upper limit, the

**Table 1.** Land use classes considered in this paper

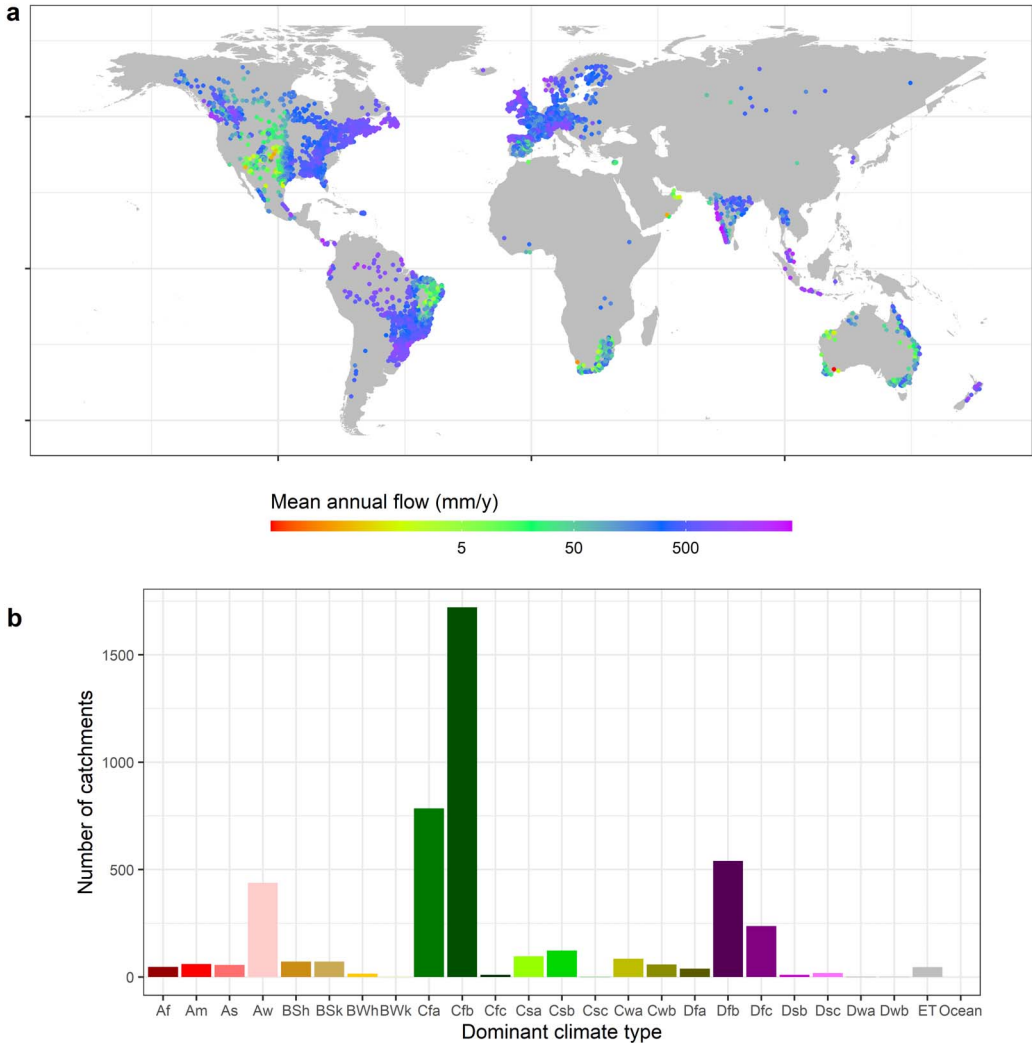
IPCC classes	Original legend from CCI-LC maps
1. Agriculture	1. Rainfed cropland 2. Irrigated cropland 3. Mosaic cropland (dominant) with natural vegetation (tree, shrub, herbaceous cover) 4. Mosaic natural vegetation (dominant) and cropland
2. Forest	5. Tree cover, broadleaved, evergreen 6. Tree cover, broadleaved, deciduous 7. Tree cover, needle-leaved, evergreen 8. Tree cover, needle-leaved, deciduous 9. Tree cover, mixed leaf type (broadleaved and needle-leaved) 10. Mosaic tree and shrub (dominant) with herbaceous cover
3. Wetland	11. Tree cover, flooded, fresh, or brackish water 12. Tree cover, flooded, saline water 13. Shrub or herbaceous cover, flooded, fresh-saline, or brackish water
4. Grassland	14. Mosaic herbaceous cover (dominant)/tree and shrub 15. Grassland
5. Settlement	16. Urban
6. Other	17. Shrubland 18. Lichens and mosses 19. Sparse vegetation (tree, shrub, herbaceous cover) 20. Bare areas 21. Water

**Table 2.** Equations of the water balance models used in the study

Model	Equation		Parameters	Reference
Schreiber	$\frac{E}{P} = 1 - e^{-PET/P}$	(1)	None	Schreiber [1904]
Ol'Dekop	$\frac{E}{P} = \frac{PET}{P} \cdot \tanh\left(\frac{P}{PET}\right)$	(2)	None	Ol'Dekop [1911]
Budyko	$\frac{E}{P} = \left[ \frac{PET}{P} \cdot \tanh\left(\frac{P}{PET}\right) \cdot (1 - e^{-PE/P}) \right]^{0.5}$	(3)	None	Budyko [1974]
Tixeront-Fu	$\frac{E}{P} = 1 + \frac{PET}{P} - \left(1 + \frac{PET}{P}\right)^{1/\omega}$	(4)	$\omega$	Tixeront [1964]; Fu [1981]; Zhang et al. [2004]

$P$  represents the mean annual precipitation,  $E$  represents the mean annual evaporation and  $PET$  represents the mean annual potential evaporation.





**Figure 1.** (a) Location of the 4539 hydrometric stations of the catchment set and mean annual flow over the record period and (b) dominant Koppen climate type distribution of the catchment set.

equation is only slightly modified. This objective function is minimized during the calibration process.

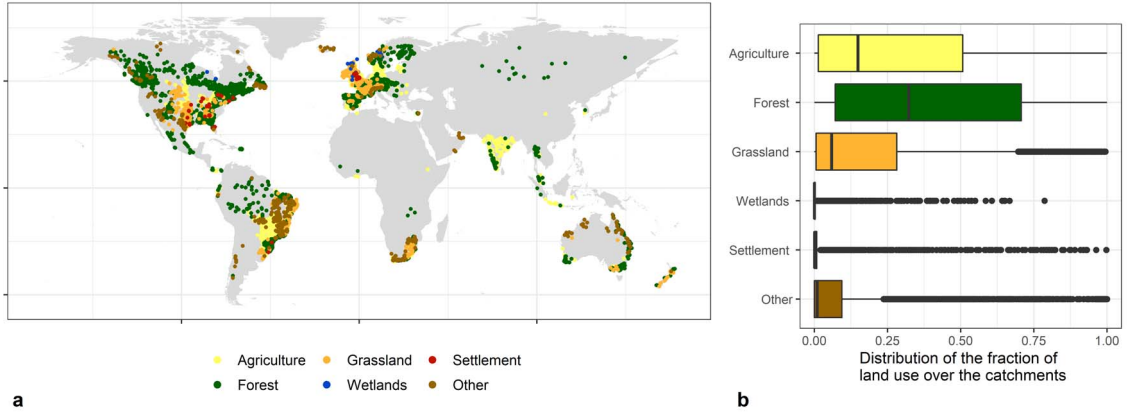
$$RMSE = \sqrt{\sum_{i=1}^{n^{BV}} (\hat{y}_i - y_i)^2} \quad (6)$$

with  $y_i = E_i/P_i$  represents the mean evaporation ratio for catchment  $i$ . The simulated mean catchment ratio  $\hat{y}_i$  is obtained by applying (5) and the “observed” mean catchment evaporation ratio  $y_i$  is obtained by solving the water budget at the catchment scale (7), considering that at the pluriennial

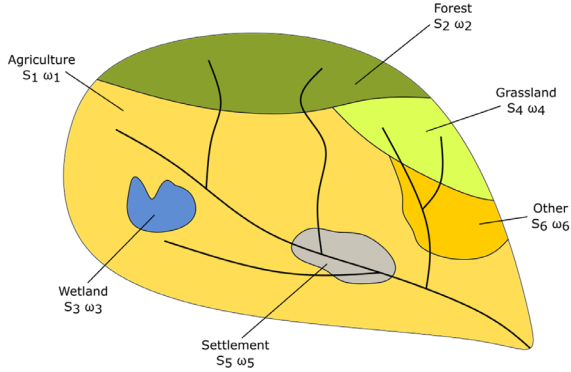
time scale, runoff ( $Q$ ) is equal to precipitation ( $P$ ) less evaporation ( $E$ ).

$$y = \frac{E}{P} = \frac{P - Q}{P} \quad (7)$$

We assessed the robustness of the calibrated parameters by performing 500 realizations of split-sample tests. For each realization, half of the catchments are randomly selected from the catchment set and are then used for calibration while the other half is used to assess the model performance in validation. This procedure also allows assessing the uncertainty



**Figure 2.** Map of catchments' (a) dominant land use and (b) distribution of fraction cover for each of the 6 land use classes over the catchment set.



**Figure 3.** Spatial averaging of the contribution of  $E/P$  for each land use class in the Tixeront-Fu model (5).

of the calibrated parameter values by analyzing the spread of the distribution of the parameter across the 500 realizations.

#### 2.4. *Attributing changes in evaporation rates to climate and land use changes*

For each catchment presenting more than 20 years of data ( $n = 2462$ ), the record period was split into two independent periods (noted "1" and "2"). The change in observed evaporation ratio ( $\Delta y$ ) between periods 1 and 2 was decomposed into a change related to climatic variability ( $\Delta y^c$ ) between the two periods, and a residual change ( $\Delta y^r$ ) that can be com-

pared to the estimated change due to land use conversions ( $\Delta y^{LU}$ ):

$$\Delta y = y_2 - y_1 = \Delta y^c + \Delta y^r \quad (8)$$

To determine the change of evaporative ratio due to climatic variability ( $\Delta y^c$ ), the theoretical evaporative ratio of the second subperiod is determined. To do so, the parameter  $\omega$  is calibrated over the first subperiod, and the Tixeront-Fu equation is applied using this parameter value and the aridity index of the second subperiod. The change related to climatic variability ( $\Delta y^c$ ) was then estimated by subtracting the evaporative ratio of the first subperiod to the theoretical evaporative ratio of the second subperiod:

$$\Delta y^c = \hat{y}\left(\omega_1, \frac{PET_2}{P_2}\right) - y_1 \quad (9)$$

By combining (8) and (9), the residual change can be estimated:

$$\Delta y^r = y_2 - \hat{y}\left(\omega_1, \frac{PET_2}{P_2}\right) \quad (10)$$

Using  $\omega$  parameter values derived from land use fractions for each subperiod, it is possible to determine a theoretical change of  $\omega$  due to land use changes between the two subperiods. This is done by exploiting the calibrated  $\omega$  values for each land use stemming from the spatial calibration of Section 2.3, noted hereafter  $\omega_1^{LU}$  and  $\omega_2^{LU}$ . These prior estimations of  $\omega$  parameter are then used to make vary the parameter value obtained by local calibration over the first subperiod by applying simply a ratio of change:

$$\widehat{\omega}_2 = \omega_1 \cdot \frac{\omega_2^{LU}}{\omega_1^{LU}} \quad (11)$$

And the change in evaporative ratio due to land use is then computed:

$$\Delta y^{LU} = \hat{y}(\hat{\omega}_2, \frac{PET_2}{P_2}) - \hat{y}(\omega_1, \frac{PET_2}{P_2}) \quad (12)$$

### 3. Results

#### 3.1. Catchment water balance analysis under the Budyko framework

In this section, the physical plausibility of the water balance at each catchment is assessed. Three categories of catchments are identified on the Budyko graph and located in Figure 4.

299 catchments (i.e. 6% of the catchment set) present inconsistent water balance ( $Q > P$ ). Those catchments are mainly located in high relief contexts (Figure 4) where problems in estimating precipitation (and particularly underestimation of snowfall) accurately are frequent. Other plausible reasons are uncertainties in streamflow estimation, wrong catchment delineation, and/or unknown interbasin groundwater flows. 150 catchments (i.e. 3% of the catchment set) present inconsistent energy balance ( $E > Rn$ ).  $Rn$  was chosen as a more physical limit of  $E$  compared to potential evaporation ( $PET$ ) since analysis on  $PET$  might be largely influenced by the formulation used to derive  $PET$ . As a comparison, 364 catchments present  $E$  larger than Penman–Monteith  $PET$ . These catchments are mainly located in humid energy-limited regions, where latent heat flux (or  $E$ ) are deemed close to net radiation (or  $PET$ ). Except for this trivial observation, no clear patterns emerged from the location of these catchments since the reasons for violating the energy budget are several, including measurement errors of precipitation, streamflow, or net radiation. The remaining 4538 (i.e. 91% of the initial catchment set) present consistent water and energy balance. Still, these catchments can present other inconsistencies (which remain unknown here), and many of these catchments are close to the theoretical limits in terms of water and energy budget. Since there is no automatic/objective way to investigate the plausibility of water and energy balances on the remaining 4538 catchments, no further investigation was conducted on these catchments and the analyses presented in the forthcoming sections are restricted to those 4538 catchments.

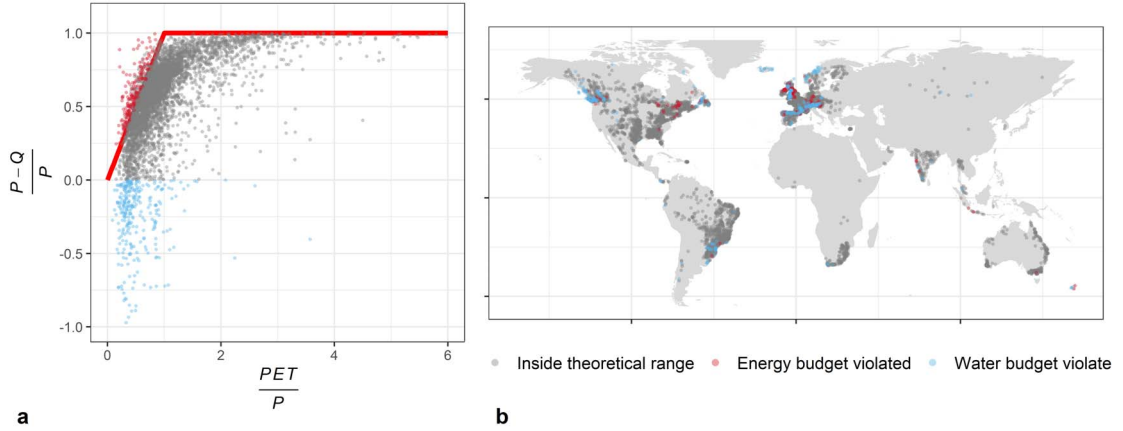
#### 3.2. A land use dependent model using a large sample of catchments

Calibration results indicate relatively similar RMSE values for the different tested water balance models (Table 3). RMSE values range from 0.150 (for the Tixeront–Fu with calibrated  $\omega$  specific for 6 land use classes) to 0.181 (for the Ol'Dekop model). Increasing the number of free parameters allows to achieve more satisfying performance, but the calibration of a single parameter for all land use classes only very marginally increases performance compared to the original non-calibrated models. The improvement brought by including land use information is about 0.003 in terms of RMSE (from 0.153 for the Tixeront–Fu model with the same  $\omega$  calibrated for all land uses to 0.150 for the Tixeront–Fu model with calibrated  $\omega$  specific for 6 land use classes), representing a relative decrease of 2%. The coefficient of determination  $R^2$  provides a similar assessment and it increases from 0.544 for the non-calibrated Schreiber model to 0.571 for the Tixeront–Fu model with 6 land use classes.

The differences in terms of prediction of the evaporative ratios are quite small but visible when analyzing the correlation between observed and simulated evaporative ratios (Figure 5). The analysis of the residuals of the Tixeront–Fu model with calibrated  $\omega$  specific for 6 land use classes does not allow identifying some specific conditions for which the model fails (Figure 6). The model appears less efficient on small catchments compared to large catchments and it also encounters difficulties in cold regions. The distributions of the residuals are relatively homogenous along with land use classes. Catchments where wetlands are dominant, are poorly modeled but these catchments are very few ( $n = 10$ ).

For all models, the RMSE obtained in validation are quite similar to those obtained in calibration (Figure 7), showing that the models are spatially robust, i.e., the spatial transferability of the model parameters is satisfying. This is reassuring and confirms that the dataset is broad enough to draw stable and significant conclusions. In validation, the advantage of a parameter-specific land use classification remains.

The parameter values obtained by calibrating the Tixeront–Fu model equation with 6 different land use classes are rather spread over the range of possible values, from close to unity to 5 (Figure 8), and rel-



**Figure 4.** (a) Catchment water balance in the non-dimensional Budyko graph and (b) location of their corresponding hydrometric stations. Catchments where the evaporation ratio is lesser than zero indicate that the mean annual runoff is greater than the mean annual precipitation received by this catchment, hence violating the water budget (blue points), while catchments where latent heat flux exceeds net radiation (red points) violate the energy budget. Note that 31 catchments present evaporative ratio lesser than  $-1.0$  and are not displayed on panel (a).

**Table 3.** Calibration results over the entire catchment set

Model	RMSE	$R^2$	Number of calibrated parameters
Schreiber [1904]	0.153	0.544	0
Ol'Dekop [1911]	0.181	0.561	0
Budyko [1974]	0.158	0.557	0
Tixeront–Fu with the same $\omega$ calibrated for all land uses	0.153	0.555	1
Tixeront–Fu with calibrated $\omega$ specific for 6 land use classes	0.150	0.571	6

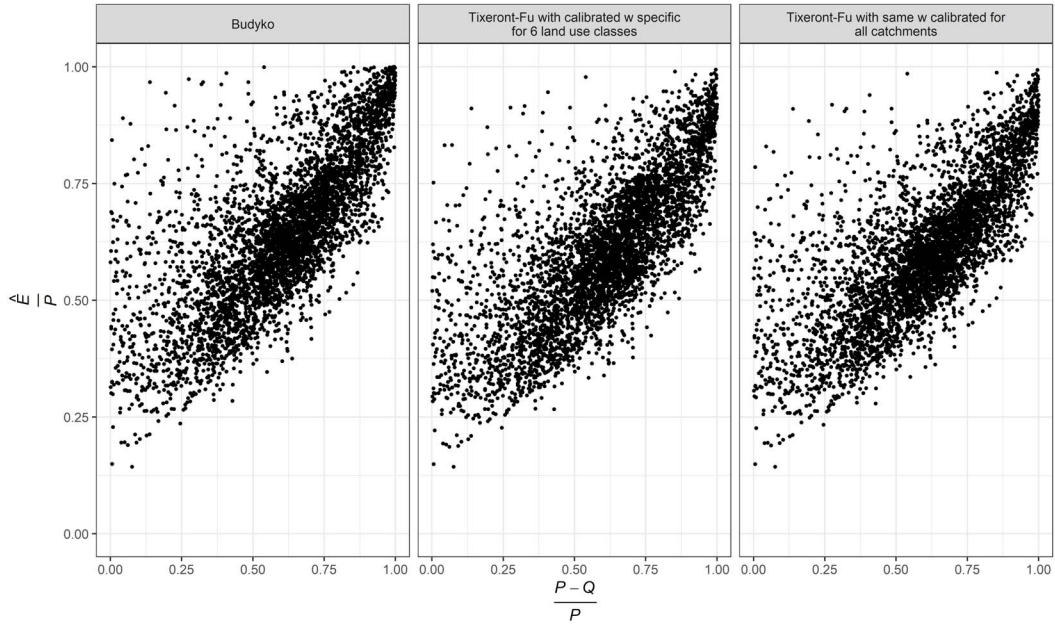
atively consistent with *a priori* expectations. The “Settlement” land use class reaches a low median parameter value of 1.78 (interquartile range: 1.72–1.83) indicating low evaporation ratios, while flooded vegetation gets a median parameter value of 5.00 (interquartile range: 3.54–5.00), indicating evaporation close to potential evaporation, which is physically consistent given that water is not a limiting factor on these landscapes. Forest and grassland show close parameter values, with median values at 2.08 and 2.11 and interquartile ranges of 2.06–2.10 and 2.09–2.14, respectively. Among the most common land use, agriculture gets a remarkably high parameter value, with a median at 2.81 (interquartile range: 2.79–2.84). The large range of parameter values for the land use class “Other” is probably explained by the large diversity of land use composing this class, and their diverse behavior relative to evaporation

(e.g. sparse vegetation, shrublands, bare areas, see Table 1). As a consequence, the median parameter value is close to the optimal lumped value, with a median at 2.33 (interquartile range 2.27–2.43).

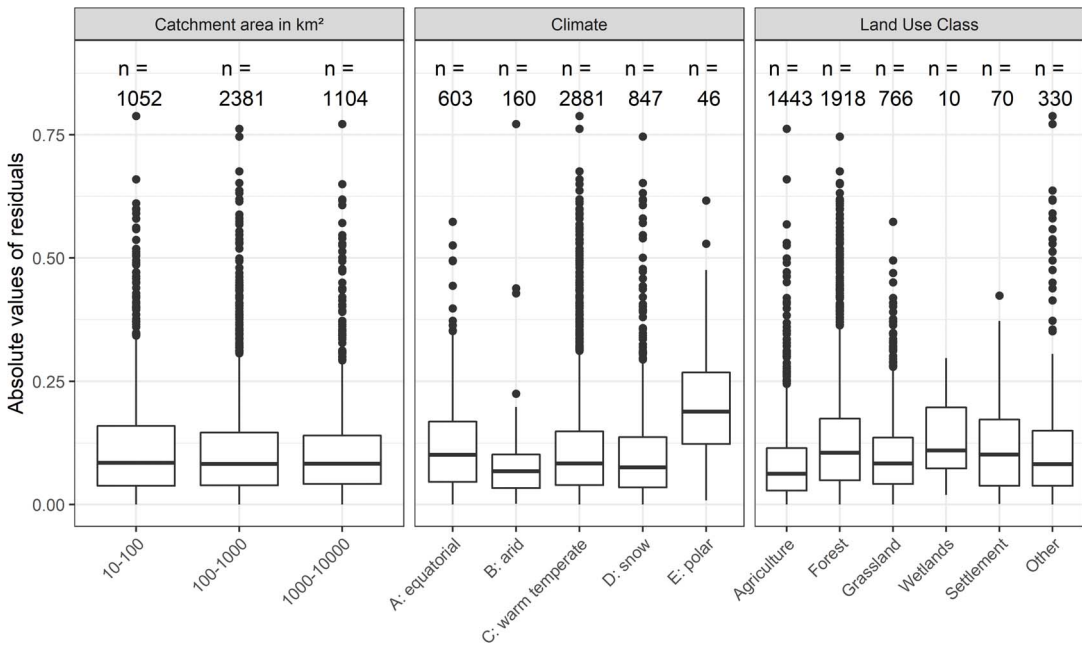
### 3.3. Trading space for time: analyzing model behavior on transient catchments

In this section, we assess the ability of the Tixeront–Fu model equation with 6 different land use classes to predict the effect of land use changes on evaporation ratios. To do so, we compared the change in evaporative ratio due to land use (12), to the residual change in evaporative ratio computed as the difference between the total change, and the change due to climate variations across the two subperiods (10).

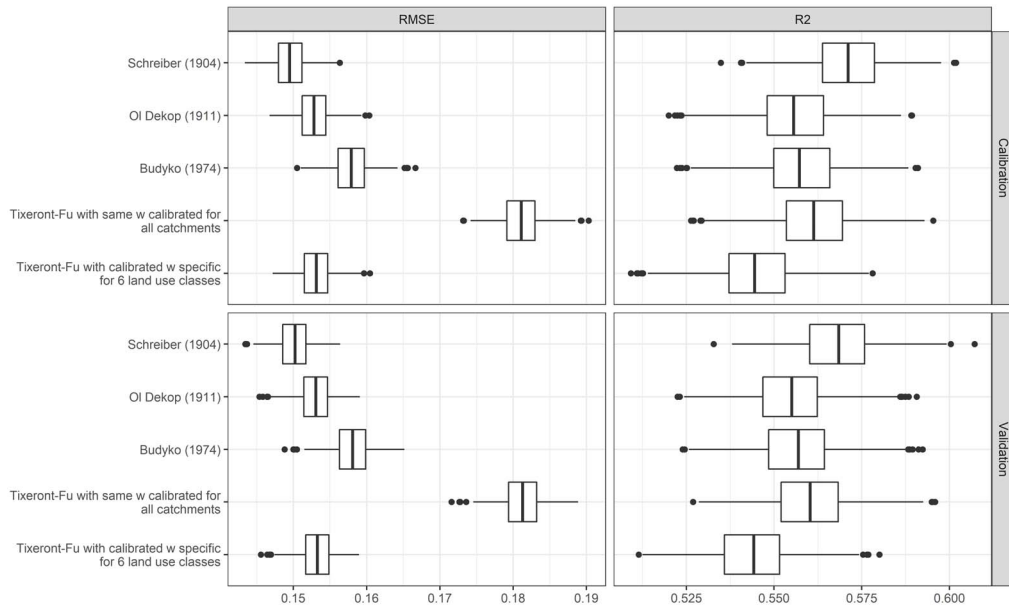
For many catchments, the Tixeront–Fu model parametrized to take into account explicitly land use



**Figure 5.** Observed and simulated evaporative ratios by three different models with zero to six calibrated parameters. Results are shown in calibration mode.



**Figure 6.** Residuals of the Tixeront–Fu model with calibrated  $\omega$  specific for 6 land use classes. The distributions of the absolute values of the residuals are represented by boxplots according to the characteristics of the catchment in terms of drainage area, climate class, and dominant land use class. The lower and upper hinges correspond to the first and third quartiles and the upper and lower whiskers extend from the hinge to the largest value no further than  $1.5 * IQR$  from the hinge (where IQR is the interquartile range).



**Figure 7.** Cross-validation results for the three benchmark models and the two calibrated Tixeront–Fu models. The boxplots represent the distribution of the two assessment criteria (RMSE and  $R^2$ ) for the 500 realizations of the split sample test. The lower and upper hinges correspond to the first and third quartiles and the upper and lower whiskers extend from the hinge to the largest value no further than  $1.5 * \text{IQR}$  from the hinge (where IQR is the interquartile range).

typology fails to predict change in evaporative ratio due to land use change (Figure 9). Generally speaking, the estimated change is quite small compared to the observed change, meaning either that the model inappropriately reproduces the role of land use on evaporative ratios or that the observed change is due to factors other than climate and land use.

To address more specifically the former hypothesis, we restrict our analysis to those catchments that experienced important gross land use change. Three typical land use conversions are considered: (i) the conversion from forest to arable land, to constitute this sample, we retained the catchments for which the fraction of arable land increased up to 0.1 while the fractions of forest decreased up to 0.1; (ii) the conversion from arable land to forest, to constitute this sample, we retained the catchments for which the fraction of forest increased up to 0.1 while the fractions of arable lands decreased up to 0.1; (iii) the conversion of vegetated lands to urban areas, to constitute this sample, we retained the catchments for which the fraction of urban area increased up to 0.1 while the fractions of other vegetated areas (forest,

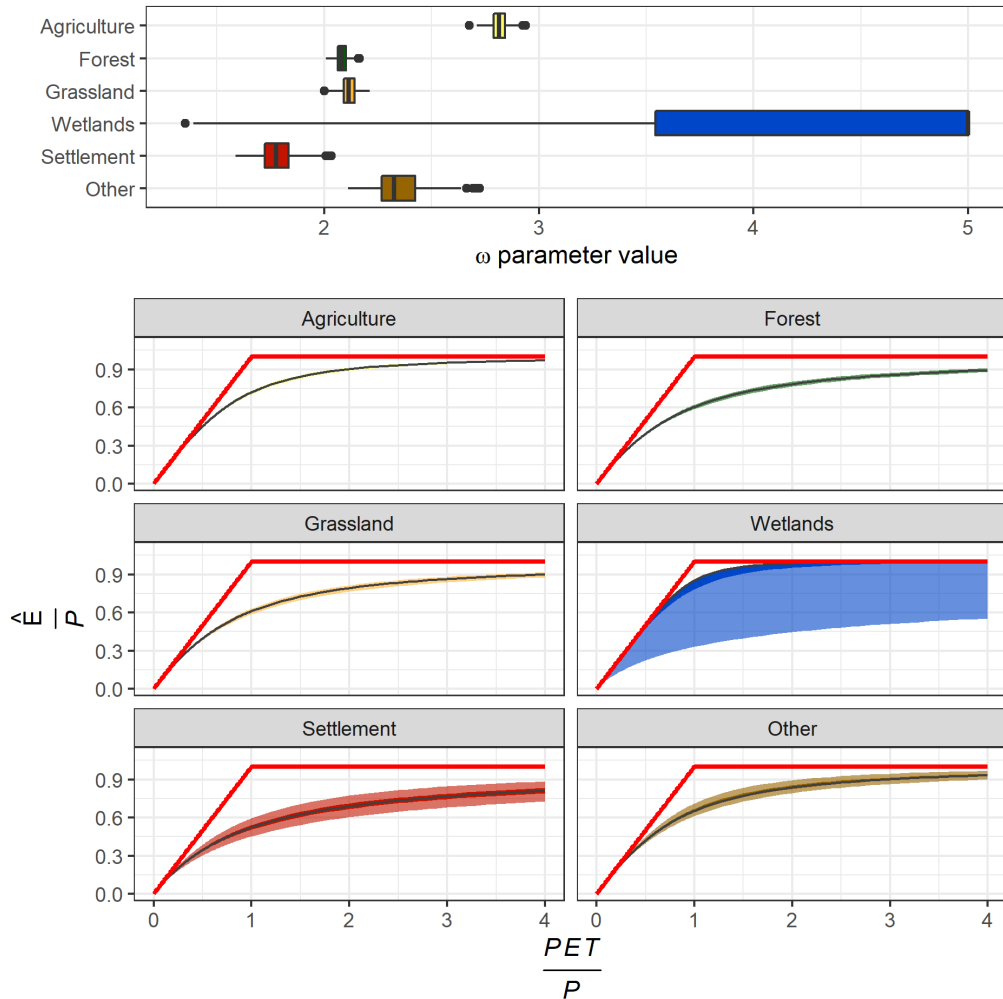
arable and grasslands) decreased up to 0.1. This leads to the selection of 52 catchments: 27 urbanizing, 10 from agriculture to forest, and 15 from forest to agriculture.

Since the selected catchments are indeed experiencing important land use change, the simulated change of evaporative ratio is more visible (Figure 10) than when considering the whole catchment set. However, the model still fails to detect the sign and intensity of change for the conversion from arable land to forest and vice-versa. Urbanization is the sole type of conversion for which the model can reproduce the effect on evaporative ratio, with a decreasing trend of evaporation with urbanization.

## 4. Discussion

### 4.1. Comparison with previous model parametrization and performance

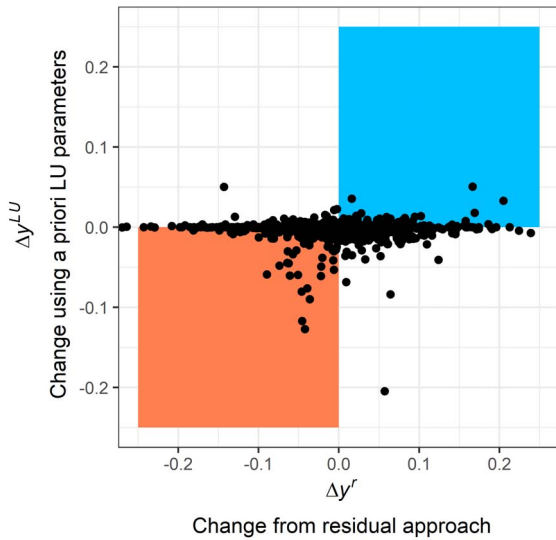
Before comparing the modeling results of this study with previous attempts, it is important to keep in mind that such comparison is complex due to (i) the



**Figure 8.**  $\omega$  parameter values for 6 land use classes (upper). The boxplots represent the distribution of the 500 split-sample tests. The lower and upper hinges correspond to the first and third quartiles and the upper and lower whiskers extend from the hinge to the largest value no further than  $1.5 * \text{IQR}$  from the hinge (where IQR is the interquartile range). Corresponding Budyko curves (lower) with predictive bounds, based on the interquartile ranges.

relatively small number of large sample studies, (ii) the different target variables within these studies, and (iii) the different data prefiltering methods. We found that 449 catchments (9% of the total catchment sample) did not fall within the theoretical range for the evaporative ratio. This ratio is similar to previous studies on the global scale: Peel et al. [2010] and Zhou et al. [2015] found ratios of 19% and 10% for global streamflow datasets constituted by 861 and 1928 worldwide catchments, respectively. The choice of including these catchments or not for developing

and assessing a hydrological model is not trivial. On the one hand, including these catchments may lead to greater uncertainties in the estimations of model parameters value but on the other hand, excluding these catchments poses the question of the validity of the remaining catchments, that are still susceptible to have water balance issues. For these reasons, some authors included further data checks, including assessments of the impacts of dams or preliminary estimation of interbasin groundwater flows [de Lavenne and Andreassian, 2018], which may affect signifi-



**Figure 9.** Observed and simulated changes in evaporative ratios as estimated by the residual approach and the Tixeront–Fu model with 6 different land use classes. Dots located in blue and red rectangles correspond to catchments for which the evaporative trend sign is correctly predicted either positive or negative, respectively.

cantly water balance for some geological settings.

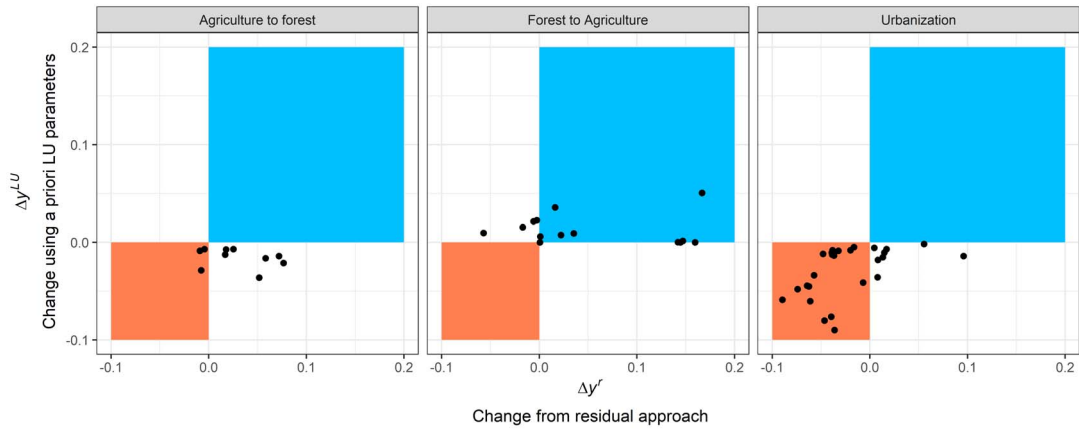
This choice also has important consequences on the performance of the model since water balance formulations from the Budyko framework are not able to reproduce the behaviors of the catchment that do not fall within the theoretical range. As a consequence, studies that prefilter hydroclimatic data before assessing model performance show remarkably better modeling results than those without or with basic filtering (Table 4). Probably the most similar study in terms of data is the study by Beck et al. [2015] who developed dedicated neural networks for a set of flow characteristics including mean annual runoff (QMEAN) and runoff yield (i.e. QMEAN/PMEAN). To compare our results with previous studies, we compute additional performance metrics, including  $R^2$  and RMSE on several mean annual variables. The obtained performance metrics are very similar to those obtained by Beck et al. [2015] with a very different modeling approach (Table 4).

#### 4.2. Is the role of land use significant and relevant?

As stated in Section 3.2, the improvement brought by incorporating specific  $\omega$  parameter value for each land use is modest: RMSE on evaporative ratio-y varied from 0.153 for the Tixeront–Fu model with the same  $\omega$  calibrated for all land uses to 0.150 for the Tixeront–Fu model with calibrated  $\omega$  specific for 6 land use classes. Previous attempts to improve Budyko formulations by adding free parameters also report marginal improvements: Oudin et al. [2008] improved RMSE on the evaporative ratio of 0.007 by including vegetation parameters while de Lavenne and Andreassian [2018] got an improvement of 0.006 by including a seasonality parameter. Considering the much larger and diverse catchment set of the present study, we may consider that the improvement of model performance agrees with previous studies and pointed out the difficulty to improve Budyko-type formulations. One may argue that the modest improvement might come from the fact that land use is not as important as expected in explaining catchment-scale water budgets. From the literature results, it is not clear whether the Tixeront–Fu parameter shall be related only to land use (and particularly vegetation types) or also to other physiographic and/or climatic characteristics. Previous attempts to regionalize  $\omega$  leads to some contrasting results: while Li et al. [2013] found that  $\omega$  may be closely related to vegetation coverage, Xu et al. [2013] noticed that the variance of  $\omega$  explained by vegetation indices (namely the NDVI) seems low and Abatzoglou and Ficklin [2017] found that vegetation metrics did not improve the determination of  $\omega$  compared to some other descriptors such as the ratio of soil water holding capacity to precipitation and topographic slope.

While modest, the improvement of the model performance obtained in Section 3.2 appears significant: a Welch Two Sample  $t$ -test on RMSE distributions of the Tixeront–Fu model with the same  $\omega$  calibrated for all land uses and the Tixeront–Fu model with calibrated  $\omega$  specific for 6 land use classes leads to a  $p$ -value below 0.005. Therefore, interpreting the parameter values for each land use class makes sense. The  $\omega$  parameter values obtained in this study can be compared in range and in order with previous studies (Table 5). Investigating the  $\omega$  parameter of the





**Figure 10.** Observed and simulated changes in evaporative ratios as estimated by the residual approach and the Tixeront–Fu model with 6 different land use classes for three types of land use changes. Dots located in blue and red rectangles correspond to catchments for which the evaporative trend sign is correctly predicted either positive or negative, respectively.

**Table 4.** Comparison of model performance with previous studies

Reference	This study	This study	Beck et al. [2015]	de Lavenne and Andreassian [2018]	Oudin et al. [2008]	Sanford and Selnick [2013]
Modeling approach	Tixeront–Fu model with fixed $w$	Optimized model with 6 parameters	Neural network	Tixeront–Fu model including a seasonality index	Parametrized Budyko model	Climate-based regression
Number of catchments	4539	4539	4079	171	1508	838
Filtering of data	Basic	Basic	Basic	Advanced	Advanced	Basic
Spatial scale	Global	Global	Global	France	Global (but mainly France, US and UK)	US
RMSE (y)	0.153	0.150		0.051	0.052	
RMSE (QMEAN) (mm/yr)	172.1	169.6		51.4	52.9	
RMSE (sqrtQ)	3.87	3.82	3.91			
$R^2$ (ET_MEAN)	0.611	0.621				0.882
$R^2$ (sqrtQ)	0.800	0.804	0.88			

Tixeront–Fu water balance model on a selection of lysimeter data, Teuling et al. [2019] found for settlement the value of 1.3, for grassland/cropland 1.7, while forest presents the larger value. Zhang et al. [2004] differentiated two land use: grassland and forest with values of 2.55 and 2.84, respectively. The comparison of absolute values is biased by methodology, including the choice of the potential evaporation equation or the way to determine the parameter

values, but it appears that  $w$  parameter values are inconsistent in terms of relative values for forest, grassland, and cropland: the value we obtained for agriculture and forest are particularly high and low respectively, compared to these previous studies. Interestingly, Williams et al. [2012] using a synthesis of evapotranspiration measured across a global network of flux towers also found that grasslands and croplands have on average a higher evaporative ratio

than forested landscapes, which tends to corroborate our findings.

#### 4.3. *Why trading space for time is inefficient?*

The land use dependent model developed in this study appears inappropriate to describe the change in time of the evaporative ratio of the majority of catchments, even for those catchments that experienced important land use change. This is not good news for hydrologists who are asked to develop models able to integrate climate and land use scenarios, to help water resources planning and management [Milly et al., 2008]. In our opinion, this relative failure may be due to either the oversimplified Budyko framework, or the fact that land use loses its interest when trading space for time. These points are discussed hereafter.

While the use of the Budyko framework is advocated by many authors for attribution studies [Wang and Hejazi, 2011], its crudeness may not allow it to model land use change effect on evapotranspiration. To fit the simplicity of the Budyko framework, we designed in this paper a semi-lumped approach by combining the responses of the diverse land use fractions in a catchment. We hypothesized that the hydrological behavior of a land use class may be formulated by adjusting a single parameter of the Tixeront–Fu model by considering that this behavior for each land use class is homogenous across environmental conditions. This is probably a too crude assumption but a more detailed description of the soil-vegetation-atmosphere hydrological processes would come with heavier parametrization. Indeed, the idea that assessing land use changes can only be made through distributed model with physical bases is well-spread in literature [Beven, 2002, Wijesekara et al., 2012].

The pluriannual time step makes it impossible to integrate the seasonality effects of climate and vegetation activity. Some studies intend to parametrize the climate seasonality effect [de Lavenne and Andreassian, 2018, Potter and Zhang, 2009], but the seasonal patterns of vegetation transpiration might also be taken into account within these parametrizations, as perennial vegetation generally transpires throughout the spring, summer, and autumn seasons, while the majority of the transpiration from crops occurs during the summer.

In contrast with the one-parametrized model, the effects of vegetation change can be subtle and the impacts of some major land use changes are still subject to debate [Andréassian, 2004, Beck et al., 2013]. Land use changes might also influence water balance only until a long time period so that the new vegetation cover reaches maturity [Kuczera, 1987]. Wang [2014] proposed to distinguish two types of human-induced changes: the changes that yield measurable effect on water balance just after the activity is in places such as surface water abstraction/diversion and storage, and the changes that yield measurable effect only until a long time period. Our results suggest that urbanization may belong to the latter since its impact on the evaporative ratios is relatively well detected and quantified by the model proposed in this study. Conversely, vegetation changes may belong to the former type of changes, with highly case-specific changes.

Finally, the space for time trading concept presents inherent caveats [Berghuijs and Woods, 2016] due to the non-orthogonality of climate, land use, and other physiographic descriptors such as topography, and soil types [Troch et al., 2015, van Dijk et al., 2012, Xu et al., 2013]. As perennial vegetation may be viewed in equilibrium with environmental conditions, the spatially calibrated model parameters are likely related to other factors like vegetation, e.g. water holding capacity, slope. When applied to non-stationary conditions, vegetation covers may change while the other characteristics remain, at least for some time, which may explain to a certain extent the failure of the model under non-stationary conditions.

## 5. Conclusion

Empirical assessments of the impact of land use on water balance are not straightforward. While the Budyko framework is often advocated as a simple but efficient way to distinguish the effects of climate and land use on water balance, very few studies applied it on large samples of non-stationary catchments. In this paper, we adapted the Budyko framework to develop a land use dependent water balance model using a large sample of 4539 catchments worldwide and we then applied it under non-stationary conditions, by trading time for space. Despite slight but significant improvements of the model to differentiate the role of several land use across catchments, the

**Table 5.**  $\omega$  parameter values for several land use classes given by the present study and two other studies

Data and methods	This study Catchment scale with fractional land use	Zhang et al. [2004] Catchment scale with dominant land use	Teuling et al. [2019] Lysimeter data
$\omega$ for agriculture	2.81 (IQ range: 2.79–2.84)		1.7
$\omega$ for forest	2.08 (IQ range: 2.06–2.10)	2.84	[2.3–3.1]
$\omega$ for grassland	2.11 (IQ range : 2.09–2.14)	2.55	1.7
$\omega$ for wetlands	5.00 (IQ range: 3.54–5.00)		
$\omega$ for settlement	1.78 (IQ range 1.72–1.83)		1.3
$\omega$ for other	2.33 (IQ range 2.27–2.43)		

Note that Teuling et al. [2019] used a rather different PE equation and acknowledged that this may lead to lower  $\omega$  values.

model fails to predict both the sign and the magnitude of change in evaporative ratios due to land use changes, except the case of urbanization. In 1994, de Marsily pointed out that the constraints of Greek tragedy (uniqueness of place, time, and action) applied to empirical hydrological models. Even with the progress in model developments and availability of worldwide hydro-climatic data, we shall acknowledge that de Marsily's view is still relevant.

### Conflicts of interest

Authors have no conflict of interest to declare.

### Acknowledgements

This study utilized data from several sources. For streamflow data and catchment delineation, we used GSIM database (<https://doi.pangaea.de/10.1594/PANGAEA.887477>) that makes available data from a large number of worldwide catchments. For the U.S. catchments, streamflow data were collected from the USGS website (available at <http://waterdata.usgs.gov/>). For the French catchments, flow measurements were extracted from HydroPortail (available at <https://www.hydro.eaufrance.fr/>). For the UK catchments, streamflow and contours data were collected from the National River Flow Archive (<https://nrfa.ceh.ac.uk/data/>). Precipitation and potential evaporation data were obtained from the TerraClimate database (available at <https://www.climatologylab.org/terraclimate.html>). We thank Antoine Carlier for his help on Figure 3, the invited editors to provide us the opportunity to contribute to this special issue, Thibault Mathevet and the anonymous reviewer for their relevant suggestions.

### References

- Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., and Hegewisch, K. C. (2018). TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Sci. Data*, 5, article no. 170191.
- Abatzoglou, J. T. and Ficklin, D. L. (2017). Climatic and physiographic controls of spatial variability in surface water balance over the contiguous United States using the Budyko relationship. *Water Resour. Res.*, 53, 7630–7643.
- Andréassian, V. (2004). Waters and forests: from historical controversy to scientific debate. *J. Hydrol.*, 291, 1–27.
- Beck, H. E., Bruijnzeel, L. A., van Dijk, A. I. J. M., McVicar, T. R., Scatena, F. N., and Schellekens, J. (2013). The impact of forest regeneration on streamflow in 12 mesoscale humid tropical catchments. *Hydrol. Earth Syst. Sci.*, 17, 2613–2635.
- Beck, H. E., de Roo, A., and van Dijk, A. I. J. M. (2015). Global maps of streamflow characteristics based on observations from several thousand catchments. *J. Hydrometeorol.*, 16, 1478–1501.
- Berghuijs, W. R. and Woods, R. A. (2016). Correspondence: Space-time asymmetry undermines water yield assessment. *Nat. Commun.*, 7, article no. 11603.
- Beven, K. (2002). Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system. *Hydrol. Process.*, 16, 189–206.
- Budyko, M. I. (1974). *Climate and Life*. Academic Press, New York.

- de Lavenne, A. and Andreassian, V. (2018). Impact of climate seasonality on catchment yield: A parameterization for commonly-used water balance formulas. *J. Hydrol.*, 558, 266–274.
- de Marsily, G. (1994). Quelques réflexions sur l'utilisation des modèles en hydrologie. *Rseau*, 7, 219–234.
- Do, H. X., Gudmundsson, L., Leonard, M., and Westra, S. (2018). The global streamflow indices and metadata archive (GSIM) – Part 1: The production of a daily streamflow archive and metadata. *Earth Syst. Sci. Data*, 10, 765–785.
- Donohue, R. J., Roderick, M. L., and McVicar, T. R. (2007). On the importance of including vegetation dynamics in Budyko's hydrological model. *Hydrol. Earth Syst. Sci.*, 11, 983–995.
- ESA (2017). Land cover CCI product user guide version 2.
- Falcone, J. A. (2011). *GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow*. USGS Unnumbered Series. U.S. Geological Survey, Reston, VA.
- Fu, B. P. (1981). On the calculation of the evaporation from land surface. *Sci. Atmos. Sin.*, 5, 23–31.
- Gupta, H. V., Perrin, C., Blöschl, G., Montanari, A., Kumar, R., Clark, M., and Andréassian, V. (2014). Large-sample hydrology: a need to balance depth with breadth. *Hydrol. Earth Syst. Sci.*, 18, 463–477.
- Jaramillo, F., Cory, N., Arheimer, B., Laudon, H., van der Velde, Y., Hasper, T. B., Teutschbein, C., and Uddling, J. (2018). Dominant effect of increasing forest biomass on evapotranspiration: interpretations of movement in Budyko space. *Hydrol. Earth Syst. Sci.*, 22, 567–580.
- Klemeš, V. (1986). Operational testing of hydrologic simulation models. *Hydrol. Sci. J.*, 31, 13–24.
- Kuczera, G. (1987). Prediction of water yield reductions following a bushfire in ash-mixed species eucalypt forest. *J. Hydrol.*, 94, 215–236.
- Leleu, I., Tonnelier, I., Puechberty, R., Gouin, P., Viquendi, I., Cobos, L., Foray, A., Baillon, M., and Ndima, P.-O. (2014). La refonte du système d'information national pour la gestion et la mise à disposition des données hydrométriques. *La Houille Blanche*, 100, 25–32.
- Li, D., Pan, M., Cong, Z., Zhang, L., and Wood, E. (2013). Vegetation control on water and energy balance within the Budyko framework. *Water Resour. Res.*, 49, 969–976.
- Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., and Stouffer, R. J. (2008). Stationarity is dead: whither water management? *Science*, 319, 573–574.
- Ol'Dekop, E. M. (1911). On evaporation from the surface of river basins. In *Transactions on Meteorological Observations. Lur-evskogo*. University of Tartu, Tartu, Estonia.
- Oudin, L., Andréassian, V., Lerat, J., and Michel, C. (2008). Has land cover a significant impact on mean annual streamflow? An international assessment using 1508 catchments. *J. Hydrol.*, 357, 303–316.
- Peel, M. C. and Blöschl, G. (2011). Hydrological modelling in a changing world. *Prog. Phys. Geogr.: Earth Environ.*, 35, 249–261.
- Peel, M. C., McMahon, T. A., and Finlayson, B. L. (2010). Vegetation impact on mean annual evapotranspiration at a global catchment scale. *Water Resour. Res.*, 46, article no. W09508.
- Potter, N. J. and Zhang, L. (2009). Interannual variability of catchment water balance in Australia. *J. Hydrol.*, 369, 120–129.
- Roderick, M. L. and Farquhar, G. D. (2011). A simple framework for relating variations in runoff to variations in climatic conditions and catchment properties. *Water Resour. Res.*, 47, article no. W00G07.
- Sanford, W. E. and Selnick, D. L. (2013). Estimation of evapotranspiration across the conterminous United States using a regression with climate and land-cover data. *J. Am. Water Resour. Assoc.*, 49, 217–230.
- Schreiber, P. (1904). Über die Beziehungen zwischen dem Niederschlag und der Wasserführung der Flüsse in Mitteleuropa. *Z. Meteorol.*, 21, 441–452.
- Singh, R., Wagener, T., van Werkhoven, K., Mann, M. E., and Crane, R. (2011). A trading-space-for-time approach to probabilistic continuous streamflow predictions in a changing climate – accounting for changing watershed behavior. *Hydrol. Earth Syst. Sci.*, 15, 3591–3603.
- Teuling, A. J., de Badts, E. A. G., Jansen, F. A., Fuchs, R., Buitink, J., Hoek van Dijke, A. J., and Sterling, S. M. (2019). Climate change, reforestation/afforestation, and urbanization impacts on evapotranspiration and streamflow in Europe. *Hydrol. Earth Syst. Sci.*, 23, 3631–3652.
- Thirel, G., Andréassian, V., and Perrin, C. (2015). On the need to test hydrological models under chang-

- ing conditions. *Hydrol. Sci. J.*, 60, 1165–1173.
- Tixeront, J. (1964). Pr evision des apports des cours d'eau (Prediction of streamflow). In *IAHS Publication, Presented at the General Assembly of Berkeley*, pages 118–126. IAHS, Gentbrugge.
- Troch, P. A., Lahmers, T., Meira, A., Mukherjee, R., Pedersen, J. W., Roy, T., and Vald es-Pineda, R. (2015). Catchment coevolution: A useful framework for improving predictions of hydrological change? *Water Resour. Res.*, 51, 4903–4922.
- Turc, L. (1954). Le bilan d'eau des sols : relation entre les pr ecipitations, l' vapotranspiration et l' coulement. *Ann. Agron.*, 5, 491–595.
- van Dijk, A. I. J. M., Pe a Arancibia, J. L., and (Sam-purno) Bruijnzeel, L. A. (2012). Land cover and water yield: inference problems when comparing catchments with mixed land cover. *Hydrol. Earth Syst. Sci.*, 16, 3461–3473.
- Wang, D. and Hejazi, M. (2011). Quantifying the relative contribution of the climate and direct human impacts on mean annual streamflow in the contiguous United States. *Water Resour. Res.*, 47, article no. W00J12.
- Wang, X. (2014). Advances in separating effects of climate variability and human activity on stream discharge: An overview. *Adv. Water Resour.*, 71, 209–218.
- Wijesekara, G. N., Gupta, A., Valeo, C., Hasbani, J.-G., Qiao, Y., Delaney, P., and Marceau, D. J. (2012). Assessing the impact of future land-use changes on hydrological processes in the Elbow River watershed in southern Alberta, Canada. *J. Hydrol.*, 412–413, 220–232.
- Williams, C. A., Reichstein, M., Buchmann, N., Baldocchi, D., Beer, C., Schwalm, C., Wohlfahrt, G., Hasler, N., Bernhofer, C., Foken, T., Papale, D., Schymanski, S., and Schaefer, K. (2012). Climate and vegetation controls on the surface water balance: Synthesis of evapotranspiration measured across a global network of flux towers. *Water Resour. Res.*, 48, article no. W06523.
- Xu, X., Liu, W., Scanlon, B. R., Zhang, L., and Pan, M. (2013). Local and global factors controlling water-energy balances within the Budyko framework. *Geophys. Res. Lett.*, 40, 6123–6129.
- Zhang, L., Hickel, K., Dawes, W. R., Chiew, F. H. S., Western, A. W., and Briggs, P. R. (2004). A rational function approach for estimating mean annual evapotranspiration. *Water Resour. Res.*, 40, article no. W02502.
- Zhou, G., Wei, X., Chen, X., Zhou, P., Liu, X., Xiao, Y., Sun, G., Scott, D. F., Zhou, S., Han, L., and Su, Y. (2015). Global pattern for the effect of climate and land cover on water yield. *Nat. Commun.*, 6, article no. 5918.





---

Research article

Geo-hydrological Data & Models

# Multi-objective assessment of hydrological model performances using Nash–Sutcliffe and Kling–Gupta efficiencies on a worldwide large sample of watersheds

Thibault Mathevet<sup>Ⓜ,\*,a,b</sup>, Nicolas Le Moine<sup>Ⓜ,c</sup>, Vazken Andréassian<sup>Ⓜ,d</sup>, Hoshin Gupta<sup>Ⓜ,e</sup>  
and Ludovic Oudin<sup>Ⓜ,c</sup>

<sup>a</sup> EDF, 4 allée du Lac de Tignes, 73290 La Motte Servolex, France

<sup>b</sup> Visiting research scholar at Hydrology and Atmospheric Sciences, University of Arizona, in 2014

<sup>c</sup> Sorbonne Université, CNRS, EPHE, UMR 7619 METIS, Case 105, 4 place Jussieu, 75005 Paris, France

<sup>d</sup> Université Paris-Saclay, INRAE, UR HYCAR, 1 Rue Pierre-Gilles de Gennes, 92160 Antony, France

<sup>e</sup> Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA

*E-mails:* thibault.mathevet@edf.fr (T. Mathevet),  
nicolas.le\_moine@sorbonne-universite.fr (N. Le Moine), vazken.andreassian@inrae.fr  
(V. Andréassian), hoshin@arizona.edu (H. Gupta),  
ludovic.oudin@sorbonne-universite.fr (L. Oudin)

**Abstract.** We introduce a new diagnosis tool that is well suited to analyzing simulation results over large samples of watersheds. It consists of a modification of the classical Taylor diagram to simultaneously visualize several error components (based on bias, standard deviation or squared errors) that are commonly used in efficiency criteria (such as the Nash–Sutcliffe efficiency (NSE) or the Kling–Gupta efficiency (KGE)) to evaluate hydrological model performance. We propose a methodological framework that explicitly links the graphical and numerical evaluation approaches, and show how they can be usefully combined to visually interpret numerical experiments conducted on large datasets. The approach is illustrated using results obtained by testing two rainfall-runoff models on a sample of 2050 watersheds from 8 countries and calibrated with two alternative objective functions (NSE and KGE). The assessment tool clearly highlights well-documented problems related to the use of the NSE for the calibration of rainfall-runoff models, which arise due to interactions between the ratio of simulated to observed standard deviations and the correlation coefficient. We also illustrate the negative impacts of classical mathematical transformations (square root) applied to streamflow when employing NSE and KGE as metrics for model calibration.

---

\* Corresponding author.

**Keywords.** Hydrological modeling, Large-sample hydrology, Taylor diagram, Diagnostics, Kling–Gupta efficiency, Nash–Sutcliffe efficiency.

*Manuscript received 7 April 2022, revised 30 August 2022 and 27 October 2022, accepted 24 November 2022.*

## 1. Introduction

Hydrological models are widely used for research, engineering and water resources management. Whatever the context, modelers often share common interests in quantifying the efficiency, the robustness and the realism of models structures, improving the generality and transposability of model structures across space and time scales, and designing methods to extract information from hydrological time series for model parameter identification.

A wide panel of methodologies and numerical techniques currently exist to meet these objectives. But advanced model validation remains a key issue for model users, who apply models to make decisions. The balance between the predictive power and the physical realism of models is still a matter of debate in validation approaches (see for example the debate between Konikow and Bredehoeft [1992] and de Marsily [1994], as summarized by Andréassian [2023], this issue). In the context of growing data availability [Addor *et al.*, 2020], the efficiency and generality of models can be better assessed by using large watershed samples [Gupta *et al.*, 2014], typically consisting of several hundreds to a few thousands of watersheds [Newman *et al.*, 2015, McMillan *et al.*, 2016, Lane *et al.*, 2019, Mathevet *et al.*, 2020]. The use of large watershed samples has various benefits, among which are improved understanding based on rigorous and controlled numerical experiments allowing comparison and hypothesis testing, and improved robustness of generalization based on statistical analyses of model performance and reduced sensitivity to site-specific watershed properties.

Among the efficiency criteria commonly used to quantify the adequacy between simulations and observations, the Nash–Sutcliffe efficiency [Nash and Sutcliffe, 1970; NSE] and the Kling–Gupta efficiency [Gupta *et al.*, 2009; KGE] provide summary statistics that can be calculated on long periods or sub-periods [Mathevet *et al.*, 2020]. These metrics are often applied to model simulations considering as target variables: (i) streamflow (Q), (ii) various non-linear transformations of streamflow [square root, inverse, logarithmic or Box–Cox transformations; see Santos

*et al.*, 2018] and (iii) some hydrological signatures (hydrological regime, flood distribution, drought distribution, etc.). Some major drawbacks of the NSE have long been demonstrated: the use of a poor benchmark model to assess model performances [Schaeffli and Gupta, 2007], an unbounded formulation inappropriate for statistical analyses [Mathevet *et al.*, 2006] or an unpredictable trade-off between mean bias, variability bias and correlation while calibration [Gupta *et al.*, 2009]. The KGE has been proposed to overcome some of these problems, providing a more balanced compromise between expected properties of model simulations (bias on mean flow, bias on flow variability and correlation between simulation and observation).

When working with large watershed samples, it becomes necessary to summarize the sets of efficiency criteria values to enable effective comparison of performance obtained with different model structures, time periods or calibration options, etc. Distributions, boxplots and scatterplots are often used to visually compare numerical experiments. However, while such analyses are statistically relevant, they can drastically reduce the information content of a numerical experiment, particularly when some criteria are partially correlated. Alternatively, one may consider the Taylor diagram [Taylor, 2001], which is commonly used to simultaneously visualize several error components of a model simulation, and which has been used in climate and hydrological modeling studies. Taylor diagram is usually used in its classical formulation to compare few models or few data sets on a case study, considering correlation, standard deviation and root mean square error (RMSE) [Yaseen *et al.*, 2018, Maroufpoor *et al.*, 2020].

The objective of this paper is to propose a modification of Taylor diagram to simultaneously represent components of the KGE, i.e. mean bias, variability bias and correlation. We illustrate the application of this graphical tool using results of a numerical experiment based on the methodology proposed by Mathevet *et al.* [2020] and demonstrate its usefulness to answer the following research questions:

- What are the main performances of models calibrated using NSE or KGE as objective



function (in terms of mean bias, standard deviation bias and correlation and dependency structure between these criteria)?

- Do the performances vary with model structure or watershed location?
- What are the impacts of non-linear transformations of streamflow, classically used when calibrating rainfall-runoff models?

Section 2 presents the theoretical background of Taylor diagram and its modification, Section 3 presents an overview of the experimental design, and Section 4 details the general results of the study. Section 5 presents a discussion of the results and our conclusions are given in Section 6.

## 2. Theoretical background

### 2.1. Classical Taylor diagram

This section reviews the basics underlying the construction of a Taylor diagram. Table 1 lists the symbols used throughout this paper.

A Taylor diagram [Taylor, 2001] is intended to simultaneously display several metrics on a graph, describing the discrepancies between a test field or series (e.g., the simulation) and a reference one (e.g., the observation). These metrics consist of the (sample) Pearson correlation coefficient  $R$ , the standard deviation of the simulation  $\hat{\sigma}$  (resp. the ratio of simulated to observed standard deviation  $\hat{\sigma}/\sigma_0$ ), and  $E'$  the root mean squared error on centered patterns (CRMSE). The CRMSE differs from  $E$ , the classical RMSE, in the removal of the bias component, according to the relation:

$$E^2 = \frac{1}{n} \sum_{i=1}^n [\hat{f}_i - f_i]^2 \quad (1)$$

$$E^2 = \frac{1}{n} \sum_{i=1}^n [(\hat{f}_i - \hat{\mu}) - (f_i - \mu_0) + (\hat{\mu} - \mu_0)]^2 \quad (2)$$

$$E^2 = \frac{1}{n} \sum_{i=1}^n [(\hat{f}_i - \hat{\mu}) - (f_i - \mu_0)]^2 + \frac{2}{n} (\hat{\mu} - \mu_0) \times \left[ \sum_{i=1}^n (\hat{f}_i - \hat{\mu}) - \sum_{i=1}^n (f_i - \mu_0) \right] + (\hat{\mu} - \mu_0)^2 \quad (3)$$

with

$$E'^2 = \frac{1}{n} \sum_{i=1}^n [(\hat{f}_i - \hat{\mu}) - (f_i - \mu_0)]^2. \quad (4)$$

By definition of the sample mean, the terms  $\sum_{i=1}^n (\hat{f}_i - \hat{\mu})$  and  $\sum_{i=1}^n (f_i - \mu_0)$  sum up to zero, so that Equation (3) can be simplified to:

$$E^2 = E'^2 + (\hat{\mu} - \mu_0)^2 = E'^2 + \beta^2. \quad (5)$$

By further developing  $E'^2$  we get:

$$E'^2 = \frac{1}{n} \sum_{i=1}^n (\hat{f}_i - \hat{\mu})^2 + \frac{1}{n} \sum_{i=1}^n (f_i - \mu_0)^2 - 2 * \frac{1}{n} \sum_{i=1}^n (\hat{f}_i - \hat{\mu})(f_i - \mu_0) \quad (6)$$

$$E'^2 = \hat{\sigma}^2 + \sigma_0^2 - 2 * \hat{\sigma} \sigma_0 R \quad (7)$$

or its normalized version:

$$\left( \frac{E'}{\sigma_0} \right)^2 = \left( \frac{\hat{\sigma}}{\sigma_0} \right)^2 + 1 - 2 * \frac{\hat{\sigma}}{\sigma_0} R. \quad (8)$$

If we define a polar plane such that a given simulation is represented by the point  $P$ : ( $\rho = \hat{\sigma}$ ;  $\theta = \arccos R$ ), then according to the law of cosines, the third metric  $E'$  is given by the distance between  $P$  and the reference point  $P_0$ : ( $\rho = \sigma_0$ ;  $\theta = 0$ ) located on the cartesian  $x$ -axis. The same plot can be drawn using normalized values, the simulation and reference points being respectively defined by  $P'$ : ( $\rho = \hat{\sigma}/\sigma_0$ ;  $\theta = \arccos(R)$ ) and  $P_0$ : ( $\rho = 1$ ;  $\theta = 0$ ).

### 2.2. Adapting Taylor diagram to represent bias, NSE, total RMSE, or any monotonic function of sum of squared errors (SSE)

One limitation of the previous diagram is that it does not allow for visualizing the absolute bias along with the CRMSE, and not the bias component  $\beta = \mu_0 - \hat{\mu}$  (resp.  $\beta_n = (\mu_0 - \hat{\mu})/\sigma_0$ ). Though it is rarely used in the literature, Taylor proposed a method to display bias on the same frame. It consists of attaching a segment to the point, which length is equal to  $|\beta|$  (resp.  $|\beta_n|$  in the normalized version), as shown in Figure 1.

If the segment is oriented perpendicular to the direction ( $P_0P$ ), then the squared hypotenuse  $P_0P'^2$  of the right triangle created in this way satisfies the condition:

$$P_0P'^2 = P_0P^2 + PP'^2 \quad (9)$$

$$P_0P'^2 = E'^2 + \beta^2 \quad (10)$$

$$P_0P'^2 = E^2. \quad (11)$$

If we chose to normalize errors with the standard deviation  $\sigma_0$  of the observed streamflow series, the squared hypotenuse  $P_0P'^2$  now quantifies the total

**Table 1.** List of symbols used

Symbol	Meaning	Mathematical definition
$\mu_0$	Mean of observed streamflow	
$\hat{\mu}$	Mean of simulated streamflow	
$\sigma_0$	Standard deviation of observed streamflow	
$\hat{\sigma}$	Standard deviation of simulated streamflow	
$\beta$	Absolute mean bias	$\beta = \mu_0 - \hat{\mu}$
$\beta_n$	Normalized mean bias	$\beta_n = \frac{\mu_0 - \hat{\mu}}{\sigma_0}$
$\beta_r$	Relative mean bias	$\beta_r = \frac{\hat{\mu}}{\mu_0}$
$\alpha$	Relative variability bias	$\alpha = \frac{\hat{\sigma}}{\sigma_0}$
$E$	Root mean square error (RMSE)	
$E'$	RMSE on centered patterns (CRMSE)	$E'^2 = E^2 - \beta^2$
$\rho$	Radius (distance to origin) in the polar plane	
$\theta$	Azimet (counter-clockwise from $x$ -axis)	
$R$	Pearson correlation coefficient	
$f_i$	Observations of a given variable (streamflow in our case)	
$\hat{f}_i$	Estimations of a given variable	

*normalized* squared error and Equations (10) and (11) become respectively Equations (12) and (13):

$$P_0 P'^2 = \left(\frac{E'}{\sigma_0}\right)^2 + \beta_n^2 \quad (12)$$

$$P_0 P'^2 = \left(\frac{E}{\sigma_0}\right)^2 \quad (13)$$

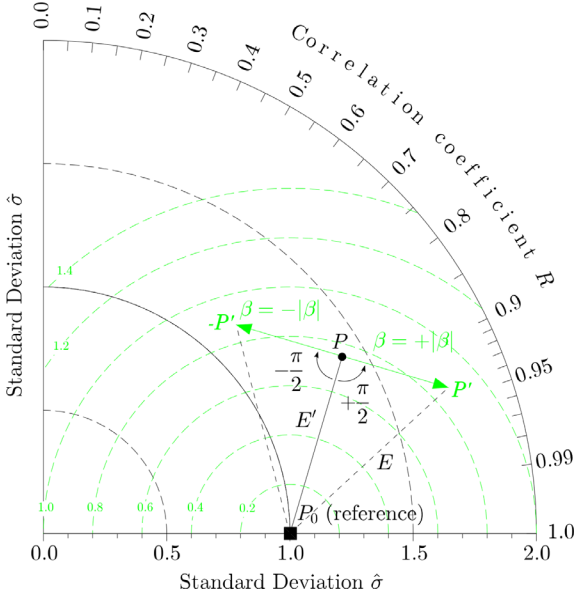
which means that with  $P'$  we now have a reading of both bias ( $|\beta|$ , leg  $PP'$  of the right triangle) and total RMSE ( $E$ , hypotenuse  $P_0P'$ ). Since the segment representing bias only has to be at a right angle with the segment reaching the reference, then we can for example choose a clockwise rotation for negative bias (i.e., an arrow pointing towards  $-P'$ , the origin of the polar plane) and counterclockwise for positive bias (i.e., an arrow pointing towards  $P'$ , the outer boundary).

Recognizing that NSE, MSE, RMSE are all monotonic functions of the total sum of squared errors, it is then easy to modify the diagram in order to display curves that are, for example, contour lines of NSE efficiency or its bounded version [Mathevet *et al.*, 2006] rather than contour lines of (C)RMSE. In this section, we show that any monotonic function of SSE

can be drawn on the normalized version of the Taylor diagram. This provides an interesting outlook of a hydrological simulation through the visualization of five assessment criteria for all the watersheds: correlation  $R$ , ratio of simulated to observed standard deviation, normalized bias  $\beta_n$ , NSE on centered patterns (CNSE) and classical NSE.

### 2.3. Prone-like regions in the modified Taylor diagram depending on the objective function

We now wish to determine in which region of the Taylor diagram a simulation result should fall, when the hydrological model is calibrated with an ordinary-least-squares (OLS) criterion such as the Nash–Sutcliffe efficiency.



**Figure 1.** Convention for representing bias and total RMSE. Green circles centered on the reference are contours of CRMSE if used with the dot  $P$ , or RMSE if used with the end tip of the green arrow ( $P'$ ). In this case where  $\sigma_0 = 1$  for the sake of simplicity, estimated standard deviation is  $\hat{\sigma} = 1.41$ , correlation  $R = 0.86$ , CRMSE  $E' = 0.75$ , bias  $\beta = \pm 0.45$  and RMSE  $E = 0.87$ .

For this, we use the decomposition of NSE proposed by Gupta *et al.* [2009], which reads:

$$\text{NSE} = 1 - \left(\frac{E}{\sigma_0}\right)^2 \quad (14)$$

$$\text{NSE} = 1 - \left(\frac{E'}{\sigma_0}\right)^2 + \beta_n^2 \quad (15)$$

$$\text{NSE} = 2 \left(\frac{\hat{\sigma}}{\sigma_0}\right) R - \left(\frac{\hat{\sigma}}{\sigma_0}\right)^2 - \beta_n^2 \quad (16)$$

$$\text{NSE} = 2\alpha R - \alpha^2 - \beta_n^2. \quad (17)$$

Equations (14)–(17) show that OLS-like criteria tend to fit the first ( $\beta = \beta_n = 0$  or in other words  $E = E'$ ) and second ( $\alpha = 1$ ) moments of the distribution of the observations in a rather straightforward way. Conversely, the role of the correlation term  $R$  that represents the ability of the model to reproduce timing and shape is mixed up with the value of  $\alpha$ . This appears clearer when deriving Equation (15) with respect to  $\alpha$  by considering  $R$  and  $\beta_n$  fixed, as shown by Gupta

*et al.* [2009]:

$$\frac{\partial \text{NSE}}{\partial \alpha} = 2(R - \alpha) = 0. \quad (18)$$

Hence, any OLS calibration will tend to equate  $\alpha$ , the ratio of simulated to observed standard deviation, with the correlation coefficient  $R$ : these two metrics are entangled. This also poses a problem since  $R$  is necessarily lesser than unity, OLS calibration will lead to  $\alpha$  less than unity, i.e. a systematic underestimation of the variance of streamflow. We will show how this dependency should translate in the Taylor diagram.

Let us consider the normalized version of the Taylor diagram, and let  $\varepsilon_{\text{OLS}}$  be the locus of the points satisfying the condition  $\alpha = R$  i.e.  $\theta = \arccos(R) = \arccos(\alpha)$ . The cartesian coordinates of those points hence verify:

$$x = \alpha \cdot \cos(\theta) = \alpha \cdot \cos(\arccos(\alpha)) = \alpha^2 \quad (19)$$

$$y = \alpha \cdot \sin(\theta) = \alpha \cdot \sin(\arccos(\alpha)). \quad (20)$$

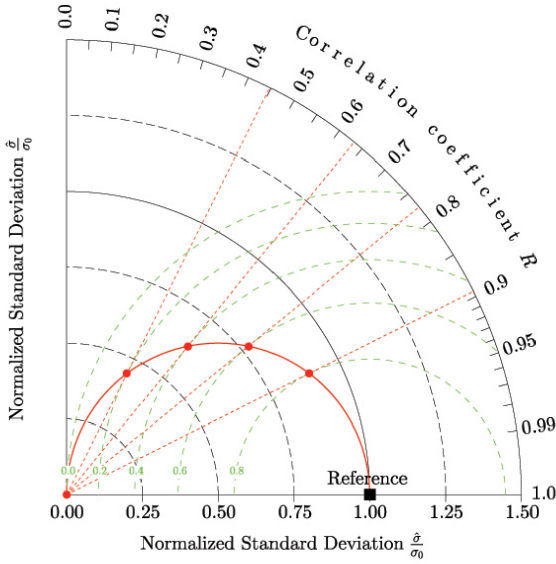
It is then easy to see that  $\varepsilon_{\text{OLS}}$  is the semi-circle centered at  $(x = 1/2; y = 0)$ , with radius  $1/2$ . Indeed, for any point in  $\varepsilon_{\text{OLS}}$  we have:

$$\left(x - \frac{1}{2}\right)^2 + y^2 = x^2 - x + \frac{1}{4} + y^2 \quad (21)$$

$$\left(x - \frac{1}{2}\right)^2 + y^2 = \alpha^4 - \alpha^2 + \frac{1}{4} + \alpha^2 \sin^2(\arccos(\alpha)) \quad (22)$$

$$\begin{aligned} \left(x - \frac{1}{2}\right)^2 + y^2 &= \alpha^2(\alpha^2 - 1) \\ &\quad + \frac{1}{4} + \alpha^2(1 - \cos^2(\arccos(\alpha))) = \frac{1}{4}. \end{aligned} \quad (23)$$

This result is illustrated in Figure 2. For every point of this circle (red dots), the radial through this point (red dashed line) is tangent to the NSE (or RMSE, or SSE, etc.) contour passing through this same point (green dash circle). We can call this curve the “*least-squares attractor in the absence of bias*” and in the following sections we will show that indeed models calibrated with a monotonic function of the sum of squared errors tend to align along this semi-circle in the polar plane. Since the red curve is systematically under the ideal fit of normalized standard deviation ( $\alpha = (\hat{\sigma}/\sigma_0) = 1$ ), this illustrates the inherent problems of underestimation of the variance of OLS schemes. As a consequence of the normalization of the axes of standard deviation, the green dashed lines now represent isolines of NSE, instead of the isolines of RMSE in Figure 1. Figure 2 shows clearly the unfair trade-off occurring during an OLS



**Figure 2.** Construction of the locus of the points satisfying  $\partial\text{NSE}/\partial\alpha = 0$  or, synonymously,  $\alpha = R$  in red. The green circles are contours of NSE.

scheme between  $R$  and the ratio of standard deviation since the least-squares attractor tends to favor a higher correlation at the expense of a systematic underestimation of  $\alpha$ . In its present form, note that the proposed modified Taylor diagram does not allow the representation of negative NSE. This was also a problem with the classical Taylor diagram that does not allow the representation of a negative correlation coefficient.

### 3. Experimental design

Here we use the same numerical experiment as in Mathevet *et al.* [2020], which is briefly summarized in this section.

#### 3.1. Watershed sample

A large sample of 2050 watersheds spanning over eight countries is used in this paper. This sample is mainly (80%) represented by watersheds in France, USA and Australia. Other watersheds are situated in Italy, Laos, Sweden, Switzerland and UK. This sample covers a variety of climatological, physical and

hydrological characteristics (see Table 2). Time series consisted of mean daily rainfall, air temperature and streamflow. Climatic data are averages at the watershed scale, but the way these averages were computed is variable between the national sub-samples of watersheds. Potential evapotranspiration (PE) was computed using the temperature-based formula proposed by Oudin *et al.* [2005].

#### 3.2. Hydrological models

Two conceptual rainfall-runoff models are used in this study: the GRX model [slightly modified from GR4J model, Le Moine, 2008 and Pushpalatha *et al.*, 2011] and the MRX model [modified from MOR-DOR model, Garçon, 1996 and Garavaglia *et al.*, 2017]. These two models have already been intensively tested, both for research and operational applications, in France and worldwide [Mathevet *et al.*, 2020]. A number of studies have shown that their structures can be efficient, with comparable performance in simulation, forecast and extrapolation modes. A snow accumulation and melt routine [from Garavaglia *et al.*, 2017] is also used in this study to represent snow processes of mountainous watersheds. Note that these two models differ in their level of complexity with 9 free parameters for GRX and 22 free parameters for MRX.

#### 3.3. Calibration and evaluation metrics

Two classical metrics were used:

- Nash–Sutcliffe efficiency (Equation (14));
- Kling–Gupta efficiency (Equation (24));

$$\text{KGE} = 1 - \sqrt{(\beta_r - 1)^2 + (\alpha - 1)^2 + (R - 1)^2}, \quad (24)$$

where  $\beta_r$  and  $\alpha$  assess the mean and variability bias of the long-term water balance, while  $R$  assesses the temporal dynamics of model simulations.

NSE and KGE metrics were used for model calibration. NSE, KGE,  $\beta_r$ ,  $\alpha$  and  $R$  were used for model evaluation. On a limited sub-sample of 635 watersheds (30% of the whole sample, to reduce computing time), a square root transformation of streamflow time-series was used to compute NSE and KGE metrics for model calibration.

**Table 2.** Distributions of characteristics of the 2050-watershed sample

Characteristics	Distribution percentiles				
	0.05	0.25	0.50	0.75	0.95
Watershed area (km <sup>2</sup> )	31	102	255	734	2785
Mean annual total precipitation (P) (mm/yr)	626	800	950	1175	1627
Part of precipitation falling as snow (-)	0	0	0.03	0.09	0.29
Mean annual air temperature (°C)	4.9	9.2	10.5	12.1	16.2
Mean annual potential evapotranspiration (PE) (mm/yr)	471	633	686	773	998
Mean annual runoff (Q) (mm/yr)	53	207	344	541	1111
Aridity index (P/PE) (-)	0.80	1.12	1.35	1.74	2.82
Runoff coefficient (Q/P) (-)	0.08	0.24	0.36	0.48	0.79
Available time series length (yr)	11	18	33	36	55

### 3.4. Testing procedure

A classical split sample test (SST) procedure was implemented [Klemeš, 1986]. For each watershed, the available time period was divided into two independent sub-periods of equivalent length. After a one-year warm-up period to minimize state initialization errors, calibration was performed on each sub-period (first half, then second half), followed by evaluation on the other sub-period (second half, then first half). On the 2050 watershed sample, the mean temporal length of calibration and evaluation periods is 14 years. This testing procedure provided 4100 calibration and evaluation periods to assess model performance. Parameter optimization was conducted by use of a genetic algorithm [Mathevet, 2005].

## 4. Results

### 4.1. Mono-objective performance assessment using distribution or boxplot analyses

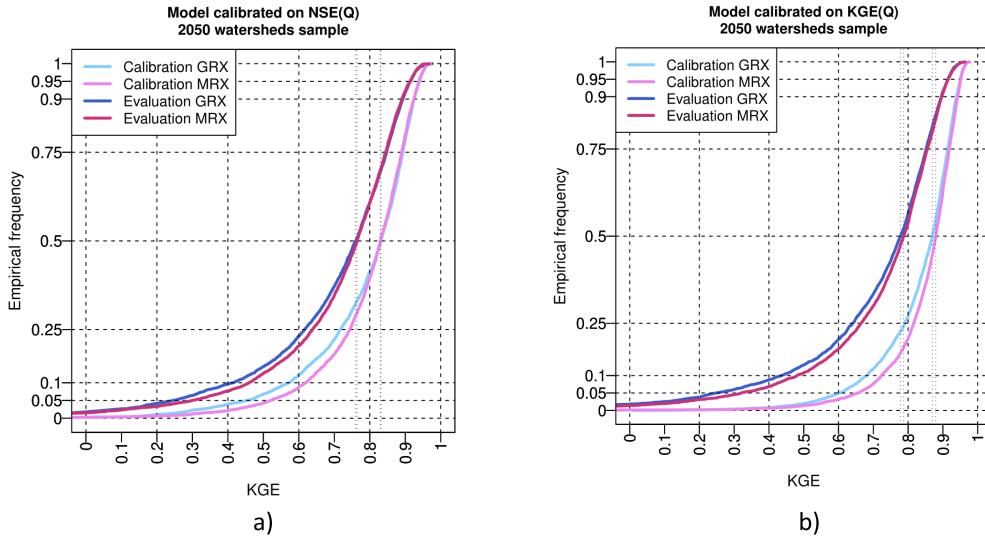
When using a large sample of watersheds, it is difficult to focus on local performance of a given model on a particular watershed, but much easier to draw general conclusions regarding comparative model performance under the hypotheses of the numeric experiments studied. Classically, samples of performance metrics are analyzed using empirical cumulative distributions (Figure 3) and distributions are summarized by boxplots (Figure 4).

Figure 3 shows the full distribution for the KGE(Q) performance metric only (for sake of brevity), while models are calibrated with NSE(Q) (Figure 3a) and KGE(Q) (Figure 3b). This figure shows that models have a similar distribution of KGE(Q) (particularly above median value) in calibration and evaluation, that models have a similar reduction of performance from calibration to evaluation and obviously that models calibrated with KGE(Q) have better performances than models calibrated with NSE(Q), while evaluated with KGE(Q).

Figure 4 shows the sample distributions of performance metrics ( $\beta_r$ ,  $\alpha$ ,  $R$ , KGE and NSE) for the calibration and evaluation periods (i.e. 4100 values per period), for each of the two models calibrated either with NSE(Q) or KGE(Q) as objective function (a detailed analysis of this numerical experiment is presented in Mathevet et al. [2020]).

When using NSE(Q) as objective function, Figure 4a shows that models are not biased in terms of mean calibration. During evaluation, models could be biased in mean but they remain not biased on average (instead the spread of the distribution increases). In terms of variability bias, Figure 4a shows that models are biased on average (mean underestimation of variability by 10%) and that the spread of the variability bias distribution increases from calibration to evaluation [Gupta et al., 2009].

When using KGE(Q) as objective function, Figure 4b shows that models are not biased in terms of mean and variability during calibration. During evaluation, models could be biased in mean and variabil-



**Figure 3.** Comparison of GRX and MRX KGE performance in calibration and evaluation using (a) NSE(Q) or (b) KGE(Q) as objective function to optimize model parameters.

ity, but they remain not biased on average (the spread of the distribution increases).

Figure 4 also shows that both models have a median correlation around 0.9 and that distributions are pretty similar for the two models and from calibration to evaluation. This means that the watershed dynamics are rather equally well represented by both models, whatever the objective function used for optimization.

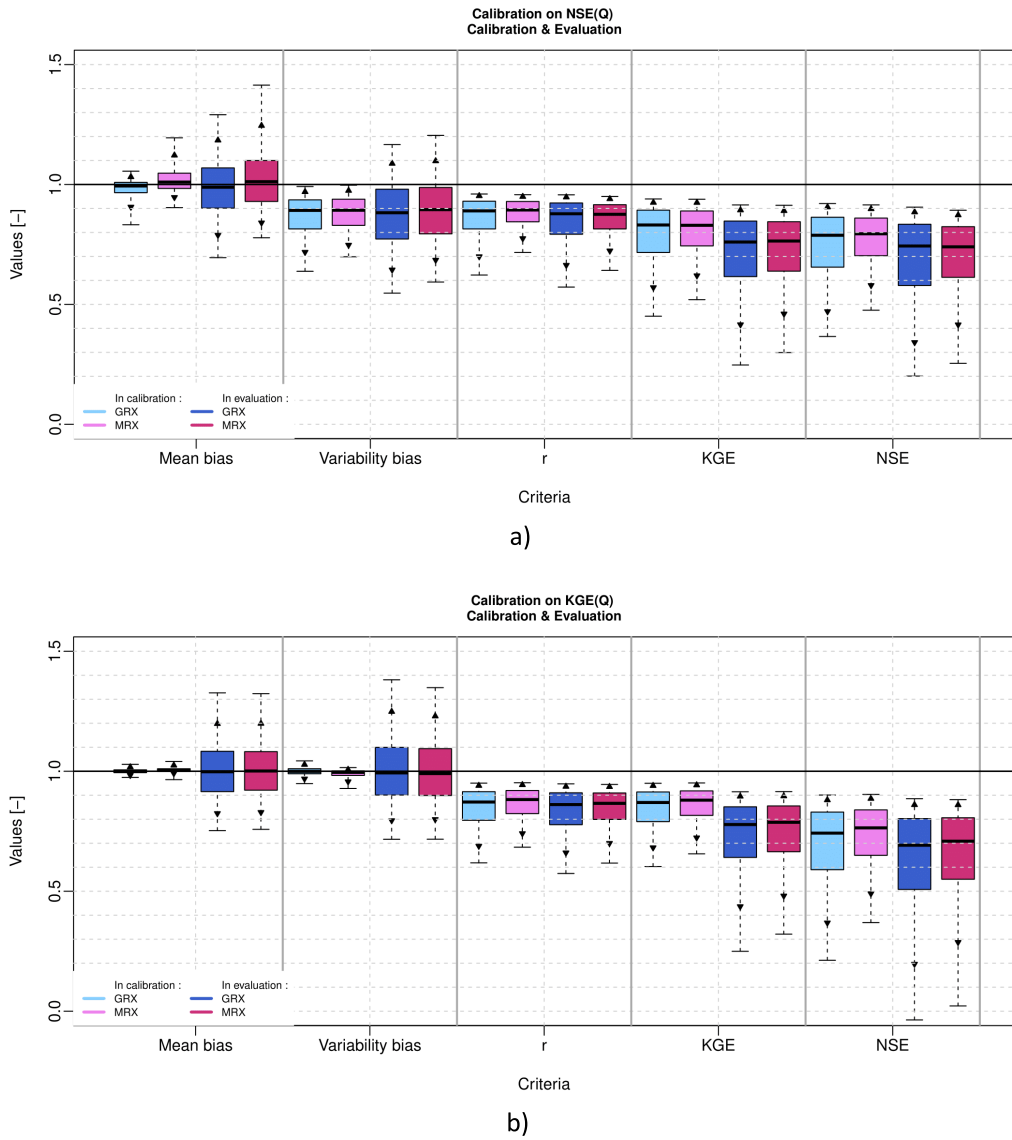
KGE(Q) and NSE(Q) distributions partially filter out some important features of individual mean and variability bias distributions and correlation distribution. While calibrated with KGE(Q), the KGE(Q) distribution is very similar to the correlation distribution since models are not biased in terms of mean and variability. From calibration to evaluation, both KGE(Q) and NSE(Q) distributions show a general decrease due to the average decline of performance concerning the representation of streamflow mean and variability.

#### 4.2. Multi-objective performance assessment using Taylor diagram

In the following part of the paper, only GRX results will be presented, since the MRX results have similar patterns and lead to similar conclusions. MRX results are presented in Appendix A.

Distributions or boxplots are useful graphical tools for the statistical analysis of model performance for a given metric. However as stated earlier (Section 2), some performance metrics might be correlated while calibrated with a given objective function (NSE(Q) as an example, or any OLS criterion).

Figure 5a shows the GRX model results when calibrated with NSE(Q) as objective function. This figure confirms that the theoretical behavior described earlier by Gupta *et al.* [2009] actually occurs when calibrating the model with the NSE(Q) criterion. In the calibration phase, hydrological models manage to get a quite small mean bias, since red arrows are barely distinguished in the left-hand side diagram of Figure 5a, given that they are very short. Mean bias in calibration is becoming greater for watersheds where models obviously fail to correctly simulate streamflow. It is generally the case of very arid watersheds, where models are less robust, with lower correlation and higher standard deviation bias [Mathevet *et al.*, 2020]. RMSE and CRMSE being very close for most of the watersheds, the colored dots characterizing the centered patterns clearly align along the blue semi-circle in the  $(R, \alpha)$  space, where the condition  $\alpha = \hat{\sigma}/\sigma_0 = R$  is satisfied. Two-dimensional contour plots show that 90% of the population lies between  $R \in [0.6, 0.98]$  and  $\alpha \in [0.5, 1.02]$ .

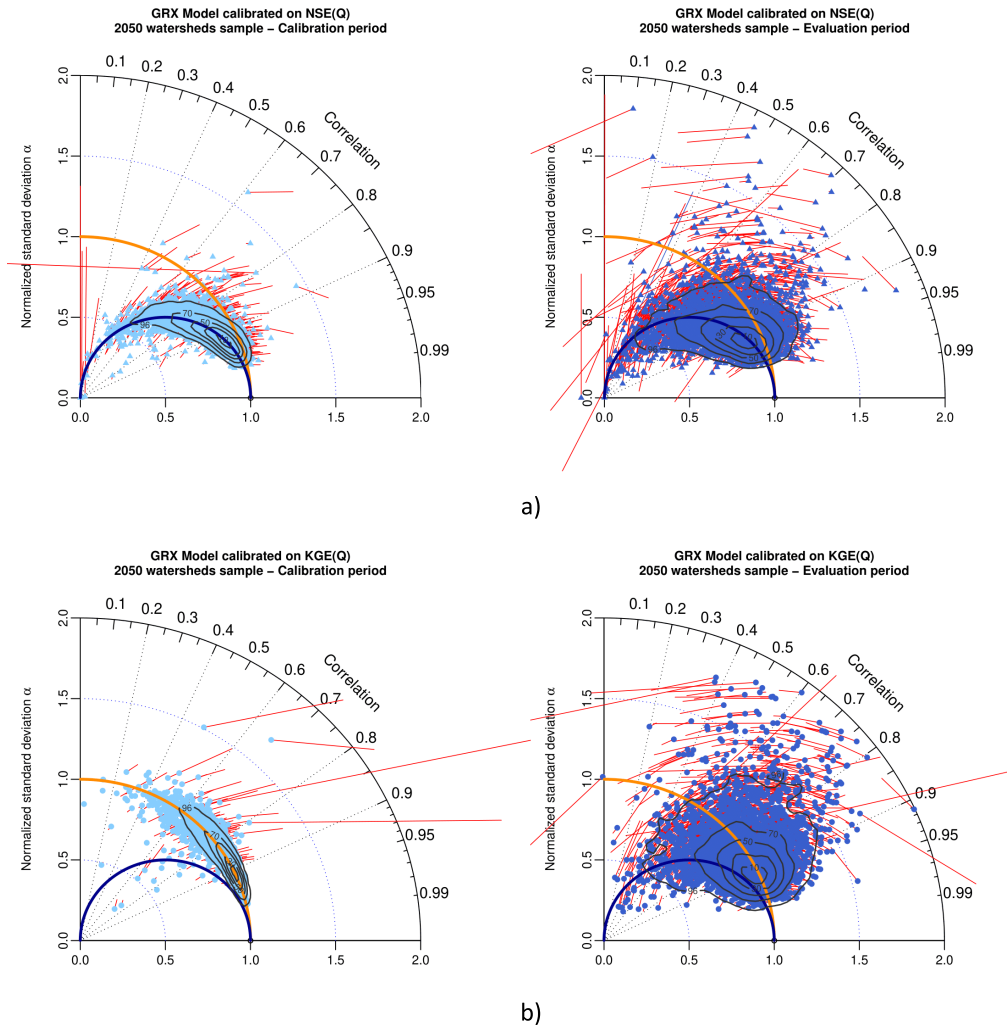


**Figure 4.** Comparison of GRX and MRX multiobjective performance (mean bias, variability bias, correlation, KGE and NSE) in calibration and evaluation using (a) NSE(Q) or (b) KGE(Q) as objective function to optimize model parameters. Boxplots represent the 5, 10, 25, 50, 75, 90 and 95 quantiles.

In the evaluation period (right-hand-side diagram of Figure 5a), the underestimation of flow variance obviously persists since colored points are still centered around the red semi-circle, though with much more scatter, and bias substantially increases. Two-dimensional contour plots show that 90% of the population lies between  $R \in [0.55, 0.97]$  and  $\alpha \in [0.4, 1.2]$ .

Figure 5b shows the results of GRX model cal-

ibrated with KGE(Q) as objective function. In the calibration period (left-hand-side diagram of Figure 5b), GRX manages to get a small mean bias (red arrows are barely distinguished in the left-hand side diagram of Figure 5b, since they are very short), apart for some particular watersheds (again very arid watersheds). While calibrated with KGE(Q), the centered pattern clearly aligns with the unbiased



**Figure 5.** Taylor diagram representing variability bias and correlation (points), and mean bias (arrows in red) for (a) GRX model in calibration (left) and evaluation (right) while calibrated with NSE(Q) as objective function, and (b) GRX model in calibration (left) and evaluation (right), when calibrated with KGE(Q) as objective function. Contour plot illustrate the two-dimensional density of points (from 10% to 90% of the sample).

normalized standard deviation semi-circle (in red) with a very limited 2D scatter (90% of the population lies between  $R \in [0.6, 0.98]$  and  $\alpha \in [0.95, 1.05]$ ). In the evaluation period (right-hand-side diagram of Figure 5b), GRX has a limited normalized standard deviation bias since colored points are still centered around the red semi-circle in the  $(R, \alpha)$  space, though with much more scatter and bias substantially increases. Two-dimensional contour plots show that 90% of the population lies between

$R \in [0.6, 0.98]$  and  $\alpha \in [0.5, 1.5]$ .

Figures 6 and 7 present results for NSE(Q) and KGE(Q) as objective functions, for the GRX model and three climate clusters as defined in Mathevet *et al.* [2020]: (a) arid with desert and steppe (A), (b) temperate with warm summer (T+WS), and (c) temperate without dry season and warm summer (T-DS+WS). These figures show that the conclusions obtained with the full 2050 watershed sample are still valid with limited climate clusters (from



100 to 800 watersheds) and patterns of results in the modified Taylor diagram are mostly the same. The large variability of model performance of arid watersheds is illustrated in Figures 6a and 7a, while the homogeneity of model performance for temperate without dry season watersheds is illustrated on Figures 6c and 7c.

As previously shown in Figures 3 and 4, while calibrated with  $NSE(Q)$  or  $KGE(Q)$ , both models have a very similar behavior and level of performances in calibration and evaluation. This is shown by the similar pattern of the 2D density contour plot on Figure 5 (resp. Appendix A, Figure A1), independently from climate clusters as shown on Figures 6 and 7 (resp. Appendix A, Figures A2, A3). The comparison of model behavior clearly shows the significant impact of the choice of the objective function ( $NSE(Q)$  or  $KGE(Q)$ ), whatever the hydrological model (GRX or MRX, with very different level of complexity) and whatever watershed properties (large sample of 2050 watersheds worldwide). As stated previously [Gupta *et al.*, 2009, Mathevet *et al.*, 2020, among others],  $KGE(Q)$  leads to much more reliable performance than  $NSE(Q)$ .

Furthermore, the Taylor diagram represents a complementary tool to distributions (Figure 3) or boxplots (Figure 4), allowing a multi-objective representation of model performance behavior on a large sample of watersheds (which is always difficult to summarize) or climate clusters. Taylor diagram (Figures 5–7) allow to display dependency structure between correlation and variability bias while calibrated with  $NSE(Q)$ .

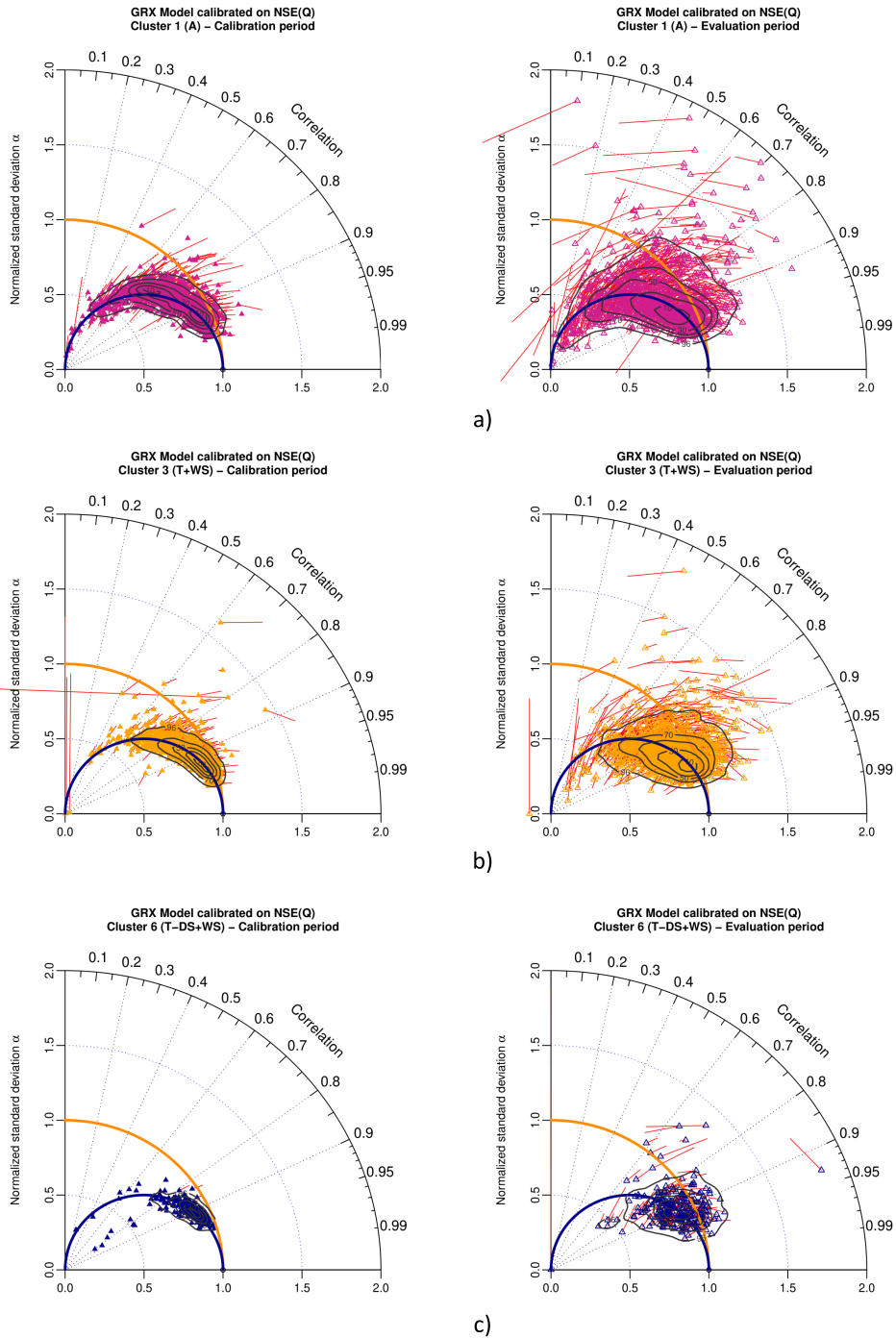
A complementary use of Taylor diagram is to explore the sub-period variability of performances of a model on a given case study, in comparison to performances estimated over the whole period. As previously explored by Mathevet *et al.* [2020], sub-period variability of performances could be high, with unbiased simulation in mean and variability at the full period scale, but with a totally different behavior at a sub-period scale. Figure 8 show the full period and sub-period (annual basis) performances for GRX model calibrated with  $KGE(Q)$  as objective function for Hérault (Y2102010), Glueyre (V4145210) and Blavet (J5202110) watersheds, during calibration and evaluation. These three watersheds come from the French sample (Banque Hydro stations) and were selected because they represent different types

of climate (Hérault: mediterranean climate with famous “Cévenols” events, Glueyre: mountainous climate, Blavet: oceanic climate) and geological context (Hérault: limestones, Glueyre: granites, metamorphic, volcanic rocks and Blavet: sandstones, metamorphic rocks). Hence, these three watersheds exhibit two different behavior, with a significant sub-period scatter of mean and variability bias for Hérault and Glueyre, while Blavet keep a rather limited and homogeneous mean and variability bias all along time. From calibration to evaluation, sub-period performance patterns exhibited on Taylor diagrams remain mostly the same, with a slight increase of the spread. Taylor diagrams clearly illustrate that even if the overall bias in the calibration or evaluation periods can be quite small, the bias on sub-period intervals can be significant (up to divided/multiplied by 2, on Hérault and Glueyre). These results shows that the use of full-period aggregate performance metrics may not provide sufficient discrimination to properly assess model behavior on sub-periods [Mathevet *et al.*, 2020]. Similar results for MRX are shown in Appendix A, Figure A4.

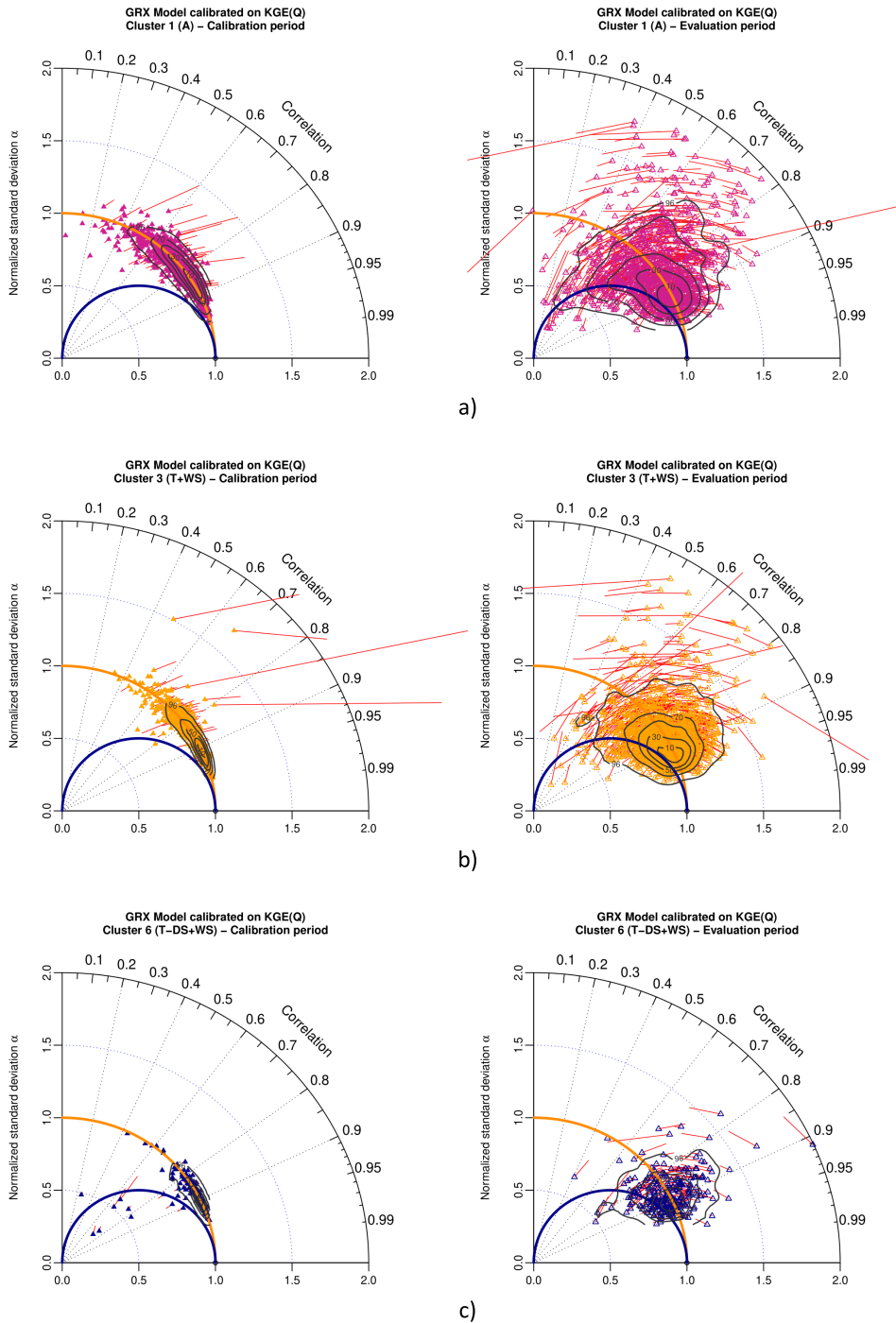
#### 4.3. *Impact of square root transformation on multi-objective assessment of model performances*

We now present a numerical experiment performed on a limited sample of 635 watersheds (randomly selected). GRX and MRX models were calibrated using a classical mathematical transformation, i.e. square root (RQ) applied to streamflow, using  $NSE(Q \& RQ)$  and  $KGE(Q \& RQ)$  as objective functions. This transformation is classically used to better represent a larger range of streamflows (i.e. low-flows) and reduce the weight of high-flows during calibration [Pushpalatha *et al.*, 2012]. Another transformation has been tested (log transformation), but results are not shown due to documented numerical issues of this transformation [Santos *et al.*, 2018].

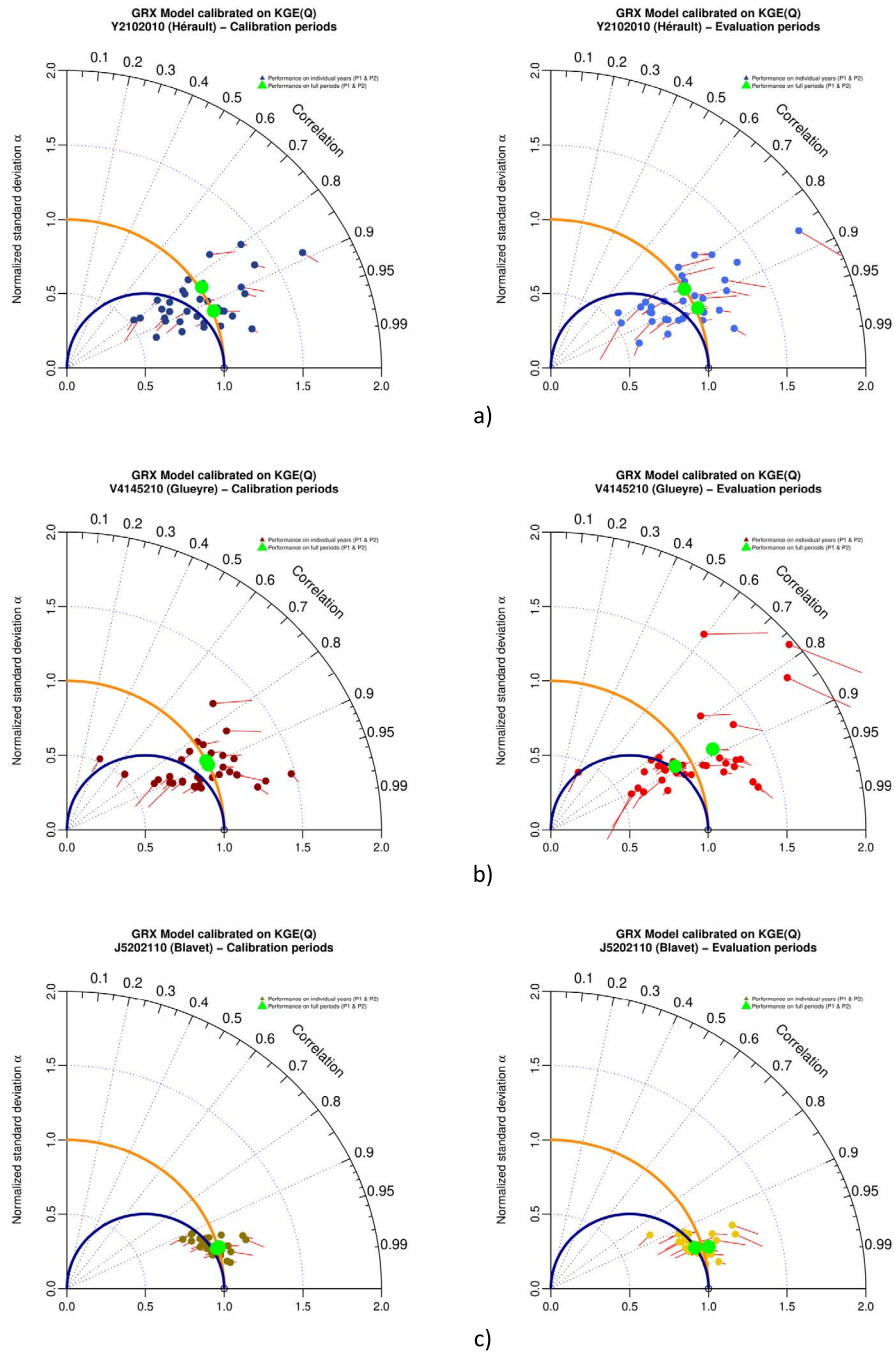
Figure 9 presents Taylor diagrams with GRX model performance in calibration and evaluation for  $NSE(Q)$  (a) and  $NSE(RQ)$  (b). Figure 10 presents Taylor diagrams with GRX model performance in calibration and evaluation for  $KGE(Q)$  (a) and  $KGE(RQ)$  (b). Similar results for MRX are shown in Appendix A, Figures A5 and A6.



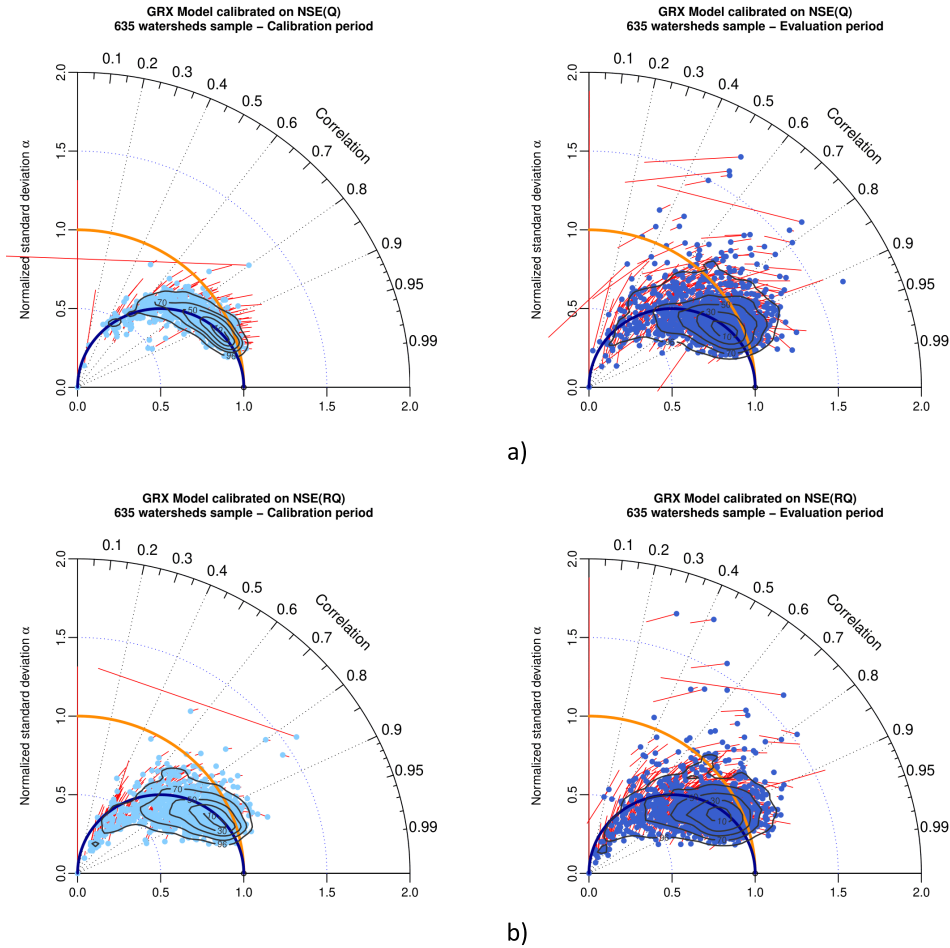
**Figure 6.** Taylor diagram representing variability bias and correlation (points) and mean bias (red arrows) for GRX model in calibration (left) and evaluation (right) for (a) arid with desert and steppe cluster (586 watersheds), (b) temperate with warm summer cluster (785 watersheds), and (c) temperate without dry season and warm summer (125 watersheds), when calibrated with NSE(Q) as objective function. Contour plots illustrate the two-dimensional density of points (from 10% to 90% of the sample).



**Figure 7.** Taylor diagram representing variability bias and correlation (points) and mean bias (red arrows) for GRX model in calibration (left) and evaluation (right) for (a) arid with desert and steppe cluster (586 watersheds), (b) temperate with warm summer cluster (785 watersheds) and (c) temperate without dry season and warm summer (125 watersheds), when calibrated with KGE(Q) as objective function. Contour plots illustrate the two-dimensional density of points (from 10% to 90% of the sample).



**Figure 8.** Taylor diagram representing variability bias and correlation (points) and mean bias (red arrows) for GRX model in calibration (left) and evaluation (right) for (a) Hérault watershed, (b) Glueyre watershed and (c) Blavet watershed, when calibrated with KGE(Q) as objective function. Points either represent performances on individual years (color) or on the full periods (green color). These three watersheds had been selected because generations of Ghislain de Marsily and Pierre Hubert students (DEA National d'Hydrologie) swan, practiced kayak or celebrate New socio-hydrological years (NASH) in their beautiful waters.

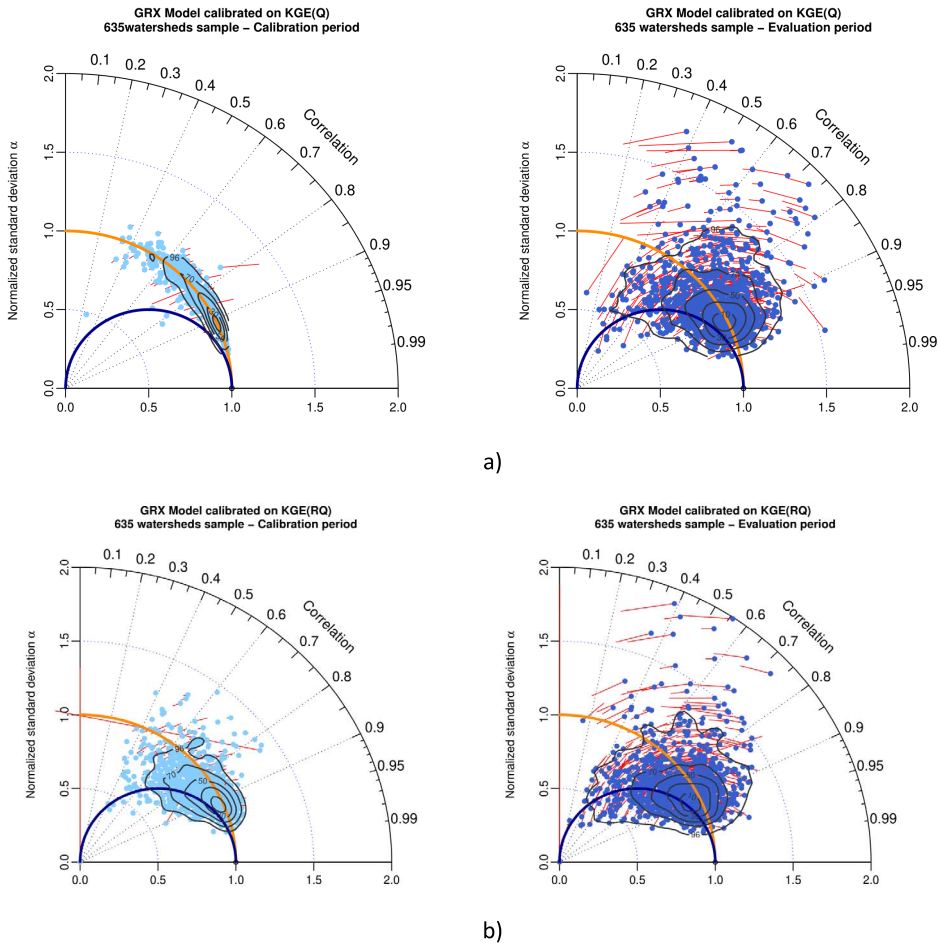


**Figure 9.** Taylor diagram representing variability bias and correlation (points) and mean bias (red arrows) for GRX model in calibration (left) and evaluation (right), when calibrated with (a) NSE(Q) and (b) NSE(RQ) as objective function. The contour plot illustrates the two-dimensional density of points (from 10% to 90% of the sample).

When using the NSE metric as objective function for parameter optimization, in calibration, Figure 9b left (square root transformation) compared to Figure 9a left (natural streamflow) shows that this classical numerical transformation of streamflow increases the normalized standard deviation bias, with a majority of points located slightly to significantly under the blue semi-circle line with an increasing spread, for square root transformation of streamflow. Two-dimension contour plots surface of 90% of the population increase from NSE(Q) to NSE(RQ). In evaluation, Figure 9b right (square root transformation), compared to Figure 9a right (natural streamflow) shows that the pattern found in calibration re-

main the same concerning the normalized standard deviation bias, with an increase of the spread of the results (as shown by the two-dimensional contour plots surface of 90% of the population).

When using the KGE metric as objective function for parameter optimization, in calibration, Figure 10b left (square root transformation) compared to Figure 10a left (natural streamflow) shows that this classical numerical transformation of streamflow increases the normalized standard deviation bias, with a majority of points located between the red semi-circle line and the blue semi-circle line for square root transformation. The spread of the results increases slightly for square root transformation. Two-



**Figure 10.** Taylor diagram representing variability bias and correlation (points) and mean bias (red arrows) for GRX model in calibration (left) and evaluation (right), while calibrated with (a) KGE(Q) and (b) KGE(RQ) as objective function. The contour plot illustrates the two-dimensional density of points (from 10% to 90% of the sample).

dimension contour plots surface of 90% of the population increases from KGE(Q) to KGE(RQ). In evaluation, Figure 10b right (square root transformation) compared to Figure 10a right (natural streamflow) shows that the pattern found in calibration remain the same concerning the normalized standard deviation bias, with an increase of the spread of the results (as shown by the two-dimensional contour plots surface of 90% of the population).

Comparison of Figure 9a (calibration with NSE(Q)), Figure 10a (calibration on KGE(Q)) and Figure 10b (calibration on KGE(RQ)) allows to compare patterns of results (in calibration and evalu-

ation) for different objective functions. Concerning the variability bias, this comparison shows that pattern of calibration with KGE(RQ) lies in-between patterns of KGE(Q) (no bias) and NSE(Q) ( $\alpha = R$ ).

## 5. Discussion

Based on a worldwide sample of 2050 watersheds, two conceptual rainfall-runoff models (GRX and MRX), two metrics classically used in hydrology

(Nash–Sutcliffe efficiency and Kling–Gupta efficiency) and a split sample test, this paper investigated the usefulness of a modified Taylor diagram to analyze the results of large sample experiments.

We first proposed a modified Taylor diagram to make clear the theoretical reasons why OLS schemes tend to bias model simulations towards an underestimated standard deviation of streamflows. Empirical evidence of this behavior was provided in an intensive numerical experiment over a large watershed sample. NSE(Q) is still largely used by hydrological modelers and we encourage the use of KGE(Q) instead of NSE(Q). The clear advantage of KGE(Q) instead of NSE(Q) is that it deals with the optimization process involving several objective functions (correlation, ratio of means, and ratio of standard deviations) independently while NSE(Q) implicitly involved a prescribed trade-off between correlation and ratio of standard deviation. As a consequence, we showed that calibrating hydrological models with NSE(Q) led to significant variability biases on watersheds where models have difficulties representing the temporal behavior of the watershed. KGE(Q) tended to reduce long-term mean and variability bias, independently from the models' level of correlation. These problems related to NSE (and more generally to OLS schemes) were already reported in previous studies [see e.g. Gupta et al., 2009] but the illustration provided by the modified Taylor diagram is a valuable add-on to these studies. Besides these numerical problems, NSE does put more weight on high flows, in the sense that high flow days have a major contribution to the total squared error. As with all least-squares-based schemes, using NSE implicitly considers uncorrelated error time series and homoscedastic errors [see e.g. Kavetski et al., 2003] and thus does not use the whole information concerning the temporal structure of the observed discharge time series. KGE allows to partially cope with this problem by including more explicitly the errors in variance. Other smarter error models (e.g. weighted least square accounting for heteroscedasticity or correlation in the residuals) may be considered and visualizing their optimization results under the proposed Taylor diagram may help to understand how the trade-off between correlation, the ratio of means, and the ratio of standard deviation are modified compared to NSE(Q) or KGE(Q).

Interestingly, we showed that the biased optimization of OLS schemes towards reduced streamflow variability was independent on the model used and the environmental settings of the watersheds. When modeling a single watershed, modified Taylor diagrams can be represented for all hydrological years individually since the use of full-period aggregate performance metrics may not provide sufficient discrimination to properly assess model behavior on sub-periods. Figure 8 showed that the optimization process may be different according to the studied watershed in the way it deals with annual biases and compensation along the calibration period. Thus, we believe that the modified Taylor diagram can be an interesting tool to investigate the ability of the model to reproduce individual years, in complement to other existing graphical tools [see e.g. Coron et al., 2015]. This diagnosis allows the modelers to figure out the ability of the model for extrapolation.

Last, we showed that calibrating hydrological models with NSE metric applied to square root of streamflows increases the initial pitfalls of using the NSE metric on natural streamflow. This transformation increases the normalized standard deviation bias of simulations while increasing the spread of the results in the  $(R, \alpha)$  space. It is also shown that calibrating hydrological models with KGE metric applied to the square root of streamflow reduces the benefits of using the KGE metric on streamflow. Again, this transformation increases the normalized standard deviation bias of simulations, while increasing the spread of the results in the  $(R, \alpha)$  space. As stated before [Santos et al., 2018], we consider that using KGE metric on square root has significant drawbacks and should be avoided since the main drawbacks of the NSE metric are introduced again.

## 6. Conclusion

We proposed in this paper a novel diagnostic tool for hydrological model simulations. The proposed graphical tool is based on the Taylor diagram and we adapted this diagram to visualize commonly used assessment criteria in hydrological modeling.

The adapted Taylor diagram illustrates some well-known drawbacks of OLS schemes widely used in the scientific community and particularly the fact that the standard deviation of the simulated streamflow

is systematically underestimated. While the use of alternative objective functions such as the KGE(Q) overcomes these problems, future works may investigate the outputs of hydrological models optimized with other objective functions that consider heteroscedasticity and/or based on flow signatures. Representing the outputs of these more complex objective functions into the modified Taylor diagram may help to understand how the calibration process deals with the trade-off between correlation, the ratio of standard deviation, and the ratio of means.

### Conflicts of interest

Authors have no conflict of interest to declare.

### Acknowledgements

Authors would like to thank Charles Perrin for his valuable support and constructive comments on the manuscript. Authors would like to warmly acknowledge Editors for their invitation to contribute to this special issue and the three anonymous reviewers for their very constructive remarks.

Authors acknowledge: Météo France, SCHAPI-Banque Hydro, EDF, Laurent Coron, Nicolas Le Moine and Audrey Valery for the French data sets, Jai Vaze and Francis Chiew for the Australian data sets (CSIRO), John Schaake and Qingyun Duan for the American (MOPEX) data set, Audrey Valery for the Swiss (Météo Suisse and OFEV) and Swedish (SMHI) data sets, Berit Arheimer for the Swedish (SMHI) data sets, and Barry Croke and Ian Littlewood for the English data sets (TDMWG).

Codes used to produce Taylor diagrams of this paper will be made available in the airGR package [Coron *et al.*, 2017, 2022].

### Appendix A.

Appendix A shows the results for MRX model of numerical experiments:

- Figure A1 (corresponding to Figure 11 for GRX): Taylor diagram for calibration with NSE(Q) and MSE(Q);

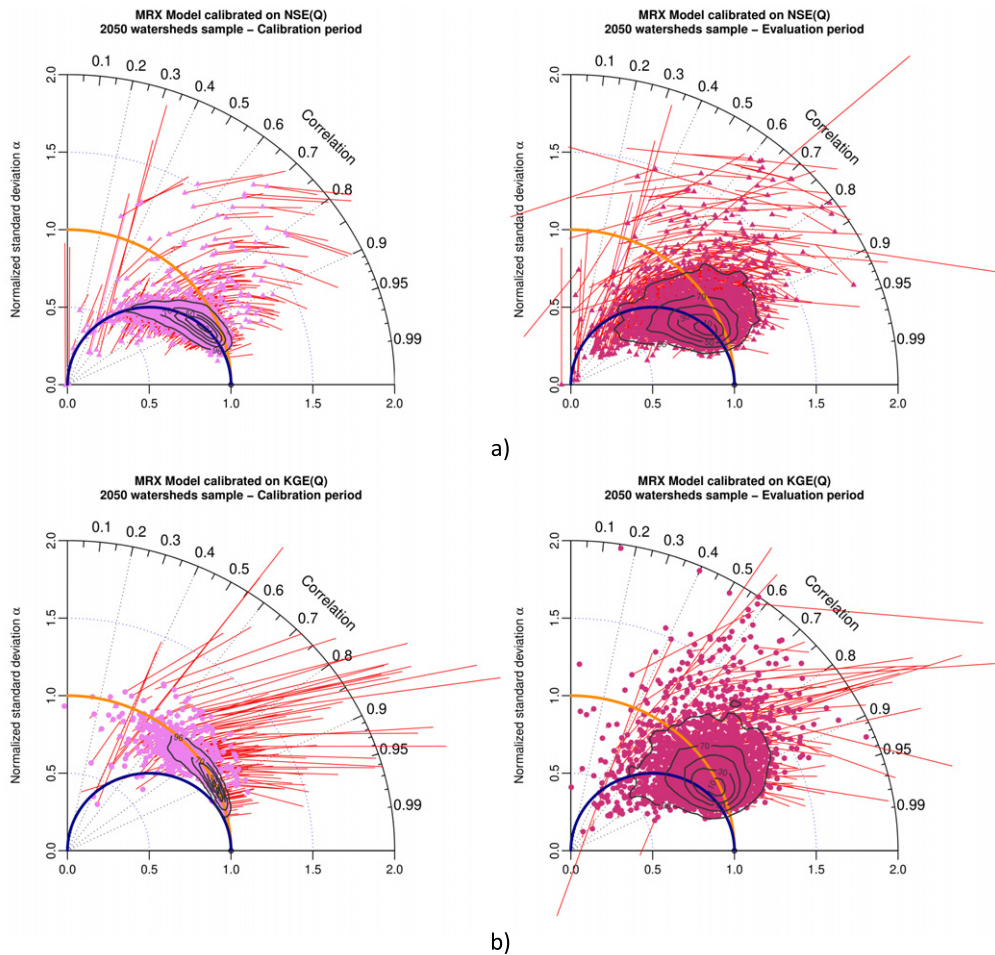
- Figure A2 (corresponding to Figure 12 for GRX): Taylor diagram for calibration with NSE(Q) for 3 different climate clusters;
- Figure A3 (corresponding to Figure 13 for GRX): Taylor diagram for calibration with KGE(Q) for 3 different climate clusters;
- Figure A4 (corresponding to Figure 14 for GRX): Taylor diagram for calibration with KGE(Q) for 3 different watersheds in France (Hérault, Glueyre and Blavet);
- Figure A5 (corresponding to Figure 15 for GRX): Taylor diagram for calibration with NSE(Q) and NSE(RQ);
- Figure A6 (corresponding to Figure 16 for GRX): Taylor diagram for calibration with KGE(Q) and KGE(RQ).

These figures show that numerical experiments with MRX model lead to similar results and conclusions to those obtained with GRX model (patterns on Taylor diagrams are very similar). These analyses with GRX and MRX models confirm the generality of our results, independent from the rainfall-runoff model structure.

### References

- Addor, N., Do, H. X., Alvarez-Garreton, C., Coxon, G., Fowler, K., and Mendoza, P. A. (2020). Large-sample hydrology: recent progress, guidelines for new datasets and grand challenges. *Hydrol. Sci. J.*, 65(5), 712–725.
- Andréassian, V. (2023). On the (im)possible validation of hydrogeological models. *C. R. Géosci.*, 355(S1). Online first.
- Coron, L., Andréassian, V., Perrin, C., and Le Moine, N. (2015). Graphical tools based on Turc–Budyko plots to detect changes in catchment behaviour. *Hydrol. Sci. J.*, 60, 1394–1407.
- Coron, L., Delaigue, O., Thirel, G., Dorchies, D., Perrin, C., and Michel, C. (2022). airGR: suite of GR hydrological models for precipitation-runoff modelling. <https://CRAN.R-project.org/package=airGR>. R package version 1.7.0.
- Coron, L., Thirel, G., Delaigue, O., Perrin, C., and Andréassian, V. (2017). The suite of lumped GR hydrological models in an R package. *Environ. Modell. Softw.*, 94, 166–171.





**Figure A1.** Taylor diagram representing variability bias and correlation (points) and mean bias (arrows in red) for (a) MRX model in calibration (left) and evaluation (right) while calibrated with NSE(Q) as objective function, and (b) MRX model in calibration (left) and evaluation (right), when calibrated with KGE(Q) as objective function. Contour plots illustrate the two-dimensional density of points (from 10% to 90% of the sample).

de Marsily, G. (1994). Quelques réflexions sur l'utilisation des modèles en hydrologie. [Tribune libre]. *Revue des sciences de l'eau / J. Water Sci.*, 7(3), 219–234.

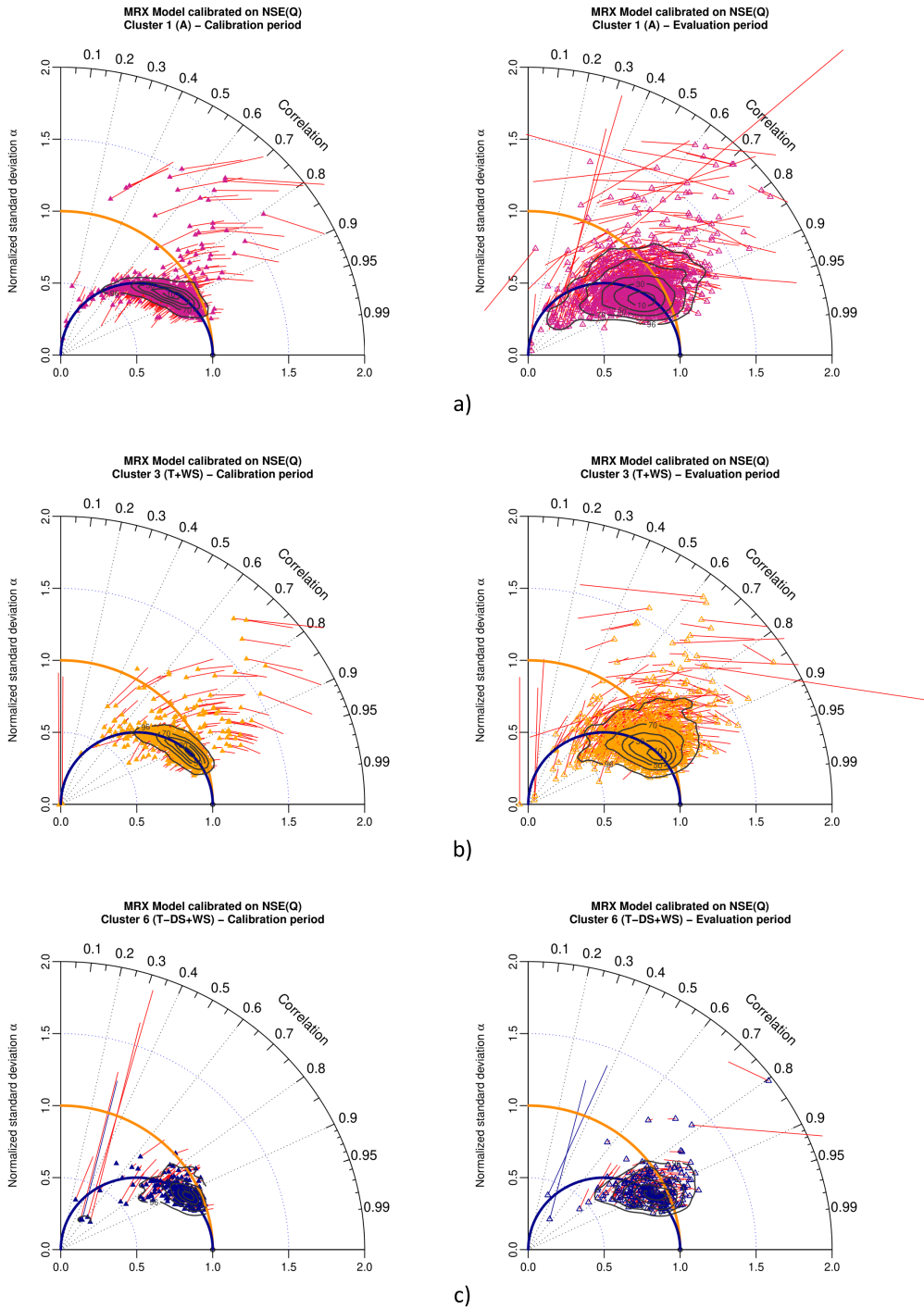
Garavaglia, F., Le Lay, M., Gottardi, F., Garçon, R., Gailhard, J., Paquet, E., and Mathevet, T. (2017). Impact of model structure on flow simulation and hydrological realism: from a lumped to a semi-distributed approach. *Hydrol. Earth Syst. Sci.*, 21, 3937–3952.

Garçon, R. (1996). Prévision opérationnelle des ap-

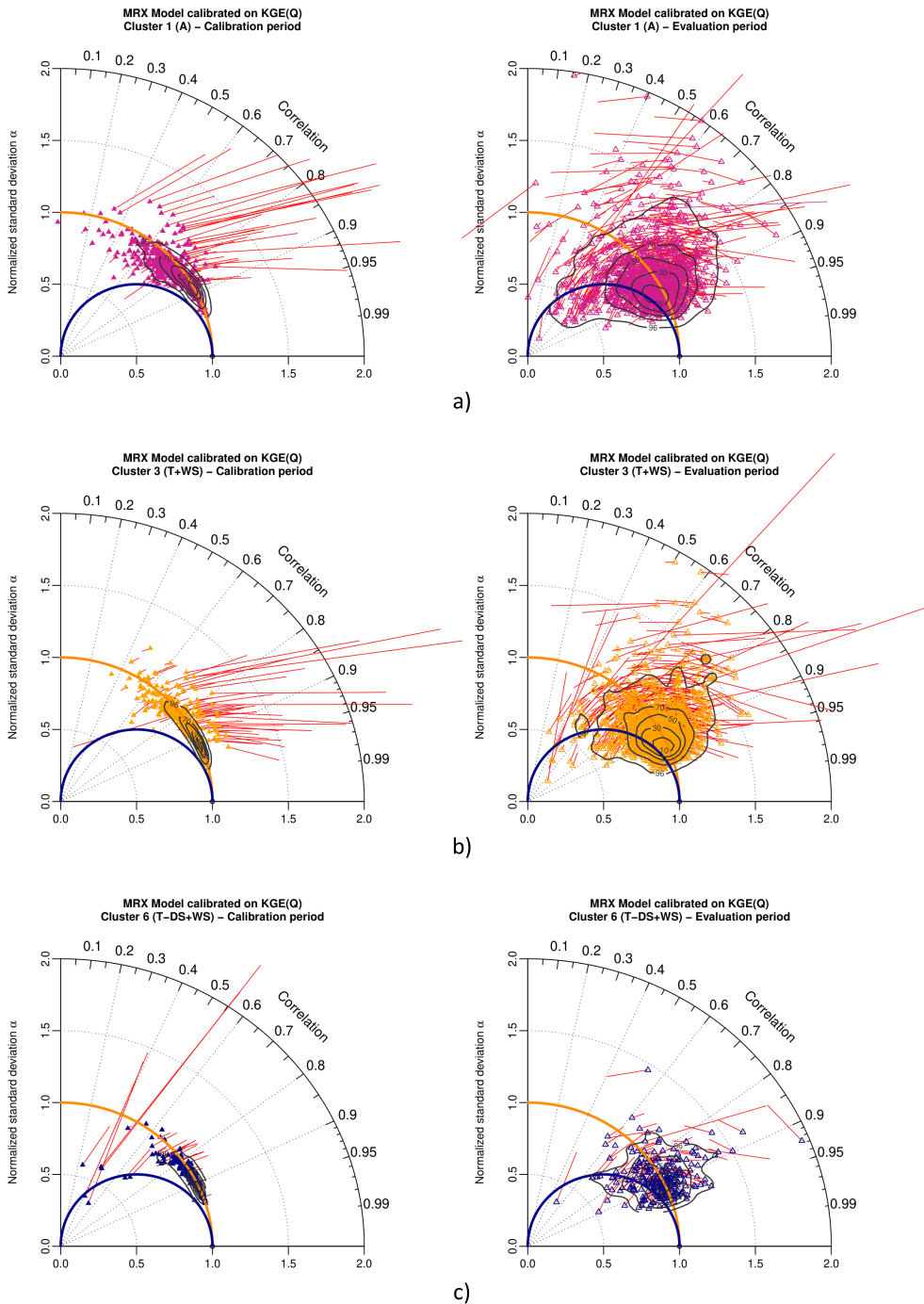
ports de la Durance à Serre-Ponçon à l'aide du modèle MORDOR. Bilan de l'année 1994–1995. *Houille Blanche*, 5, 71–76.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez-Baquero, G. F. (2009). Decomposition of the mean squared error & NSE performance criteria: implications for improving hydrological modelling. *J. Hydrol.*, 377, 80–91.

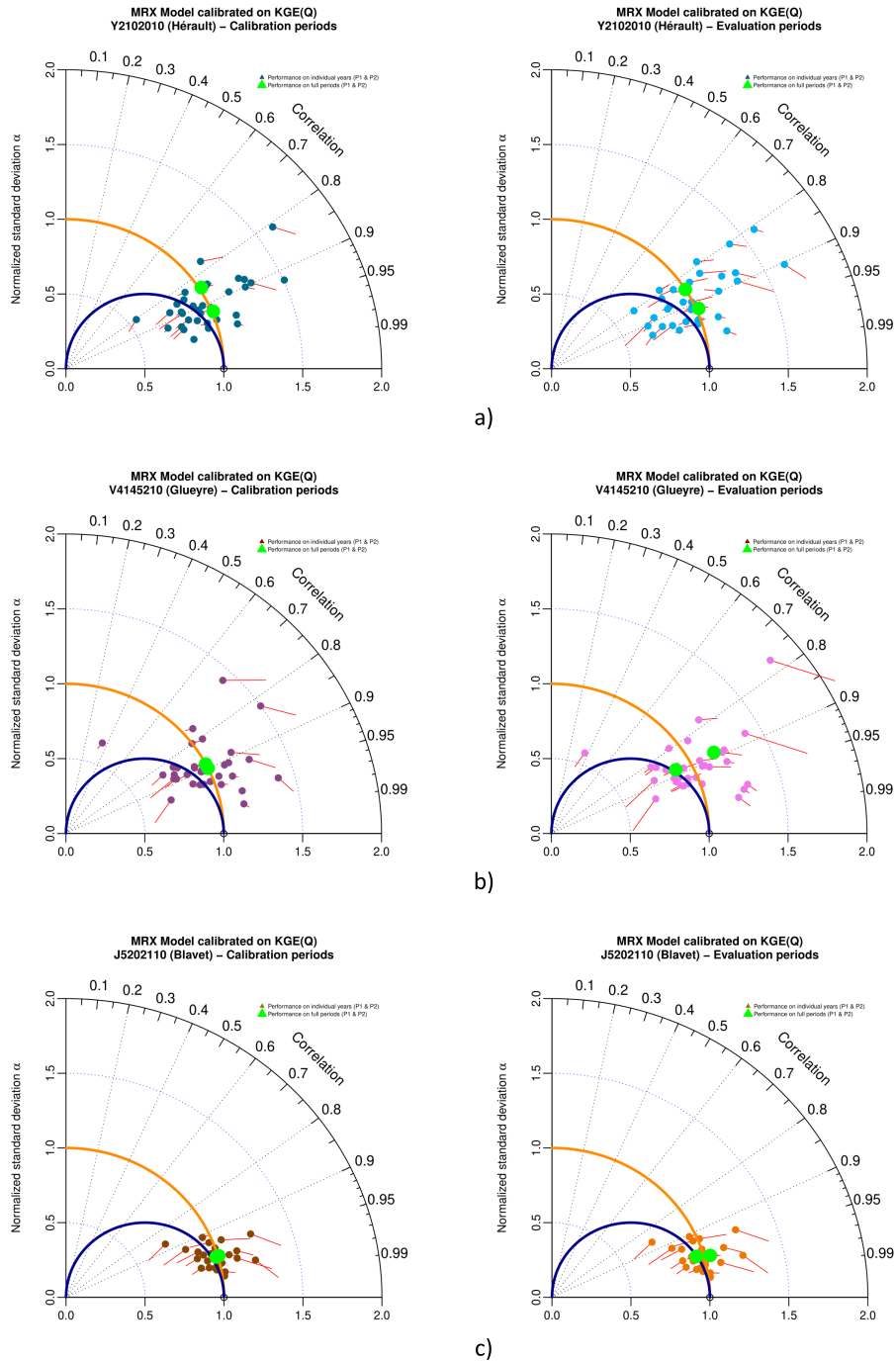
Gupta, H. V., Perrin, C., Blöschl, G., Montanari, A., Kumar, R., Clark, M., and Andréassian, V. (2014). Large-sample hydrology: a need to balance depth



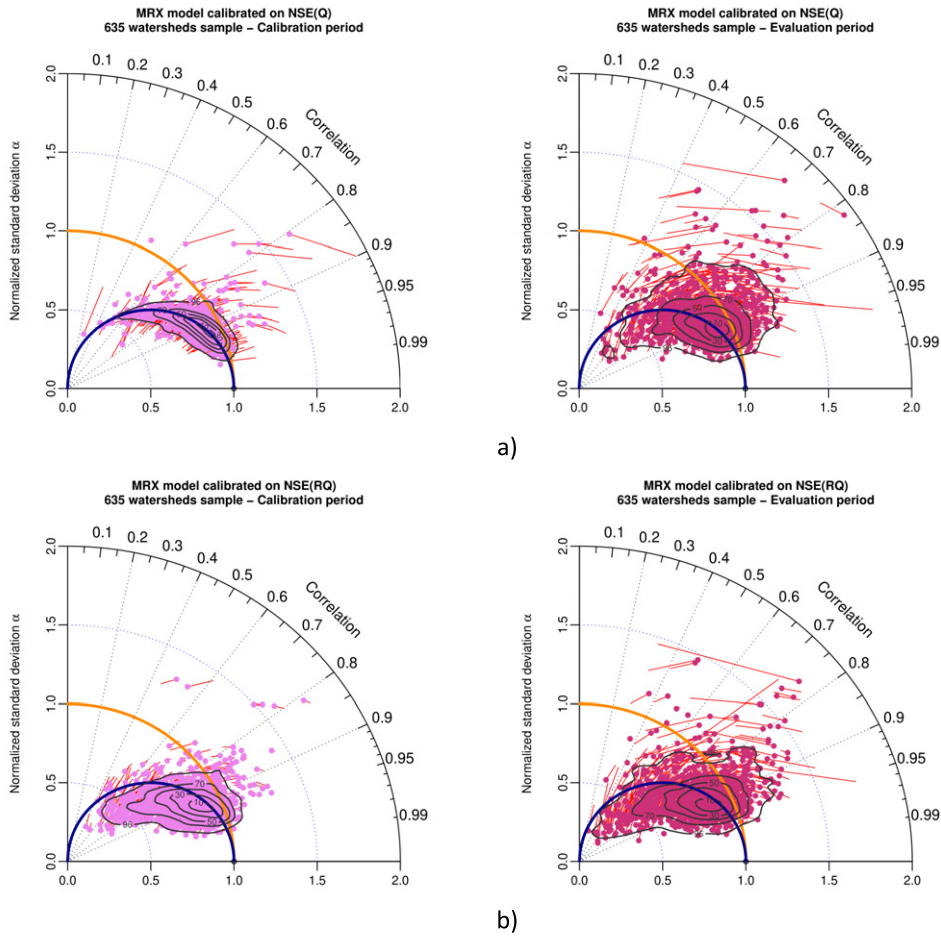
**Figure A2.** Taylor diagram representing variability bias and correlation (points) and mean bias (red arrows) for MRX model in calibration (left) and evaluation (right) for (a) arid with desert and steppe cluster (586 watersheds), (b) temperate with warm summer cluster (785 watersheds) and (c) temperate without dry season and warm summer (125 watersheds), when calibrated with NSE(Q) as objective function. Contour plots illustrate the two-dimensional density of points (from 10% to 90% of the sample).



**Figure A3.** Taylor diagram representing variability bias and correlation (points) and mean bias (red arrows) for MRX model in calibration (left) and evaluation (right) for (a) arid with desert and steppe cluster (586 watersheds), (b) temperate with warm summer cluster (785 watersheds) and (c) temperate without dry season and warm summer (125 watersheds), when calibrated with KGE(Q) as objective function. Contour plots illustrate the two-dimensional density of points (from 10% to 90% of the sample).



**Figure A4.** Taylor diagram representing KGE mean bias, variability bias and correlation for MRX model in calibration (left) and evaluation (right) for (a) Hérault watershed, (b) Glueyre watershed and (c) Blavet watershed, while calibrated with KGE(Q) as objective function. Points either represent performances on individual years (color) or on the full periods (green color). These three watersheds had been selected because generations of Ghislain de Marsily and Pierre Hubert students (DEA National d'Hydrologie) swan, practiced kayak or celebrate New socio-hydrological years (NASH) in their beautiful waters.



**Figure A5.** Taylor diagram representing variability bias and correlation (points) and mean bias (red arrows) for MRX model in calibration (left) and evaluation (right), when calibrated with (a) NSE(Q) and (b) NSE(RQ) as objective function. The contour plot illustrates the two-dimensional density of points (from 10% to 90% of the sample).

with breadth. *Hydrol. Earth Syst. Sci.*, 18, 463–477.

Kavetski, D., Franks, S. W., and Kuczera, G. (2003). Confronting input uncertainty in environmental modelling. In Duan, Q., Gupta, V. H., Sorooshian, S., Rousseau, N. A., and Turcotte, R., editors, *Calibration of Watershed Models*. American Geophysical Union, Washington, DC.

Klemeš, V. (1986). Operational testing of hydrological simulation models. *Hydrol. Sci. J.*, 31(1), 13–24.

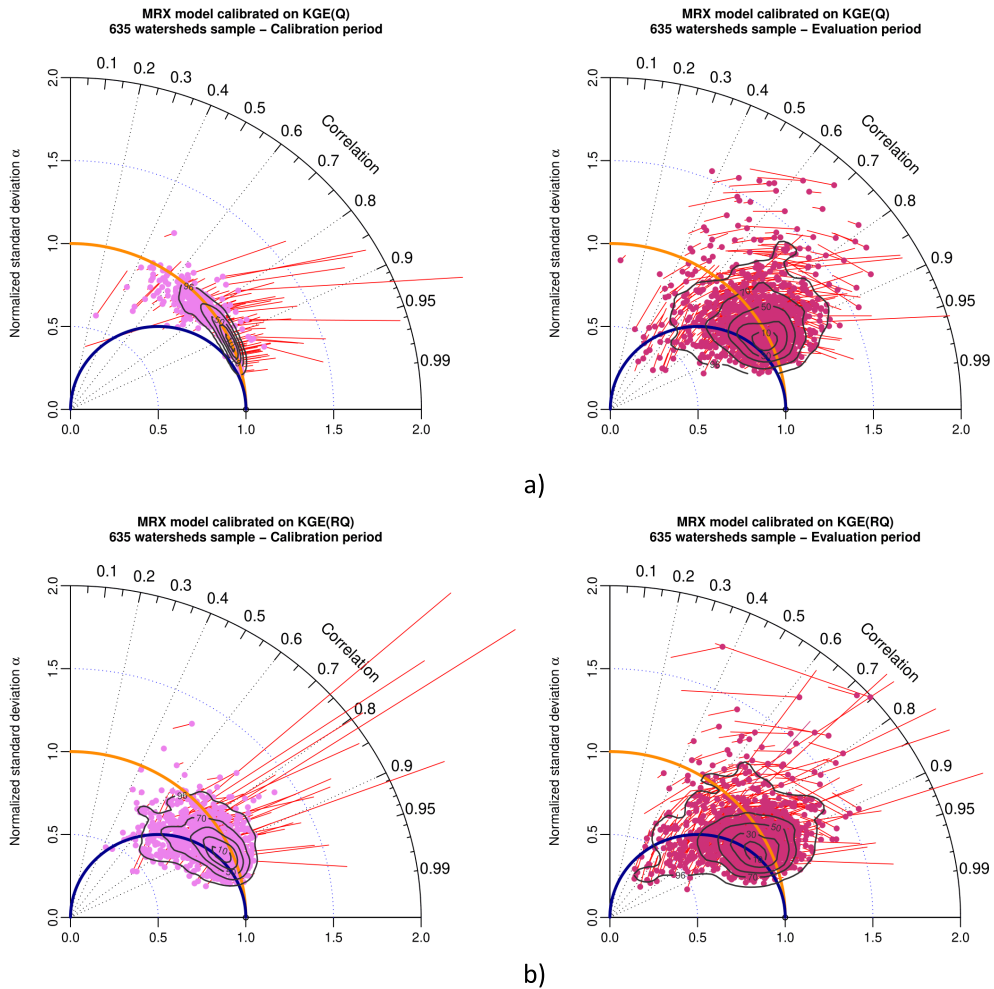
Konikow, L. F. and Bredehoeft, J. D. (1992). Groundwater models cannot be validated. *Adv. Water Resour.*, 15, 75–83.

Lane, R. A., Coxon, G., Freer, J. E., Wagener, T., Johnes,

P. J., Bloomfield, J. P., Greene, S., Macleod, C. J. A., and Reaney, S. M. (2019). Benchmarking the predictive capability of hydrological models for river flow and flood peak predictions across over 1000 catchments in Great Britain. *Hydrol. Earth Syst. Sci.*, 23, 4011–4032.

Le Moine, N. (2008). *Le bassin versant de surface vu par le souterrain : une voie d'amélioration des performances et du réalisme des modèles pluie-débit ?* Thèse de doctorat, Université Pierre et Marie Curie (Paris), Cemagref (Antony). 324 pages.

Maroufpoor, S., Bozorg-haddad, O., and Maroufpoor, E. (2020). Reference evapotranspiration estimating



**Figure A6.** Taylor diagram representing variability bias and correlation (points) and mean bias (red arrows) for MRX model in calibration (left) and evaluation (right), when calibrated with (a) KGE(Q) and (b) KGE(RQ) as objective function. The contour plot illustrates the two-dimensional density of points (from 10% to 90% of the sample).

based on optimal input combination and hybrid artificial intelligent model: hybridization of artificial neural network with grey wolf optimizer algorithm. *J. Hydrol.*, 588, article no. 125060.

Mathevet, T. (2005). *Quels modèles pluie-débit globaux pour le pas de temps horaire ? Développement empirique et comparaison de modèles sur un large échantillon de bassins versants*. Thèse de doctorat, ENGREF (Paris), Cemagref (Antony), France. 463 pages.

Mathevet, T., Gupta, H., Perrin, C., Andréassian, V., and Le Moine, N. (2020). Assessing the perfor-

mance and robustness of two conceptual rainfall-runoff models on a worldwide sample of watersheds. *J. Hydrol.*, 585, article no. 124698.

Mathevet, T., Michel, C., Andréassian, V., and Perrin, C. (2006). *A Bounded Version of the Nash-Sutcliffe Criterion for Better Model Assessment on Large Sets of Basins*. IAHS Red Books Series No 307. IAHS, Wallingford, Oxfordshire, UK.

McMillan, H. K., Booker, D. J., and Cattoën, C. (2016). Validation of a national hydrological model. *J. Hydrol.*, 541, 800–815.

Nash, J. and Sutcliffe, J. (1970). River flow forecasting

- through conceptual models part I. A discussion of principles. *J. Hydrol.*, 10, 282–290.
- Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., Viger, R. J., Blodgett, D., Brekke, L., Arnold, J. R., Hopson, T., and Duan, Q. (2015). Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: data set characteristics and assessment of regional variability in hydrologic model performance. *Hydrol. Earth Syst. Sci.*, 19, 209–223.
- Oudin, L., Hervieu, E., Michel, C., Perrin, C., Andréassian, V., Anctil, E., and Loumagne, C. (2005). Which potential evapotranspiration input for a lumped rainfall-runoff model?: Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall-runoff modelling. *J. Hydrol.*, 303, 290–306.
- Pushpalatha, R., Perrin, C., Le Moine, N., and Andréassian, V. (2012). A review of efficiency criteria suitable for evaluating low-flow simulations. *J. Hydrol.*, 420–421, 171–182.
- Pushpalatha, R., Perrin, C., Le Moine, N., Mathevet, T., and Andréassian, V. (2011). A downward structural sensitivity analysis of hydrological models to improve low-flow simulation. *J. Hydrol.*, 411(1–2), 66–76.
- Santos, L., Thirel, G., and Perrin, C. (2018). Technical note: pitfalls in using log-transformed flows within the KGE criterion. *Hydrol. Earth Syst. Sci.*, 22, 4583–4591.
- Schaefli, B. and Gupta, H. V. (2007). Do Nash values have value? *Hydrol. Process.*, 21, 2075–2080.
- Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, 106(D7), 7183–7192.
- Yaseen, Z. M., Awadh, S. M., Sharafati, A., and Shahid, S. (2018). Complementary data-intelligence model for river flow simulation. *J. Hydrol.*, 567, 180–190.







Research article

Geo-hydrological Data & Models

# Assessing water and energy fluxes in a regional hydrosystem: case study of the Seine basin

Deniz Kilic<sup>\*,a</sup>, Agnès Rivière<sup>\*,a</sup>, Nicolas Gallois<sup>\*,a</sup>, Agnès Ducharne<sup>\*,b</sup>, Shuaitao Wang<sup>\*,b</sup>, Philippe Peylin<sup>\*,c</sup> and Nicolas Flipo<sup>\*,a</sup>

<sup>a</sup> Mines Paris, PSL University, Centre for Geosciences and Geoengineering, Fontainebleau, France

<sup>b</sup> UMR 7619 METIS, Sorbonne Université, CNRS, EPHE, Paris, France

<sup>c</sup> Laboratoire des Sciences du Climat et de l'Environnement (LSCE), Gif-sur-Yvette, France

*E-mails:* deniz.kilic@minesparis.psl.eu (D. Kilic), agnes.riviere@minesparis.psl.eu (A. Rivière), nicolas.gallois@minesparis.psl.eu (N. Gallois), agnes.ducharne@upmc.fr (A. Ducharne), shuaitao.wang@upmc.fr (S. Wang), peylin@lsce.ipsl.fr (P. Peylin), nicolas.flipo@minesparis.psl.eu (N. Flipo)

**Abstract.** While it is well accepted that climate change and growing water needs affect long-term sustainable water resources management, performing accurate simulations of water cycle and energy balance dynamics at regional scale remains a challenging task.

Traditional Soil-Vegetation-Atmosphere-Transfer (SVAT) models are used for numerical surface water and energy simulations. These models, by conception, do not account for the groundwater lower boundary that permits a full hydrosystem representation. Conversely, while addressing important features such as subsurface heterogeneity and river-aquifer exchanges, groundwater models often integrate overly simplified upper boundary conditions ignoring soil heating and the impacts of vegetation processes on radiation fluxes and root-zone uptakes. In this paper, one of the first attempts to jointly model water and energy fluxes with a special focus on both surface and groundwater at the regional scale is proposed on the Seine hydrosystem (78,650 km<sup>2</sup>), which overlays one of the main multi-aquifer systems of Europe.

This study couples the SVAT model ORCHIDEE and the process-based hydrological-hydrogeological model CaWaQS, which describes water fluxes, via a one-way coupling approach from ORCHIDEE toward CaWaQS based on the blueprint published by de Marsily et al. [1978]. An original transport library based on the resolution of the diffusion/advection transport equation was developed in order to simulate heat transfer in both 1D-river networks and pseudo-3D aquifer systems. In addition, an analytical solution is used to simulate heat transport through aquitards and streambeds. Simulated ORCHIDEE surface water and energy fluxes feed fast surface runoff and slow recharge respectively and then is used as CaWaQS forcings to compute river discharges, hydraulic heads and temperature dynamics through space and time, within each of the hydrosystem compartments. The tool makes it possible to establish a fully consistent water and energy budget over a period of 17 years. It also simulates temperature evolution in each aquifer and evaluates that river thermal regulation mostly relies by order of importance on short wave radiations (109.3 W·m<sup>-2</sup>), groundwater fluxes (48.1 W·m<sup>-2</sup>) and surface runoff (22.7 W·m<sup>-2</sup>).

\* Corresponding authors.

**Keywords.** Water and energy fluxes, Numerical simulation, Seine hydrosystem, Regional scale, Groundwater, River, SVAT.

*Manuscript received 28 June 2022, revised 14 September 2022, accepted 15 September 2022.*

## 1. Introduction

Nowadays, “*the water we use comes to 95% from the renewable water cycle*” [de Marsily, 2020, 2009]. Hydrosystems, defined as the overlapping combination of atmosphere, soil surface and subsurface, where water flows, host the water cycle and are an essential resource for humans and ecosystems. Hydrosystems sustain the global ecology and meet the societal needs of drinking water, energy and food production. In terms of water resources, providing sufficient quantities for all requirements is not the only challenge. In order to be considered as a resource, the water quality, which encompasses physical, chemical and biological properties, is also essential. Our current understanding of hydrosystems is based primarily on numerical models [de Marsily, 1986] in which physical processes are simulated under a range of historical or potential conditions to quantify storage as well as fluxes within a given compartment or between multiple compartments of the system.

Climate change will modify the water cycle and energy balance processes [Taylor et al., 2013]: soil infiltration, groundwater recharge [York et al., 2002], river discharge [Gudmundsson et al., 2021] and water temperature [Webb and Nobilis, 2007, Michel et al., 2022]. Moreover, rising temperatures increase evaporative demand over land [Berg et al., 2016], which limits the amount of water for aquifer system recharge. Groundwater withdrawals, river uses and indirect effects of irrigation and land use changes [Wagner et al., 2010] also increase stresses on hydrosystems. Despite the importance of quantifying hydrosystem fluxes for water resources management, little is known about the prevalent temperature regime and heat fluxes of hydrosystems at the regional scale in the light of climate change [Taylor et al., 2013].

Many studies have been carried out to determine the effects of climate warming on stream temperature [Moatar and Gailhard, 2006, Ducharme, 2008, Webb et al., 2008, van Vliet et al., 2013, Michel et al., 2020, Seyedhashemi et al., 2022] and aquatic species [Isaak et al., 2020]. Groundwater temperature is influenced by the temperature of the infiltrating wa-

ter and by the conduction of heat from the surface domain. Seasonal variations in shallow aquifers below urban areas coincide with various anthropogenic heat sources, such as sealed surfaces or subsurface building infrastructures for instance [Böttcher and Zosseder, 2022]. This results in highly dynamic and complex thermal conditions, and the identification of the influencing heat sources is not straightforward [Ferguson and Woodbury, 2007]. From the land surface downward, the amplitude of groundwater temperatures decreases and average temperatures tend towards the yearly average ground surface temperature [Bense and Kooi, 2004]. In areas with strong upward seepage, groundwater temperature is carried into streams. Therefore, groundwater seepage to streams is known to moderate summer and winter stream temperatures and to create so-called local thermal refugia [Kaandorp et al., 2019] and climate refugia [Briggs et al., 2018] for aquatic biota. Temperature is a major determinant for drinking water quality, since it influences physical, chemical and biological processes, such as absorption, decay [Monteiro et al., 2017], growth and competition [Prest et al., 2016]. Water treatment processes are also influenced by water temperature, which must be between 10 °C and 15 °C to ensure the best potabilization conditions. Although the link between energy and water fluxes is clear and the processes are well known [Anderson, 2005], studies on water resources include energy fluxes only *via* the evapotranspiration sink term of the water balance [Singh, 2014].

Understanding the coupled water and energy fluxes of a hydrosystem at the regional scale has benefits beyond the understanding of physical processes. First, quantifying all fluxes at the hydrosystem scale allows decision makers to develop water policies [Flipo et al., 2020] regarding water demand and energy generation or to mitigate the risks of contamination [Briggs et al., 2014] or climate change [Taylor et al., 2013], and is therefore crucial. Especially, the impact of climate change on densely populated hydrosystems is highly uncertain, while the demand for water and energy increases rapidly with population growth [de Marsily, 2009]. Second, critical infrastructure such as nuclear power, water treat-

ment plants depend on water temperature for cooling and uncertainties due to hydrological extremes [Boogert and Dupont, 2005] pose a risk within the water–energy–food nexus framework. Third, proper utilisation of geothermal sources, which has a great potential to replace fossil fuels, requires the heat budget estimation. Therefore, examining how water budget and subsurface energy may change in response to both climate-driven and anthropogenic effects is crucial to a holistic water management in terms of quantitative and qualitative issues.

Between the different approaches for water temperature modeling [see Caissie, 2006], deterministic models offer the benefit of scenario testing for management purposes, which is critical in understanding the impact of climate change on river basins. Most attempts are local 0D models, often applied in many points [e.g. Bustillo *et al.*, 2014], or 1D models mostly used in small river basins [e.g. Wondzell *et al.*, 2019]. Notable exceptions are the worldwide modelling of Van Vliet *et al.* [2012], van Vliet *et al.* [2013] and the high-resolution basin scale model of Beaufort *et al.* [2016] in the Loire basin. In all these applications, groundwater temperature is never explicitly modelled, and groundwater contribution to river temperature is either overlooked or approximated by empirical relationships. The same applies to input heat fluxes from rainfall and/or surface runoff. Qiu *et al.* [2019] and Loinaz *et al.* [2013] attempted to simulate river temperatures explicitly, although these applications were limited to small watershed scales. A challenge in the explicit simulation of river and groundwater temperatures is establishing the temperature of recharge and runoff.

To address these issues, a numerical tool has been developed to jointly model water and energy flows with a focus on surface and groundwater at regional scale. The tool couples the SVAT (Soil-Vegetation-Atmosphere-Transfer) model ORCHIDEE and the hydrological–hydrogeological model CaWaQS. ORCHIDEE is capable of simulating coupled water and energy fluxes at surface. The novelty of our approach is to combine surface temperatures and recharge and runoff fluxes simulated by ORCHIDEE with the process based CaWaQS model to explicitly simulate river and aquifer water and heat fluxes. For this purpose, an original transport library based on the resolution of the diffusion/advection transport equation was developed to simulate heat transfer in both 1D river

networks, and pseudo-3D aquifer systems. In addition, an analytical solution is used to simulate heat transport through aquitard layers and streambed.

We selected the Seine River basin to test our approach, as it is under increased stress due to climate change and it possesses distributed data set for process-based models. The Seine River basin, one of the major aquifer systems in Europe, faces major water resource issues due to a variety of uses, e.g. drinking water, cooling and heating of buildings by geothermal or river-based energies, agricultural withdrawals, cooling of the nuclear power plants, that provide electricity in the region, among others.

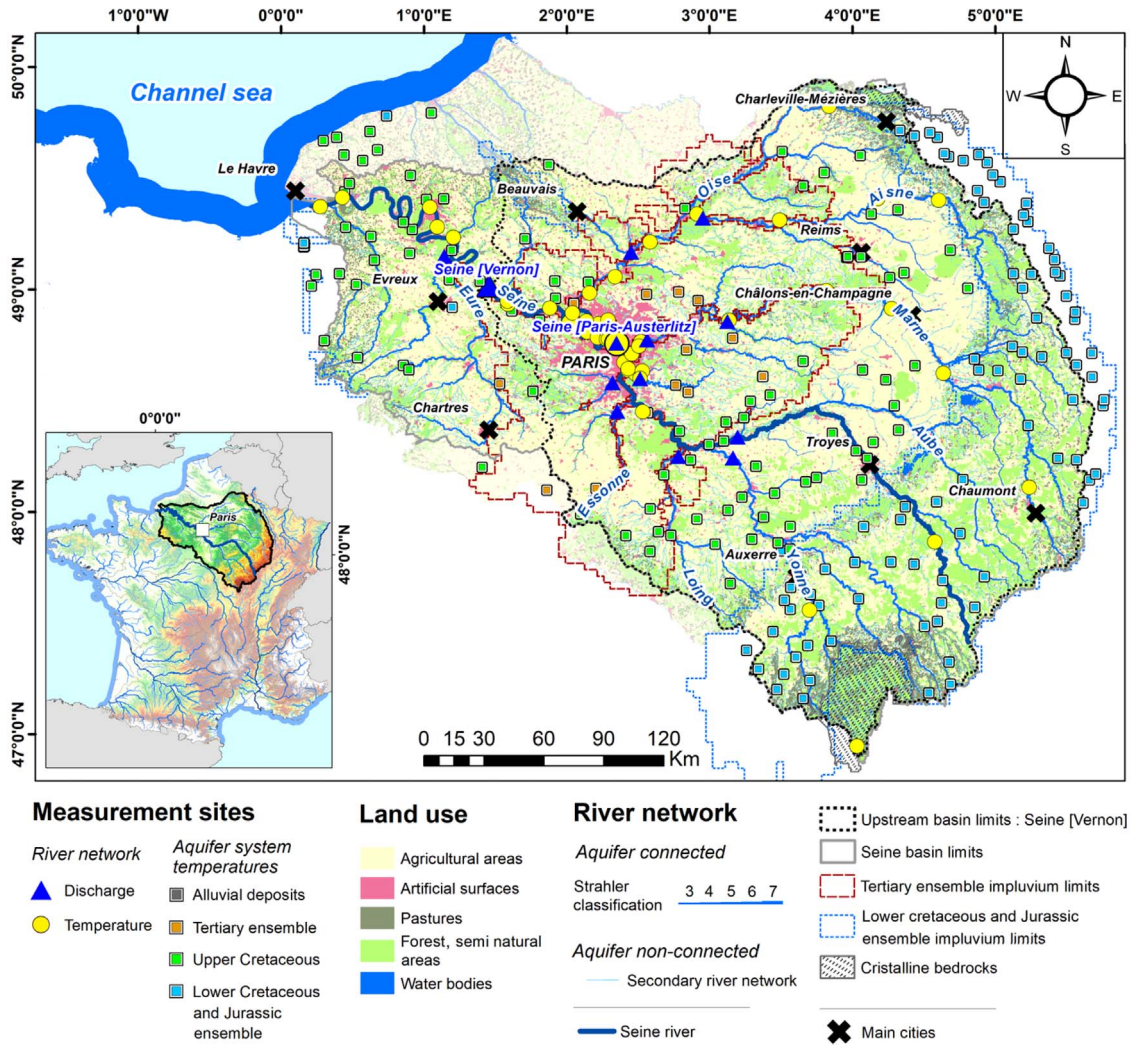
In the following, the numerical tool is first presented with a focus on SVAT–subsurface coupling. The calibration of surface water balance is demonstrated. Dynamics and orders of magnitude of water and heat budgets are detailed in the case of long-term simulations, to answer three research questions: (i) what is the hydrological and thermal functioning of the Seine hydrosystem in current climate conditions? (ii) What are the predominant heat sources in the Seine aquifer system and river network, resulting in the formation of extensive subsurface urban heat islands on groundwater temperature? (iii) What is the role of the surface runoff and river–aquifer exchanges in the river thermal load?

## 2. Material and methods

### 2.1. Study area: the Seine hydrosystem

The Seine basin ( $\approx 78,650 \text{ km}^2$ ) is entirely located within the Paris sedimentary basin in northern France. The geological settings consists of concentric tertiary sedimentary formations (alternating clay, sandstone and limestone), Cretaceous Chalk and Jurassic limestones lying on a basement of Hercynian crystalline rock formations, mainly outcropping at the extreme south-east (Morvan) and north-east (Ardennes) limits of the basin. The highest altitude is 856 m above sea level (asl) and 90% of the basin is below 300 m. The river network shows gentle slopes and climate does not exhibit sharp geographical gradients.

The climatic regime of the basin is pluvial oceanic, modulated by seasonal variations in evapotranspiration. The mean annual precipitation rate is 825 mm over the 1980–2020 period [SAFRAN dataset, Vidal



**Figure 1.** General overview of the Seine basin area (78,650 km<sup>2</sup>). Land use as provided by the 2018 *CorineLandCover* database is used as map background. Blue triangles and yellow circles represent river discharge and temperature measurement locations, respectively. Colored dashed lines indicate the extent of the entire aquifer system. Aquifer temperatures stations are grouped according to the main aquifer ensemble, they are associated with.

et al., 2010]. Flow at the downstream Seine river was determined at the gauging station of Vernon before the estuary. The mean discharge at the station was 490 m<sup>3</sup>·s<sup>-1</sup> over the 1980–2020 period (HYDRO database). Furthermore, variographic studies have shown the stationarity of both groundwater and river water stock over this time period [Flipo et al., 2012].

Currently, the Seine basin is the most urbanized and industrialized basin in France [Flipo et al., 2020]. It includes 17 million inhabitants (25% of the

national population), with 10 million in the Paris conurbation alone, and 40% of national industrial activities, while a large zone around the huge Paris conurbation is oriented towards the mass production of cereals and industrial crops (Figure 1). This scale of activity creates an enormous energy demand, which can be partially offset by exploiting geothermal energy [Bayer et al., 2019]. However, this potential is yet to be quantified by estimating the energy budget of the aquifer system. Given the large population and

food production of the Seine basin, water resources are of high strategic importance, as about 1 km<sup>3</sup> of groundwater is extracted every year over the whole basin [Flipo *et al.*, 2022].

## 2.2. Numerical tools and model development

### 2.2.1. ORCHIDEE (tag 2.2, rev. 6533)

ORCHIDEE is a SVAT model, initially designed to be coupled to an atmospheric global circulation model [Krinner *et al.*, 2005], but it can also be used in stand-alone mode, using an input meteorological data set. ORCHIDEE tag 2.2 was developed as part of the IPSL-CM6 model [Boucher *et al.*, 2020] to contribute to the sixth phase of the Coupled Model Inter-comparison Project (CMIP6) [Eyring *et al.*, 2016].

The horizontal resolution is imposed by the atmospheric grid, and the energy and water budgets of the land surface are calculated in each grid-cell at a sub-hourly time step (15-min in the coupled model, 30-min in stand-alone mode) to properly account for the diurnal variations of incoming radiation. These budgets depend strongly on land cover, described in each grid-cell as a mosaic of 15 plant functional types (PFTs) including bare soil, evergreen and deciduous trees, C3 and C4 grasses and crops. All PFTs share the same equations but with different parameters. The proportion of each PFT is prescribed in each grid cell from land cover maps derived from satellite observations, with yearly updates [Lurton *et al.*, 2020]. Two options are compared in this work regarding the evolution of leaf area index (LAI), which can be either prescribed via input maps of monthly LAI, or calculated dynamically at a daily time step by the so-called STOMATE module, which redistributes the net primary production (photosynthesis minus autotrophic respiration) owing to PFT dependent rules of allocation and phenology.

The vertical soil moisture redistribution is modelled by a multi-layer solution of Richards equation, using a 2-m soil discretized into 11 soil layers, with free drainage as a bottom boundary condition. Infiltration and infiltration-excess surface runoff are derived from a Green-Ampt parameterization [d'Orgeval *et al.*, 2008]. The unsaturated values of hydraulic conductivity and diffusivity are given by the Van Genuchten–Mualem relationships. In each grid cell, the corresponding parameters (including

saturated hydraulic conductivity  $K_{\text{sat}}$  and porosity) are deduced from soil texture as detailed in Tafasca *et al.* [2020].  $K_{\text{sat}}$  (in  $\text{m}\cdot\text{s}^{-1}$ ) further undergoes a vertical exponential decay to account for the effects of soil compaction and bioturbation:

$$K_{\text{sat}}(z) = K_{\text{sat}}^* \exp(-c_K(z - z_{\text{lim}})), \quad (1)$$

where  $z$  is the depth below the soil surface (m),  $K_{\text{sat}}^*$  is the reference value of  $K_{\text{sat}}$  based on soil texture ( $\text{m}\cdot\text{s}^{-1}$ ),  $z_{\text{lim}}$  is the depth at which the decrease of  $K_{\text{sat}}$  starts (by default equal to 0.3 m), and  $c_K$  is the decay factor (in  $\text{m}^{-1}$ , equal to 2 by default).

Evapotranspiration is composed of four sub-fluxes: sublimation, interception loss, soil evaporation, and transpiration, tightly coupled to photosynthesis. Soil evaporation and transpiration depend on soil moisture via stress factors reducing the effective rates compared to the potential rate, which itself depends on an aerodynamic resistance (and stomatal resistance for transpiration) and the water vapor gradient between the surface and the atmosphere. Evapotranspiration  $E$  (in  $\text{kg}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ ) and its energy equivalent, the latent heat flux  $LE$  (in  $\text{W}\cdot\text{m}^{-2}$ ), couple the water and the energy budgets, calculated at the grid-cell scale given by:

$$C_p \frac{\partial T_s}{\partial t} = (1 - \alpha_s) R_{\text{SWd}} + R_{\text{LWd}} - \varepsilon \sigma T_s^4 - LE - H - G. \quad (2)$$

Surface temperature  $T_s$  is the prognostic variable (K),  $t$  is time (s),  $C_p$  is the heat capacity of the surface layer ( $\text{J}\cdot\text{K}^{-1}\cdot\text{m}^{-3}$ ), while  $R_{\text{SWd}}$  and  $R_{\text{LWd}}$  are the input downward radiation fluxes ( $\text{W}\cdot\text{m}^{-2}$ ) in the visible and infra-red wavelengths, respectively. Surface albedo  $\alpha_s$  depends on the albedo of the soil and of each PFT, which are prescribed parameters, and on snow albedo, calculated as a function of snow age by a 3-layer snow module [Wang *et al.*, 2013]. Finally,  $H$  is the sensible heat flux ( $\text{W}\cdot\text{m}^{-2}$ ), and  $G$  is the ground heat flux ( $\text{W}\cdot\text{m}^{-2}$ ), which couples the surface energy budget to the soil temperature profile, resulting from a 1D heat diffusion equation. Soil freezing is neglected since the Seine River basin is predominantly temperate. The vertical discretization for heat diffusion is identical to that adopted for water down to 2 m, then further extended to 18 m where a zero heat flux is assumed. The soil thermal properties (heat capacity and conductivity) of each layer depend on soil moisture (assuming a uniform profile below 2 m) and soil texture [Wang *et al.*, 2016].

### 2.2.2. CaWaQS (v3.17)

Mostly relying on the former MODCOU-NEWSAM software [Ledoux *et al.*, 1989] based on the blueprints of de Marsily *et al.* [1978], the physically-based CaWaQS3.x [Flipo *et al.*, 2022, 2007] computes, at a daily time-step, the main physical processes controlling the water budget and flow dynamics in each compartment of a hydrosystem. Each CaWaQS library deals with a specific compartment of a hydrosystem. This feature allows users to bypass certain modules when necessary and makes CaWaQS suitable for coupled applications.

CaWaQS is capable of distributing forcing fluxes over its own grid, and accepting forcing fluxes at various grid sizes. Forcing runoff fluxes are aggregated by CaWaQS to local sub-catchment areas and directly routed to the river network in which river discharges are simulated using a Muskingum scheme and water levels are simulated using a Manning–Strickler approach [Saleh *et al.*, 2011]. Saturated groundwater flows are numerically solved using a finite difference solution of the 2-D diffusivity equation on a multi-layered nested grid. Vertical exchanges between aquifer layers and stream-aquifer exchanges depend on drainance and conductance concepts respectively i.e. a 1-D steady-state approach, assumed to be linearly linked to the head difference between two aquifer layers or river water level-aquifer head difference [Flipo *et al.*, 2014]. Aquifer recharge from the unsaturated zone as well as anthropogenic withdrawals are treated as source terms.

Recent advances added a transport module, which simulates heat and solute transport in porous media (aquifer, pseudo-3D resolution) and the free surface compartment (river network, 1-D approach). Transport is described using the following advection-diffusion equation:

$$\nabla \left[ \frac{\lambda}{\rho_w C_w} \nabla T_w - q T_w \right] = \frac{\rho C}{\rho_w C_w} \frac{\partial T_w}{\partial t}, \quad (3)$$

where  $T_w$  is the water temperature (K),  $q$  is the specific discharge ( $\text{m}\cdot\text{s}^{-1}$ ),  $\lambda$  is the thermal conductivity of the porous medium ( $\text{W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$ ),  $\rho_w$  and  $\rho$  are the densities of water and the porous medium (both in  $\text{kg}\cdot\text{m}^{-3}$ ), and  $C_w$  and  $C$  are the specific heat capacities of water and porous medium ( $\text{J}\cdot\text{kg}^{-1}\cdot\text{K}^{-1}$ ).

River-atmosphere heat exchanges are described based on an energy budget equation similar to (2) but for water temperature  $T_w$ ; the various heat fluxes

were calculated following Magnusson *et al.* [2012]. Whilst the energy income from the sun is the largest source of energy for a river system, clouds, riparian vegetation and topography-related shading can reduce heat input significantly [Loicq *et al.*, 2018]. The limited permeability of aquitard layers constrains water flow to the vertical direction between aquifers [de Marsily, 1986]. Similarly, streambed flow is also dominated by vertical exchanges between the stream and its underlying aquifer layer. Consequently, heat flow between two adjacent aquifers separated by an aquitard layer or a streambed linking the river with the aquifer layer can be resolved by a 1-D analytical solution. In this study, the heat exchanges through aquitard layers and streambed were resolved by the analytical solution of Kurylyk *et al.* [2017]:

$$\alpha_i \frac{\partial^2 T_i(z, t)}{\partial z^2} - q \frac{\partial T_i(z, t)}{\partial z} = 0 \quad (i = 1, 2, \dots, n-1), \quad (4)$$

where  $z$  is the depth (m),  $n$  is the number of layers,  $\alpha_i$  is the bulk thermal conductivity  $\lambda$  of layer  $i$  divided by the volumetric heat capacity of water ( $\text{m}^2\cdot\text{s}^{-1}$ ), and  $T_i$  is the temperature of layer  $i$ . This solution is an extension of the method initially proposed by Bredehoeft and Papaopulos [1965], to allow for variations in thermal conductivity with depth. The equation is based on four fundamental assumptions: (i) heat transfer is restricted to one dimension; (ii) Darcy flux is constant and uniform throughout each layer; (iii) temperature profiles are at steady-state during each time step; and (iv) properties are homogeneous within each layer.

The aquitard interface is a 3-layered system with 3 unknowns, i.e. the temperatures of the aquitard, as well as of the top and bottom aquifers. The streambed interface is a 2-layered system with 3 unknowns, i.e., the streambed, river and aquifer temperatures. The analytical solution coefficients are incorporated as a term of the solution matrix of the porous media, thereby simultaneously resolving the heat transport in the aquitard. The river temperature in the previous time-step is used to calculate the streambed temperature at the stream-aquifer interface. The streambed temperature calculated by the analytical solution is then used as the lower boundary of the river transport simulation. The developed transport module was validated using the analytical solution of Goto *et al.* [2005] [Kilic *et al.*, 2021, Wang *et al.*, 2021].

### 2.2.3. Model coupling strategy

Coupling is achieved *via* an “offline” forcing procedure, whereby the models involved are sequentially run, thus only allowing one-way interaction, from ORCHIDEE to CaWaQS. To this end, selected ORCHIDEE simulation results are used as daily CaWaQS upper boundary conditions: drainage fluxes simulated by ORCHIDEE are prescribed as CaWaQS aquifer recharge, with a temperature equal to the one in the soil at 2-m, while surface runoff is routed to the river system, assuming its temperature is at equilibrium with the soil surface temperature (0.001 m depth, top layer).

The correspondence between the ORCHIDEE grid mesh and the elementary surface units of CaWaQS is handled by CaWaQS, *via* spatially weighted area-based arithmetic means. The coupled model scheme internally projects each element to its geographic coordinate and then passes the forcing flux information to the receiving model.

### 2.2.4. Numerical experiment design

The coupled tool was applied over the 2001–2018 period, which was noteworthy for remarkable flood events (e.g. winter/spring 2001, winter 2016–2018) as well as extremely low-flow/drought periods and heat-waves (e.g. summer 2003, summer 2006–2007).

**ORCHIDEE setup.** ORCHIDEE is run over the Seine basin in stand-alone mode, forced with the SAFRAN meteorological data set [Vidal *et al.*, 2010]. It provides the required atmospheric variables at an hourly time step, over a regular  $8 \times 8$  km grid, which is therefore the horizontal resolution of the simulations. A total of 1490 cells are used to describe the simulated domain. A constant carbon dioxide concentration (370 ppm) over the period 1980–2018 was applied.

The vegetation distribution map was derived at  $0.5^\circ$  resolution from the ESA-CCI Land Cover dataset at 300 m resolution, for the 1992–2015 period [see details in Lurton *et al.*, 2020]. Vegetation distribution outside this period was set to that of the nearest available year with a constant vegetation distribution. For simulations with prescribed LAI, the monthly values were obtained from an existing 40-year global simulation performed in stand-alone mode at  $0.5^\circ$  ( $55$  km resolution). The soil background albedo map was derived from the MODIS albedo dataset aggregated at  $0.5^\circ$  resolution. Soil texture distribution maps

were obtained from four sources: Reynolds [Reynolds *et al.*, 2000], SoilGrids [Hengl *et al.*, 2017], LUCAS [Ballabio *et al.*, 2016], and INRA [Jolivet *et al.*, 2006]. The initial state was obtained after a five-year warm-up period (1980–1985) in order to reach the land surface model equilibrium state for soil moisture and soil temperature.

**CaWaQS setup.** A pre-existing CaWaQS application over the Seine basin, which considers the main aquifer units in interaction with the Seine river and its main tributaries [Flipo *et al.*, 2022], was used here. From the oldest to the most recent, these aquifers can be grouped according to the following four main ensembles: (i) an ensemble of aquifer and aquitard units, ranging from Lower Jurassic (Hettangian stage) to Lower Cretaceous (Albian stage), (ii) the large Upper Cretaceous chalk aquifer, (iii) a Tertiary aquifer complex including the five main aquifer layers dating from Palaeocene to Miocene stages and (iv) recent alluvial deposits (Pleistocene, Holocene). The resolution of CaWaQS aquifer cells range between 100 m–3200 m using a nested-grid approach. River geometry was obtained from the SYRAH-CE [Chandesris *et al.*, 2008] database. River network water inputs account for subsurface runoff, overflows from aquifer to soil surface as well as river–aquifer exchanges. Mainly due to computation time concerns, river–aquifer exchange calculations were restrained to the main hydraulic network, which corresponds to river reaches with a Strahler order above or equal to 3 (Figure 1). The thermal parameterization (thermal conductivities, heat capacities and porosities) of the Seine basin was extracted from Dentzer [2016] and Rivière [2012]. Heat transport was simulated in the river reaches with a Strahler order above or equal to 3 (Figure 1). Soil surface temperatures obtained from ORCHIDEE were prescribed to river reaches below Strahler order 3. The initial temperature state of the aquifer system was obtained by running the coupled model for 40 years by re-cycling 2 sets of 20-year simulations in the period 1985–2005.

CaWaQS forcings were defined as follows: (i) A no-flow boundary condition is imposed at the bottom and border limits of the domain, except where the mesh boundary corresponds to an adjacent river. (ii) Groundwater withdrawals are taken into account as a daily pumping rate, inferred from the annual groundwater abstraction data of the Seine basin wa-

ter management agency (Agence de l'eau Seine-Normandie, AESN) for domestic, industrial and agricultural needs. (iii) Agricultural withdrawals are assumed to take place only during the summer period (i.e. bulk irrigation water volume is distributed as: 20% in May, 30% in June and July, and 20% in August). (iv) A no-heat flux boundary condition is imposed for the horizontal borders. (v) A Dirichlet boundary condition is assigned to the top boundary of the groundwater system for transport. (vi) Geothermal fluxes are assumed for the bottom limit of the domain [75 m·Wm<sup>-2</sup>, Dentzer, 2016].

### 2.2.5. Calibration protocol

**Calibration strategy.** A two-step calibration procedure was used to calibrate the water and heat cycle of the Seine Hydrosystem. Considering the impact of North Atlantic Oscillation (NAO) on the groundwater and river stock variation [Flipo *et al.*, 2012], a 17-year averaging period between 2001–2018 was selected to conduct the first step of the optimization procedure.

The surface hydrology calibration step consists of iterative trial-and-error adjustments of the ORCHIDEE parameters to minimize the bias between simulated and observed mean pluri-annual discharge at Vernon, the most downstream gauging station of the Seine before the estuary (Figure 1). The bias was calculated after the simulated pluri-annual mean discharge (estimated by total runoff times basin area) was corrected for yearly groundwater abstraction (annual average equivalent of 37 m<sup>3</sup>·s<sup>-1</sup> during 2008–2012, AESN) from the basin. Previous studies [Flipo *et al.*, 2022] indicate that around two-thirds of effective rainfall contributes to aquifer recharge in the basin, and this ratio was also considered as a qualitative target for the calibration.

A set of 10 ORCHIDEE simulations were chosen to illustrate this iterative process (Table 1). First, EXP1–EXP4 allow us to assess the influence of the soil texture map (with a prescribed LAI map and default parameters, as used at global scale for CMIP6). Then, the two available options to account for LAI variations (prescribed map vs dynamic simulation by the STOMATE module) can be assessed by comparing EXP4 and EXP5. Finally, EXP5–EXP10 only differ by some ORCHIDEE parameters found to be effective by several sensitivity analyses [e.g. Raoult *et al.*, 2021]: the saturated hydraulic conductivity ( $K_{\text{sat}}$ ) assigned to the dominant soil textures in the Seine River basin

(loamy sand, loam, clay loam) and its decay factor  $c_K$  (1); the visible and near-infrared leaf albedo and maximum LAI of the main PFTs in the basin. It must be noted that no calibration was carried out to improve the thermal fluxes. The temperatures of soil bottom and soil top layers were assumed to be at equilibrium with the recharge (drainage) and surface runoff temperatures, respectively.

CaWaQS parameterizations of river network and groundwater were taken from the pre-calibrated parameter set provided by Flipo *et al.* [2022]. No additional calibration of the hydrological components and groundwater heat transport of CaWaQS were undertaken, but the river heat budget was calibrated by modifying the uniform shading factor ranging from 0% to 50% applied on the net incoming radiation on the river [Beaufort *et al.*, 2016]. A 30% uniform shading, consistent with reported shading factor values [Loicq *et al.*, 2018], was selected after a series of simulations.

**Evaluation strategy.** To validate the coupled model performance, the Seine basin hydrogeological and thermal application results were compared to a set of river discharge and water level monitoring wells as well as river and aquifer temperature data, distributed among the different aquifer units (Figure 1). The Kling–Gupta Efficiency coefficient (KGE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Percent bias (PBIAS) and the Pearson correlation coefficient (CC) were used to quantitatively evaluate the misfit between simulated and observed values and dynamics.

Field data were collected between 2001–2018. The variables collected are discharge data from 183 gauging stations; hydraulic head data from 267 piezometers were used to evaluate the performance statistics of the hydrology part. Discharge and hydraulic head time series data were obtained from publicly available ADES and HYDRO national datasets. River water temperature data were assembled from various agencies operating in the basin; the dataset comprises hourly and daily observations from 62 sites [Rivière *et al.*, 2021]. During the simulation period (2001–2018), 48 stations with hourly observations were selected. The aquifer temperature data used were obtained from the national ADES database. While providing a comprehensive amount of data, records are often incomplete and heterogeneous. The ma-



**Table 1.** Parameter set applied to the 10 experiments to calibrate the ORCHIDEE model

Experiment		EXP1	EXP2	EXP3	EXP4	EXP5	EXP6	EXP7	EXP8	EXP9	EXP10
Soil distribution		I	SG	L	R	R	R	R	R	R	R
Vegetation model		pLAI	pLAI	pLAI	pLAI	ST	ST	ST	ST	ST	ST
Parameter	Class										
$c_K$		2	—	—	—	—	4	—	—	10	—
$K_{\text{sat}}$ ( $\text{m}\cdot\text{s}^{-1}$ )	Loamy sand	$4.05 \times 10^{-5}$	—	—	—	—	—	—	$1.73 \times 10^{-5}$	—	—
	Loam	$2.88 \times 10^{-6}$	—	—	—	—	—	—	$1.50 \times 10^{-6}$	—	—
	Clay loam	$7.22 \times 10^{-7}$	—	—	—	—	—	—	$4.62 \times 10^{-7}$	—	—
Near-infrared leaf albedo	TeBE	0.22	—	—	—	—	—	0.2	—	—	—
	TeBS	0.22	—	—	—	—	—	0.2	—	—	—
	NC3	0.3	—	—	—	—	—	0.25	—	—	—
	AC3	0.3	—	—	—	—	—	0.1	—	—	—
Visible-light leaf albedo	TeBE	0.06	—	—	—	—	—	0.04	—	—	—
	TeBS	0.06	—	—	—	—	—	0.04	—	—	—
	NC3	0.1	—	—	—	—	—	0.05	—	—	—
	AC3	0.1	—	—	—	—	—	0.04	—	—	—
Maximum LAI	TeBE	4.5	—	—	—	—	—	—	—	—	5.5
	TeBS	4	—	—	—	—	—	—	—	—	5
	NC3	2	—	—	—	—	—	—	—	—	3
	AC3	2	—	—	—	—	—	—	—	—	3

—: same value than the previous one, R: Reynolds, SG: SoilGrids, L: Lucas, I: INRA, pLAI: prescribed LAI, ST: STOMATE activated, TeBS: Temperature Boreal Summergreen Trees, TeBe: Temperate Boreal Evergreen Trees, NC3: Natural  $C_3$  grass, AC3: Agricultural  $C_3$  crops.

majority of the temperature data were collected on site while pumping, hence the depths of the temperature observations represent the depth of the screened well. A pre-processing was therefore applied to the dataset using the interquartile range (IQR) method, as described in Hemmerle *et al.* [2019]. From the initial data set of 9179 stations over the Seine-Normandy region, 212 stations with a minimum of 40 observations over the simulation period remained after the pre-processing.

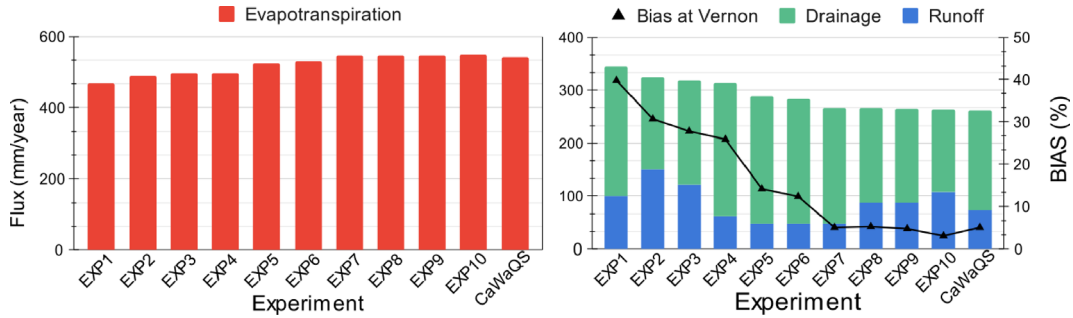
### 3. Results and discussions

#### 3.1. Calibration of the surface water cycle

The simulated water balance, shown in Figure 2, illustrates how the different parameter configurations (Table 1) influence the partitioning of the precipitation into actual evapotranspiration (AET), surface

runoff and recharge (drainage). The first four simulations (EXP1-4) show the impact of the soil texture map, which induces a 6% variation of AET (mean values between 469 and 497  $\text{mm}\cdot\text{a}^{-1}$ ), also reflected in the total runoff (effective precipitation). Soil texture also impacts the partitioning of the total runoff between drainage and surface runoff, which varies between 20 and 47% of the total runoff between EXP4 (Reynolds) and EXP2 (SoilGrids), as the SoilGrids and INRA maps include a significant fraction of silty soils that increase the surface runoff at the expense of drainage [Tafasca *et al.*, 2020]. Eventually, the Vernon discharge bias decreases from 39.8% in EXP1 to 25.9% in EXP4, so we kept the Reynolds soil texture map for further optimization.

The dynamic calculation of the vegetation dynamics by the STOMATE module (EXP5) allows a further increase in AET (+26  $\text{mm}\cdot\text{a}^{-1}$ ) compared to simulations with prescribed LAI (EXP1-4). The reason is



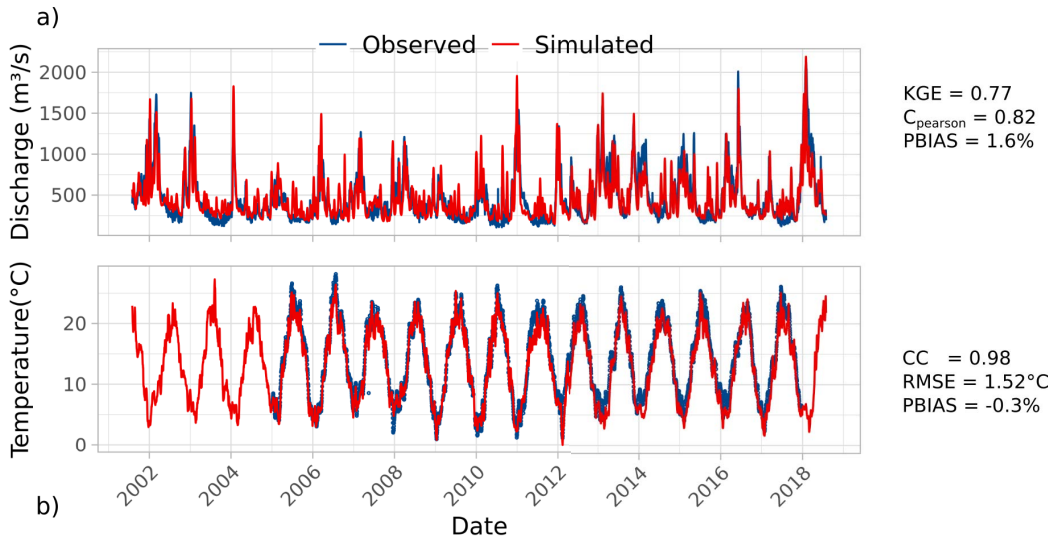
**Figure 2.** Mean yearly surface water budget components calculated by ORCHIDEE in  $\text{mm}\cdot\text{a}^{-1}$ . Bias was calculated on the multi-year mean discharge at Vernon gauging station after the adjustments on the mean yearly water abstractions are taken into account for EXP1-10. CaWaQS column represents the water balance calculated by the CaWaQS-Seine application including water abstractions.

that STOMATE calculates a higher LAI than the prescribed one, increasing the AET via a larger biomass during the growth season. The resulting decrease in total runoff leads to a further reduction of the Vernon discharge bias decreases, from 25.9% to 14.2%. The small decrease in the contribution of surface runoff to total runoff ( $-3\%$ ) can be attributed to the trade off between surface runoff and increased vegetation activity and water abstraction of roots in the topsoil. Reducing the NIR and VIS leaf albedo (from EXP6 to EXP7) also yielded a significant increase in AET ( $+16 \text{ mm}\cdot\text{a}^{-1}$ , i.e.  $+5\%$ ), because it increases the net available energy. This brings a further reduction of the discharge bias at Vernon to 5.1% in EXP7.

In contrast, implementing a sharper vertical decay of  $K_{\text{sat}}$  below 30 cm in the soil column (from EXP5 to EXP6) showed a weak impact on all the studied variables: AET improved slightly by only  $4 \text{ mm}\cdot\text{a}^{-1}$ , the discharge bias by only 1.8%, the partition of the total runoff, and the surface runoff fraction was hardly changed. This weak sensitivity to modifications of the vertical  $K_{\text{sat}}$  profile shows the importance of the top 30 cm of the soil column on the water balance rather than deeper zones [De Rosnay *et al.*, 2000]. The decrease of  $K_{\text{sat}}$  for all the dominant soil textures (EXP7–EXP8) had an even smaller impact on AET, total runoff thus on the discharge bias (even slightly worsened). Decreasing  $K_{\text{sat}}$ , however, showed a significant impact on the partition of total runoff, as it nearly doubled the contribution of surface runoff (from 18% to 33%). The additional decrease of  $K_{\text{sat}}$  in the bottom part of the soil by increasing the decay factor  $c_K$  to 10 (EXP9) had no impact on the water

and energy budget, and the increase of maximum LAI (EXP 10) had a very limited effect on AET and total runoff although it increased the proportion of surface runoff ( $+10\%$ ). The maximum LAI values in all simulations including EXP10 match the reported LAI ranges observed throughout the FluxNET network for similar biomes and climates [Krinner *et al.*, 2005]. A detailed calibration of vegetation properties is beyond the scope of this study, although ways to improve the terrestrial biosphere and energy processes in SVAT models are active research domains [Peylin *et al.*, 2016].

Eventually, these results show that the overestimation of discharge at the Vernon outlet resulted from an underestimated AET. A considerable improvement of the discharge bias came from the dynamic LAI calculations, which may partly be attributed to the coarse resolution of the input LAI map in pLAI simulations, and to a better seasonal cycle of LAI (comparison to GiMMS [Zhu *et al.*, 2013] product not shown), more consistent with the SAFRAN meteorological forcing. Although groundwater withdrawals are taken into account in the calculation of the ORCHIDEE discharge bias (see Section 2.2.5), the lack of irrigation in the simulations may also contribute to the underestimation of AET and the overestimation of river discharge since the enhancement of AET by surface water withdrawals is overlooked. Considering that ORCHIDEE simulates a natural basin contrary to heavily regulated and irrigated Seine basin, a 3.1% bias (EXP10) at the outlet of the basin (Vernon) is very satisfactory. However, when the proportions of infiltration and surface runoff are considered (see



**Figure 3.** Comparison between (a) simulated (red) and observed (blue) discharge at the outlet of Seine watershed (Vernon); (b) simulated (red) and observed (blue) river temperature at Paris (Austerlitz). From August 2001 to July 2018. Years indicate 1st of January. KGE: Kling–Gupta Efficiency coefficient; CC: Correlation Coefficient; PBIAS: Percent Bias; RMSE: Root Mean Square Error.

Section 2.2.5), EXP9 performs better with nearly two-thirds of recharge contributing to total runoff.

### 3.2. Performance of the ORCHIDEE-CaWaQS coupled simulation

The best performing ORCHIDEE simulation in terms of river discharge bias and partitioning of effective rainfall (EXP9) was selected to force the CaWaQS simulation. A coupled simulation was performed over the 17 year period (August 2001–July 2018). The simulated discharge compared satisfactorily with observed stream flow in 14 main sub-catchments, as illustrated at the outlet of the Seine basin (Vernon) in illustrated in Figure 3a. Although low flows are slightly overestimated by the coupled model, it captures the yearly dynamics properly (KGE = 0.77, CC = 0.82) and mean volume (PBIAS = 1.6%), given that the measurement error of river discharge can reach 20% [Sauer and Meyer, 1992].

Stream temperature is also well simulated along the Seine basin, as illustrated by Paris-Austerlitz station (Figure 3b). The simulated temperatures match the observed ones (CC = 0.98) with a PBIAS of -0.3% and an RMSE of 1.52 °C, although simulated temperatures are generally underestimated in both

summer and winter. Part of this uncertainty can be attributed to the uniform shading factor or the lack of urban structures that artificially modifies the heat regime on riverbanks [Ferguson and Woodbury, 2007]. The model also uses a uniform and time-constant shading factor which deviates from the prevalent seasonal dynamics and spatial distribution of the vegetation in the Seine basin (Figure 1). Model performances respectively evaluated against stream and aquifer temperature probes over the simulation period are provided in Tables 2 and 3.

The dynamics of the heat transport in the river network is well represented (CC > 0.7 for 100% of the stations) and MAE is lower than 2 °C for 88% of the stations. 96% of the river temperature stations have an RMSE lower than 4 °C. Water temperatures at streams and small rivers (<10,000 km<sup>2</sup>) show a higher mean RMSE (2.58 °C) than downstream ( $\geq 10,000$  km<sup>2</sup>) stations (1.58 °C). Prescribed temperatures on river reaches with a Strahler order below 3 strongly influence the river temperatures on streams and small rivers and contribute to the higher RMSE because the proportion of reaches with a prescribed temperature is higher than downstream reaches. A relatively high RMSE indicates a misfit between simulated and observed temperatures at the

**Table 2.** Statistical criteria (2001–2018 period) from 48 stream temperature probes distributed in the basin

Value range (°C)	RMSE	MAE	RMSE	MAE
	Station count (-)		Cumulative percentage (%)	
[0.0–2.0[	31	42	65	88
[2.0–4.0[	15	1	96	90
[4.0–6.0[	2	2	100	94
[6.0–8.0[	0	3	100	100
Value range (-)	CC	KGE	CC	KGE
	Station count (-)		Cumulative percentage (%)	
]0.7–1.0]	48	43	100	90
]0.5–0.7]	0	2	100	94
]0.4–0.5]	0	1	100	96
]0.2–0.4]	0	2	100	100

Upper table: Root Mean Square Error (RMSE), Mean Absolute Error (MAE). Lower table: Pearson correlation coefficient (CC), and Kling–Gupta Efficiency (KGE) coefficient.

**Table 3.** Statistical criteria (2001–2018 period) from 212 aquifer temperature probes distributed in the basin

Value range (°C)	RMSE	MAE	RMSE	MAE
	Piezometer count (-)		Cumulative percentage (%)	
[0.0–2.0[	93	115	44	54
[2.0–4.0[	58	37	71	72
[4.0–6.0[	16	16	79	79
[6.0–8.0[	17	16	87	87
[8.0–10.0[	8	9	91	91
≥10	20	19	100	100

RMSE: Root Mean Square Error, MAE: Mean Absolute Error.

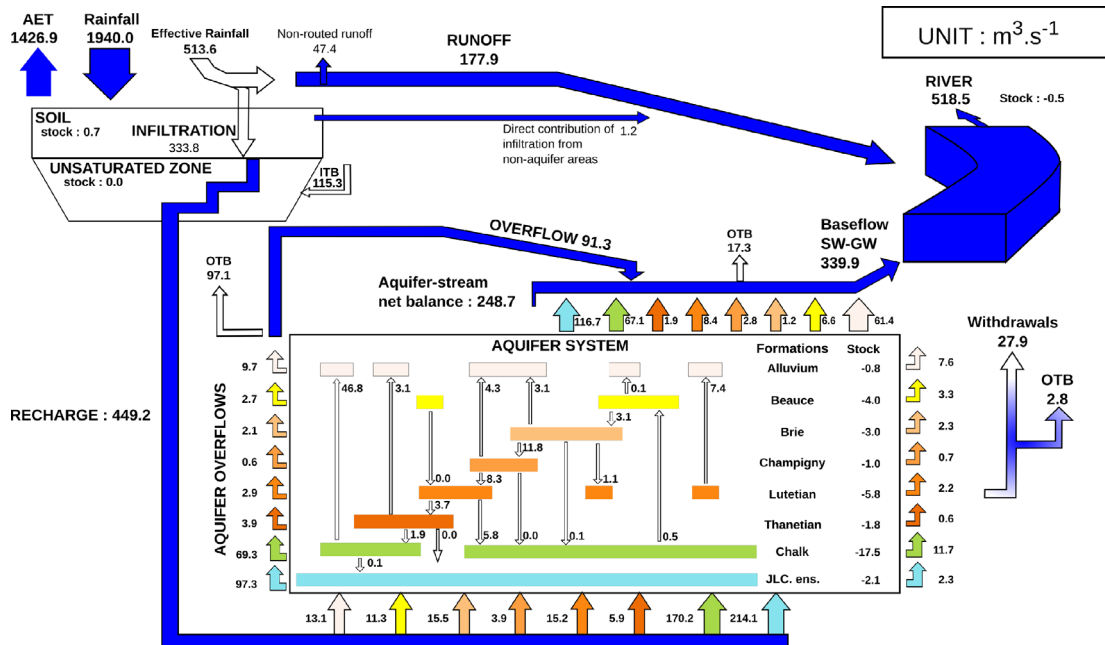
daily scale. Overall the seasonal dynamics are well represented in the river network (Figure 3b). Heat transport in the river performs comparably to other studies at the regional scale using the equilibrium temperature method [Beaufort *et al.*, 2016] in terms of RMSE.

The groundwater temperatures show a good fit as the RMSE is below 2 °C for 44% of the stations. The median RMSE calculated for the aquifer observations was 2.12 °C (ranging between 20.2 and 0.01 °C). Stations with a larger than 10 °C RMSE are focused on the Jurassic aquifer. This is largely due to a mismatch between the depth of the well and the depth of the cell. However, only 22% of the stations show a CC higher than 0.5. Only 3% of the stations show a KGE performance greater than 0.5. The limiting fac-

tor in terms of the dynamics is the lack of sufficient time series in the aquifer system. Regarding the assumptions for the estimates of the upper boundary conditions for heat transport in the groundwater system, simulated soil temperatures at 2 m depth from ORCHIDEE were applied to the groundwater system top boundary, which are quite different from the actual temperatures of groundwater at the water table [Anderson, 2005].

#### 4. Hydrological and thermal functioning of the Seine hydrosystem

The results of the best ORCHIDEE-CaWaQS simulation were used to compute the water budget (Figure 4) and the energy budget (Figure 5) of the whole



**Figure 4.** Water budget of the Seine hydrosystem over the period 2001–2018 as simulated by ORCHIDEE-CaWaQS application. Groundwater system encompasses the full extent of the aquifers. ITB: Infiltration flux for outcropping aquifer units beyond the Seine basin limits. Flows outside boundaries (OTB) are beyond the extent of the Seine basin. Flows are expressed in inter-annual average values in  $\text{m}^3 \cdot \text{s}^{-1}$ . Rainfall =  $760 \text{ mm} \cdot \text{a}^{-1}$ , AET\* =  $559 \text{ mm} \cdot \text{a}^{-1}$ , Runoff =  $69 \text{ mm} \cdot \text{a}^{-1}$ , Infiltration =  $132 \text{ mm} \cdot \text{a}^{-1}$ . Model layers: “Alluvium”: Alluvium deposits, “Beauce”: Beauce limestones ensemble, “Brie”: Brie limestones and Fontainebleau sands, “Champigny”: Champigny limestones, “Lutetian”: Lutetian limestones, “Thanetian”: Thanetian sands, “Chalk”: Upper Cretaceous chalk aquifer, “JLC. ens.”: Lower Cretaceous and Jurassic ensemble.

Seine basin, including groundwater contributions, over the period 2001–2018.

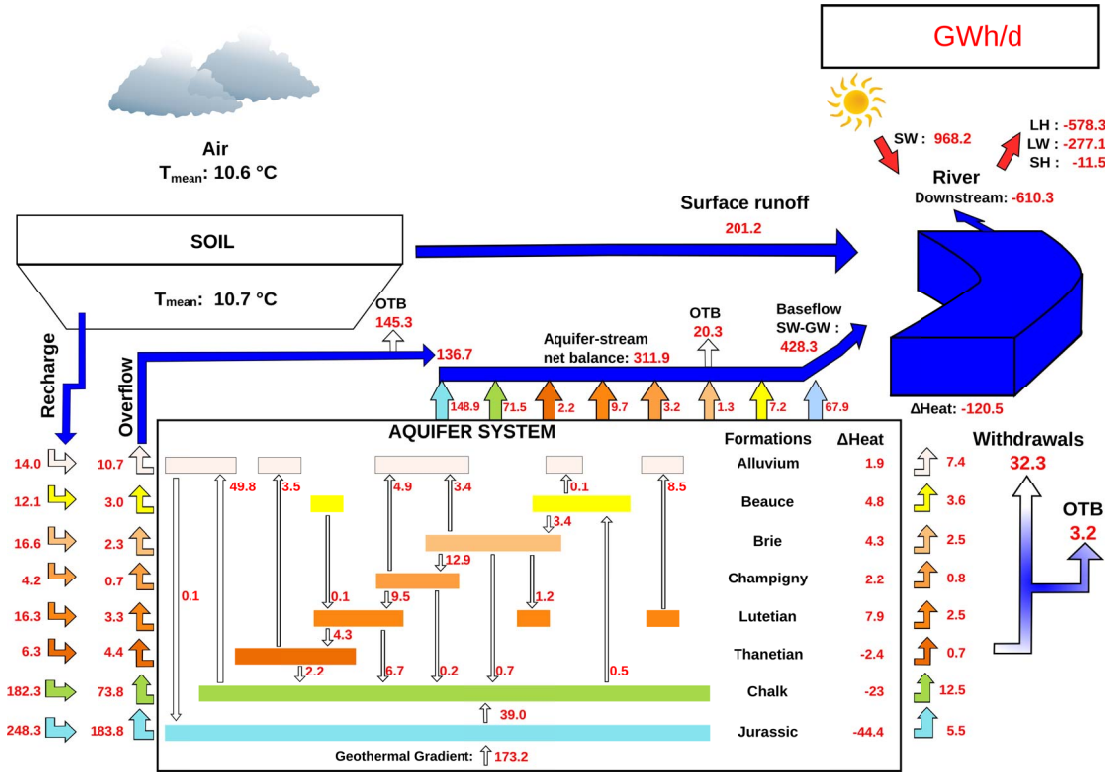
#### 4.1. Pluri-annual water budget

Figure 4 displays a pluri-annual water budget of the Seine basin. The AET represents 73.5% ( $559 \text{ mm} \cdot \text{a}^{-1}$ ) of the rainfall ( $760 \text{ mm} \cdot \text{a}^{-1}$ ). The remaining effective rainfall ( $201 \text{ mm} \cdot \text{a}^{-1}$ ) is partitioned into surface runoff and groundwater recharge, amounting to 34.6% and 65.3% of the effective rainfall, respectively. A total infiltration flow of  $449.2 \text{ m}^3 \cdot \text{s}^{-1}$  transits through the unsaturated zone, recharging the whole aquifer system. 27.8% of the aquifer recharge is lost through regional water fluxes across the Seine basin boundaries (21.6% of recharge) or withdrawn from the system by pumpings (6.2% of recharge). Within the Seine basin, river–aquifer exchanges for Strahler

orders higher than 3 supply roughly half of the discharge ( $248.7 \text{ m}^3 \cdot \text{s}^{-1}$ ) recorded at the outlet of the river network (Figure 4). What is called overflow mostly corresponds to the drainage of the groundwater system by low Strahler orders (one and two). It therefore also corresponds to stream–aquifer exchanges, and amounts to  $91.3 \text{ m}^3 \cdot \text{s}^{-1}$  (17.6% of total discharge). Finally, surface runoff contributes up to  $177.9 \text{ m}^3 \cdot \text{s}^{-1}$  (34.4% of total discharge).

#### 4.2. Pluri-annual heat budget

The pluri-annual heat budget over the period 2001–2018 is illustrated in Figure 5. The pluri-annual average temperature of the air is  $10.6 \text{ }^\circ\text{C}$  in the Seine basin, and the pluri-annual average temperature of the soil surface is  $11.3 \text{ }^\circ\text{C}$  (i.e. the temperature associated to surface runoff). Soil temperature at 2 m



**Figure 5.** Heat budget of the Seine hydrosystem over the period 2001–2018 as simulated by ORCHIDEE-CaWaQS application. Groundwater system encompasses the entire aquifer extents. Flows outside boundaries are beyond the extent of the Seine basin. Energy amounts are expressed in multi-annual average values in GWh per day and mean annual daily air temperature and soil temperature at 2 m depth are shown in °C. Δ Heat: net thermal load variation in the river network. SW: net short-wave radiation, LW: long-wave radiation, SH: sensible heat, LH: latent heat. Model layers: “Alluvium”: alluvial deposits, “Beauce”: beauce limestones ensemble, “Brie”: Brie limestones and Fontainebleau sands, “Champigny”: Champigny limestones, “Lutetian”: Lutetian limestones, “Thanetian”: Thanetian sands, “Chalk”: upper Cretaceous chalk aquifer, “Jurassic”: lower Cretaceous and Jurassic ensemble.

depth is estimated as 10.7 °C. The mean daily river temperature is estimated as 12.1 °C (Figure 5). Yearly mean aquifer temperatures range between 10.6 °C for the shallow alluvium aquifer and 23.8 °C for the deeper Jurassic aquifer. The largest heat gains in the aquifer system are due to the recharge fluxes (182.5 TWh) and the geothermal gradient (63.2 TWh) at the bottom of the basin (Figure 5). The former represents advective fluxes while the latter represents conductive fluxes. The influence of advection diminishes with depth as expected, while conduction becomes the dominant thermal process. This is due to the decrease of water fluxes with depth.

The main sources of surface advective heat losses

are due to: (i) river–aquifer exchanges (113.8 TWh), (ii) stream–aquifer exchanges for the secondary river network (Strahler orders 1 and 2) (i.e. overflow that amounts 102.9 TWh), (iii) and withdrawals (11.8 TWh). The aquifer system loses the most energy via the Jurassic, the Chalk and the alluvium aquifers. The Jurassic loses the largest advective heat fluxes via overflow and river–aquifer exchanges towards the surface (Figure 5). The Chalk aquifer and the alluvium are the second and third largest contributors to river via advective term. The Chalk aquifer is also the largest contributor to thermal losses due to water withdrawal (4.6 TWh).

Energy from short-wave radiation (net radiation,

**Table 4.** Seasonal heat fluxes ( $\text{W}\cdot\text{m}^{-2}$ ) in the Seine river network over the 2001–2018 period

Process	Type	Winter	Spring	Summer	Autumn	Average
Short-wave radiation	Conductive	39.3	142.0	180.1	75.9	109.3
Long-wave radiation	Conductive	-16.1	-42.9	-44.6	-21.5	-31.3
Latent heat	Conductive	-20.3	-69.9	-120.6	-50.4	-65.3
Sensible heat	Conductive	-0.2	-1.7	-2.1	-1.1	-1.3
Streambed conduction	Conductive	0.4	-0.2	-0.7	-0.1	0.2
Surface runoff	Advective	13.3	18.9	37.9	20.8	22.7
Streambed advection	Advective	54.0	54.8	44.3	39.5	48.1
Net gains	—	69.1	101.5	96.2	63.2	82.5

The total river length is 28,378 km. The total surface area of the river network with Strahler order higher than 3 is 369  $\text{km}^2$ . The streambed fluxes are the heat exchanges between river and aquifer. Net gains is the heat balance of the river network. Negative sign (–) indicates losing fluxes by the river.

353.4 TWh) is main non-advective source of heat gain in the river network annually (Figure 5), contributing 60% of the river thermal load during the study period. As evidenced in Figure 4, the river–aquifer exchanges for the simulated river (Strahler orders larger than 3) and overflow (Strahler orders 1 and 2) are the largest contributors to the river heat budget. The incidence of the hydrological regime on the heat balance in the river network is illustrated by the large advective heat flux via baseflow (156.3 TWh per year), which forms the second source of river heat gains. Surface runoff contributes 13% of the total thermal load in the river network. The largest non-advective heat losses in the river network are latent heat and long-wave radiation losses, accounting for 67% (211.0 TWh) and 32% (101.1 TWh), respectively. Approximately 1% of total heat losses are due to sensible heat losses in the river.

#### 4.3. Seasonal heat fluxes in the river network

Seasonal variations of each term of the heat balance in the river network are provided in Table 4. It is important to note that the heat exchanges between river and aquifer is not simulated by the model for the rivers with a Strahler orders lower than 3. This implies that the fluxes are not calculated on the same river extent as the other terms. Short-wave radiation is higher during summer and spring. It depends on the sunlight reaching to stream surface, albedo of water, and shading factor which limits the sunlight. The shading factor can vary as a function of stream

morphology, riparian vegetation, urban structures, or clouds [Loicq *et al.*, 2018]. Although a fixed shading factor was used in this study (30%), more work is needed to implement the shading factor's spatial variability due to the aforementioned factors. The order of non-advective heat sinks does not vary seasonally, as latent heat is the largest heat sink in the system, followed by long-wave radiation, sensible heat and streambed conduction, respectively. Latent heat causes on average a  $120.6 \text{ W}\cdot\text{m}^{-2}$  heat loss during summer and  $20.3 \text{ W}\cdot\text{m}^{-2}$  heat loss during winter. From the simulation, heat fluxes related to conduction in the streambed are negligible when compared to other flux terms, regardless of the season.

Seasonally, heat gains in the Seine river are mainly controlled by the short-wave radiation and river–aquifer exchanges. During winter the net heat gains related to short-wave radiation decrease to only 22% of the heat gains during the summer period (Table 4). River–aquifer exchanges sustain the heat gains of the river network throughout seasons, and become the dominant heat source during winter. Advective heat flux contributions by each aquifer unit change every season due to river and riverbed temperature variations and the magnitude of the river–aquifer exchanges. Studies usually assume that the riverbed conduction contribution is negligible with respect to the advective heat input [Magnusson *et al.*, 2012]. Our study confirms this statement as riverbed conduction reaches up to only 1.5% of river–aquifer advective fluxes.

Loinaz *et al.* [2013] has showed that at the catch-

ment scale, river–aquifer exchanges contribute to a significant fraction of heat sources. Similarly, in the Seine basin, river–aquifer exchanges contribute significantly to the river energy budget. Locally, the river–aquifer exchanges can dominate the heat balance especially during the winter (Table 4). This can create a local cooling effect during summer or a warming effect during winter; an essential aspect of the river–aquifer exchanges that creates thermal refugia for fish species [Kaandorp *et al.*, 2019]. Also, local thermal variations in the river have implications for the industrial use of water and pollution [Prest *et al.*, 2016].

Although a detailed sensitivity analysis was not feasible within the scope of this paper, a similar study at the regional scale in the Loire Basin reported that the dominant factors, in decreasing order of importance, were the shading factor, headwater temperatures, groundwater flow, discharge and river geomorphology, respectively [Beaufort *et al.*, 2016]. Our approach differs from this study, which relied on an equilibrium temperature concept, by the way that the processes are treated explicitly and streambed temperature is estimated by an analytical solution in a distributed manner. We found that the river–atmosphere exchange had the largest impact on the river heat balance at regional scale.

#### 4.4. *Implications on the water–energy–food nexus*

Water, food, and energy are the three essential elements that is needed for human survival and livelihood [De Amorim *et al.*, 2018]. The Water–Energy–Food Nexus (WEFN) has emerged since the early 2010’s as a new approach to collectively manage three finite resources, whose accessibility is increasingly threatened by climate change, population growth, urbanization, and environmental concerns [Hoff, 2011]. The need for a nexus approach has been motivated by the fact that there is a competition for water between water, food and energy production.

Our study shows that the Jurassic aquifer (Figure 5) has a good potential for the geothermal use. This fact is well known, the Dogger aquifer (part of the Jurassic aquifer) is exploited for more than 40 years via low temperature geothermal energy applications [Lopez *et al.*, 2010]. The Paris basin has a great potential for geothermal energy production

as annual thermal energy recycling is substantial (182.5 TWh via recharge and 63.2 TWh via geothermal gradient annually). For comparison, France produced 548.6 TWh electricity in 2020 [NEA, 2020]. Our findings indicate that the geothermal energy potential to offset part of this energy production is substantial. Further research is needed to assess the suitability of aquifer thermal energy systems (ATES) applications at shallower depths [Lee, 2010] in the Seine River basin.

Our coupled ORCHIDEE–CaWaQS model therefore opens the way for new research that integrates the WEFN in the sense that it can be used to assess the impact of energy and agricultural policies at the basin scale in a quantitative way, leveraging on previous studies describing the effect of agricultural practices on the Seine hydro-agrosystem with the ancestor of CaWaQS [Beaudoin *et al.*, 2018].

## 5. Conclusions and perspectives

A first coupling model of water and energy fluxes at the regional scale (96,204 km<sup>2</sup>) was carried out with a focus on the Seine hydrosystem in order to address the hydrological and thermal functioning of the hydrosystem and quantify the interactions between surface and subsurface compartments. The distribution of the thermal loads in the Seine hydrosystem including the aquifer system and the river network at the regional scale is quantified for the first time.

Overall, the coupled framework simulates the mean discharge at various sub-catchment scales in the Seine basin and hydraulic heads in the groundwater system adequately for the purposes of thermal transport simulation. For a first attempt of such a coupling between a SVAT (ORCHIDEE) and a hydrological model (CaWaQS), the platform performs satisfactorily for the simulation of both water fluxes and energy fluxes. The coupled ORCHIDEE–CaWaQS framework is designed to take advantage of broadly available gridded datasets from catchment to regional scales to estimate hydrosystem water and energy budgets. The transport module estimated the mean river temperatures and aquifer temperatures adequately.

The findings presented herein have the potential to inform biological or water quality studies [Flipo *et al.*, 2020] on the Seine basin or engineering applications such as ATES to reduce the dependency



on fossil fuel [Lee, 2010] by providing insights into key physical processes driving the heat fluxes in the river network and the aquifer system. The impact of climate change on the share of energy fluxes at the regional basin scale will also be studied and provide useful information for adaptation and mitigation strategies. The model, for example, can be integrated in an ecohydrological model to estimate the impacts of thermal variability on fish or riparian habitat [Loinaz *et al.*, 2014]. These efforts can have policy implications by guiding environmental management agencies to prioritize their restoration and exploitation areas in the Seine basin. Future coupling with an agronomic model [Beaudoin *et al.*, 2018] will also provide a complete picture within the WEFN framework [Garcia and You, 2016] at the regional scale, thereby improving our understanding of the competing aspects of limited resources.

### Conflicts of interest

The authors declare no competing financial interest.

### Dedication

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. DK: designed the research, developed the CaWaQS transport module, performed all simulations, analyzed the data and prepared the manuscript draft. AR: designed the research, developed CaWaQS transport module, designed the transport through streambed and aquitard interfaces, interpreted results, discussed findings, contributed and improved the manuscript. NG: developed the CaWaQS transport module and the required CaWaQS coupling functionalities, contributed to figures and improved the manuscript. SW: developed the CaWaQS transport module. AD: contributed to the design of the ORCHIDEE application, interpreted ORCHIDEE results and discussed findings. PP: contributed to the design of the ORCHIDEE application, interpreted ORCHIDEE results and discussed findings. NF: designed the research, managed the development of CaWaQS, interpreted results, discussed findings and improved the manuscript, managed and ensured the funding of the project.

### Acknowledgments

This work is a contribution to the PIREN Seine research program, which is part of the Zone Atelier Seine, included in the e-LTER network. We gratefully acknowledge the river temperature dataset provided by William Thomas. Groundwater withdrawals were obtained from Seine-Normandie Water Agency, over the 1994–2012 period. We would like to thank two anonymous reviewers for their input that improved the final article's quality.

### Supplementary data

Supporting information for this article is available on the journal's website under <https://doi.org/10.5802/crgeos.165> or from the author.

### References

- Anderson, M. P. (2005). Heat as a ground water tracer. *Ground Water*, 43, 951–968.
- Ballabio, C., Panagos, P., and Monatanarella, L. (2016). Mapping topsoil physical properties at European scale using the LUCAS database. *Geoderma*, 261, 110–123.
- Bayer, P., Attard, G., Blum, P., and Menberg, K. (2019). The geothermal potential of cities. *Renew. Sustain. Energy Rev.*, 106, 17–30.
- Beaudoin, N., Gallois, N., Viennot, P., Le Bas, C., Puech, T., Schott, C., Buis, S., and Mary, B. (2018). Evaluation of a spatialized agronomic model in predicting yield and N leaching at the scale of the Seine-Normandie basin. *Environ. Sci. Pollut. Res.*, 25(24), 23529–23558.
- Beaufort, A., Curie, F., Moatar, F., Ducharne, A., Melin, E., and Thiery, D. (2016). T-net, a dynamic model for simulating daily stream temperature at the regional scale based on a network topology. *Hydrol. Process.*, 30(13), 2196–2210.
- Bense, V. F. and Kooi, H. (2004). Temporal and spatial variations of shallow subsurface temperature as a record of lateral variations in groundwater flow. *J. Geophys. Res.: Solid Earth*, 109(B4), article no. B04103.
- Berg, A., Findell, K., Lintner, B., Giannini, A., Seneviratne, S. I., Van Den Hurk, B., Lorenz, R., Pitman, A., Hagemann, S., Meier, A., *et al.* (2016).

- Land–atmosphere feedbacks amplify aridity increase over land under global warming. *Nat. Clim. Chang.*, 6(9), 869–874.
- Boogert, A. and Dupont, D. (2005). The nature of supply side effects on electricity prices: the impact of water temperature. *Econ. Lett.*, 88(1), 121–125.
- Böttcher, F. and Zosseder, K. (2022). Thermal influences on groundwater in urban environments—a multivariate statistical analysis of the subsurface heat island effect in Munich. *Sci. Total Environ.*, 810, 152–193.
- Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., et al. (2020). Presentation and evaluation of the ipsl-cm6a-lr climate model. *J. Adv. Model. Earth Syst.*, 12(7), 1–52.
- Bredehoeft, J. D. and Papaopulos, I. S. (1965). Rates of vertical groundwater movement estimated from the earth's thermal profile. *Water Resour. Res.*, 1(2), 325–328.
- Briggs, M. A., Johnson, Z. C., Snyder, C. D., Hitt, N. P., Kurylyk, B. L., Lautz, L., Irvine, D. J., Hurley, S. T., and Lane, J. W. (2018). Inferring watershed hydraulics and cold-water habitat persistence using multi-year air and stream temperature signals. *Sci. Total Environ.*, 636, 1117–1127.
- Briggs, M. A., Lautz, L. K., and Hare, D. K. (2014). Residence time control on hot moments of net nitrate production and uptake in the hyporheic zone. *Hydrol. Process.*, 28(11), 3741–3751.
- Bustillo, V., Moatar, F., Ducharne, A., Thiéry, D., and Poirel, A. (2014). A multimodel comparison for assessing water temperatures under changing climate conditions via the equilibrium temperature concept: case study of the middle Loire river, France. *Hydrol. Process.*, 28(3), 1507–1524.
- Caissie, D. (2006). The thermal regime of rivers: a review. *Freshwater Biol.*, 51(8), 1389–1406.
- Chandesris, A., Mengin, N., Malavoi, J., Souchon, Y., Pella, H., and Wasson, J. (2008). Système relationnel d'audit de l'hydromorphologie des cours d'eau: principes et méthodes. Technical report.
- De Amorim, W. S., Valduga, I. B., Ribeiro, J. M. P., Williamson, V. G., Krauser, G. E., Magtoto, M. K., and de Andrade, J. B. S. O. (2018). The nexus between water, energy, and food in the context of the global risks: an analysis of the interactions between food, water, and energy security. *Environ. Impact Assess. Rev.*, 72, 1–11.
- de Marsily, G. (1986). *Quantitative Hydrogeology: Groundwater Hydrology for Engineers*. Academic Press, Cambridge, MA.
- de Marsily, G. (2009). *L'eau, un trésor en partage*. Dunod, Malakoff.
- de Marsily, G. (2020). Will we soon run out of water? *Ann. Nutr. Metab.*, 76(1), 10–16.
- de Marsily, G., Ledoux, E., Levassor, A., Poitrinal, D., and Salem, A. (1978). Modelling of large multi-layered aquifer systems: theory and applications. *J. Hydrol.*, 36, 1–34.
- De Rosnay, P., Bruen, M., and Polcher, J. (2000). Sensitivity of surface fluxes to the number of layers in the soil model used in gcms. *Geophys. Res. Lett.*, 27(20), 3329–3332.
- Dentzer, J. (2016). *Forçages environnementaux et contrôles structuraux sur le régime thermique actuel du bassin de Paris : enjeux pour la compréhension du potentiel géothermique en Ile-de-France*. PhD thesis, Université Pierre et Marie Curie, Paris VI.
- d'Orgeval, T., Polcher, J., and De Rosnay, P. (2008). Sensitivity of the west african hydrological cycle in orchidee to infiltration processes. *Hydrol. Earth Syst. Sci.*, 12(6), 1387–1401.
- Ducharne, A. (2008). Importance of stream temperature to climate change impact on water quality. *Hydrol. Earth Syst. Sci.*, 12(3), 797–810.
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E. (2016). Overview of the coupled model intercomparison project phase 6 (cmip6) experimental design and organization. *Geosci. Model Dev.*, 9(5), 1937–1958.
- Ferguson, G. and Woodbury, A. D. (2007). Urban heat island in the subsurface. *Geophys. Res. Lett.*, 34(23), 1–4.
- Flipo, N., Even, S., Poulin, M., Théry, S., and Ledoux, E. (2007). Modelling nitrate fluxes at the catchment scale using the integrated tool CAWAQS. *Sci. Total Environ.*, 375, 69–79.
- Flipo, N., Gallois, N., and Schuite, J. (2022). Regional coupled surface-subsurface hydrological model fitting based on a spatially distributed minimalist reduction of frequency-domain discharge data. *Geosci. Model Dev.*, 2022, 1–40. <https://gmd.copernicus.org/preprints/gmd-2022-24/>.
- Flipo, N., Lestel, L., Labadie, P., Meybeck, M., and Garnier, J. (2020). The Seine river basin. In *Handbook of Environmental Chemistry*, volume 90 of

- Trajectories of the Seine River Basin*, pages 1–28. Springer, Cham, Switzerland.
- Flipo, N., Monteil, C., Poulin, M., de Fouquet, C., and Krimissa, M. (2012). Hybrid fitting of a hydrosystem model: long-term insight into the Beauce aquifer functioning (France). *Water Resour. Res.*, 48(5), article no. W0550.
- Flipo, N., Mouhri, A., Labarthe, B., Biancamaria, S., Rivi re, A., and Weill, P. (2014). Continental hydrosystem modelling: the concept of nested stream-aquifer interfaces. *Hydrol. Earth Syst. Sci.*, 18, 3121–3149.
- Garcia, D. J. and You, F. (2016). The water–energy–food nexus and process systems engineering: a new focus. *Comput. Chem. Eng.*, 91, 49–67.
- Goto, S., Yamano, M., and Kinoshita, M. (2005). Thermal response of sediment with vertical fluid flow to periodic temperature variation at the surface. *J. Geophys. Res.: Solid Earth*, 110(B1), article no. B01106.
- Gudmundsson, L., Boulange, J., Do, H. X., Gosling, S. N., Grillakis, M. G., Koutroulis, A. G., Leonard, M., Liu, J., Schmied, H. M., Papadimitriou, L., Pokhrel, Y., Seneviratne, S. I., Satoh, Y., Thiery, W., Westra, S., Zhang, X., and Zhao, F. (2021). Globally observed trends in mean and extreme river flow attributed to climate change. *Science*, 371(6534), 1159–1162.
- Hemmerle, H., Hale, S., Dressel, I., Benz, S. A., Attard, G., Blum, P., and Bayer, P. (2019). Estimation of groundwater temperatures in Paris, France. *Geofluids*, 2019, article no. 5246307.
- Hengl, T., Jesus, J. M. d., Heuvelink, G. B. M., Gonzalez, M. R., Kilibarda, M., Blagotić, A., Shangquan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S., and Kempen, B. (2017). SoilGrids250m: global gridded soil information based on machine learning. *PLoS One*, 12(2), article no. e0169748.
- Hoff, H. (2011). Policy Commons [Internet]. Understanding the nexus. [cited 2022 September 28]. Available from: <https://policycommons.net/artifacts/1359033/understanding-the-nexus/1972269/>.
- Isaak, D. J., Luce, C. H., Horan, D. L., Chandler, G. L., Wollrab, S. P., Dubois, W. B., and Nagel, D. E. (2020). Thermal regimes of perennial rivers and streams in the western United States. *J. Am. Water Resour. Assoc.*, 56(5), 842–867.
- Jolivet, C., Boulonne, L., and Rati , C. (2006). *Manuel du R seau de Mesures de la Qualit  des Sols*. Unit  InfoSol, INRA, Orl ans, France, <https://hal.inrae.fr/hal-02818011>. Accessed on 03.10.2021.
- Kaandorp, V. P., Doornenbal, P. J., Kooi, H., Peter Broers, H., and de Louw, P. G. (2019). Temperature buffering by groundwater in ecologically valuable lowland streams under current and future climate conditions. *J. Hydrol.*, 3, article no. 100031.
- Kilic, D., Rivi re, A., Wang, S., Gallois, N., and Flipo, N. (2021). D veloppement du transport de chaleur au sein de la plateforme de mod lisation des hydrosyst mes CaWaQS. PIREN-Seine. <https://doi.org/10.26047/PIREN.RAPPANN.2021.VOL13>.
- Krinner, G., Viovy, N., de Noblet-Ducoudr , N., Og e, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C. (2005). A dynamic global vegetation model for studies of the coupled atmosphere–biosphere system. *Glob. Biogeochem. Cycles*, 19(1), article no. GB1015.
- Kurylyk, B. L., Irvine, D. J., Carey, S. K., Briggs, M. A., Werkema, D. D., and Bonham, M. (2017). Heat as a groundwater tracer in shallow and deep heterogeneous media: analytical solution, spreadsheet tool, and field applications. *Hydrol. Process.*, 31(14), 2648–2661.
- Ledoux, E., Girard, G., de Marsily, G., Villeneuve, J. P., and Deschenes, J. (1989). *Spatially Distributed Modeling: Conceptual Approach, Coupling Surface Water And Groundwater*. Springer Netherlands, Dordrecht.
- Lee, K. S. (2010). A review on concepts, applications, and models of aquifer thermal energy storage systems. *Energies*, 3(6), 1320–1334.
- Loicq, P., Moatar, F., Jullian, Y., Dugdale, S. J., and Hannah, D. M. (2018). Improving representation of riparian vegetation shading in a regional stream temperature model using lidar data. *Sci. Total Environ.*, 624, 480–490.
- Loinaz, M. C., Davidsen, H. K., Butts, M., and Bauer-Gottwein, P. (2013). Integrated flow and temperature modeling at the catchment scale. *J. Hydrol.*, 495, 238–251.
- Loinaz, M. C., Gross, D., Unnasch, R., Butts, M., and Bauer-Gottwein, P. (2014). Modeling ecohydrological impacts of land management and water use in the silver creek basin, Idaho. *J. Geophys. Res.: Bio-*

- geosci.*, 119(3), 487–507.
- Lopez, S., Hamm, V., Le Brun, M., Schaper, L., Boissier, F., Cotiche, C., and Giuglaris, E. (2010). 40 years of dogger aquifer management in ile-de-france, Paris basin, France. *Geothermics*, 39(4), 339–356.
- Lurton, T., Balkanski, Y., Bastrikov, V., Bekki, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Contoux, C., Cozic, A., Cozic, A., Cugnet, D., Dufresne, J.-L., éthé, C., Foujols, M.-A., Ghattas, J., Hauglustaine, D., Hu, R.-M., Kageyama, M., Khodri, M., Lebas, N., Levvasseur, G., Marchand, M., Ottlé, C., Peylin, P., Sima, A., Szopa, S., Thiéblemont, R., Vuichard, N., and Boucher, O. (2020). Implementation of the cmip6 forcing data in the ipsl-cm6a-lr model. *J. Adv. Model. Earth Syst.*, 12(4), article no. e2019MS001940.
- Magnusson, J., Jonas, T., and Kirchner, J. W. (2012). Temperature dynamics of a proglacial stream: identifying dominant energy balance components and inferring spatially integrated hydraulic geometry. *Water Resour. Res.*, 48(6), article no. W06510.
- Michel, A., Brauchli, T., Lehning, M., Schaepli, B., and Huwald, H. (2020). Stream temperature and discharge evolution in Switzerland over the last 50 years: annual and seasonal behaviour. *Hydrol. Earth Syst. Sci.*, 24(1), 115–142.
- Michel, A., Schaepli, B., Wever, N., Zekollari, H., Lehning, M., and Huwald, H. (2022). Future water temperature of rivers in switzerland under climate change investigated with physics-based models. *Hydrol. Earth Syst. Sci.*, 26(4), 1063–1087.
- Moatar, F. and Gailhard, J. (2006). Water temperature behaviour in the river Loire since 1976 and 1881. *C. R. Geosci.*, 338(5), 319–328.
- Monteiro, L., Figueiredo, D., Covas, D., and Menaia, J. (2017). Integrating water temperature in chlorine decay modelling: a case study. *Urban Water J.*, 14(10), 1097–1101.
- NEA (2020). OECD iLibrary [Internet]. Nuclear Energy Data 2019. [cited 2022 September 28]. Available from: <https://www.oecd-ilibrary.org/content/publication/1786b86b-en-fr>.
- Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P., Kuppel, S., Koffi, E., Kane, A., Maignan, F., Chevallier, F., Ciais, P., and Prunet, P. (2016). A new stepwise carbon cycle data assimilation system using multiple data streams to constrain the simulated land surface carbon cycle. *Geosci. Model Dev.*, 9(9), 3321–3346.
- Prest, E. I., Hammes, F., van Loosdrecht, M. C. M., and Vrouwenvelder, J. S. (2016). Biological stability of drinking water: controlling factors, methods, and challenges. *Front. Microbiol.*, 7, article no. 45.
- Qiu, H., Blaen, P., Comer-Warner, S., Hannah, D. M., Krause, S., and Phanikumar, M. S. (2019). Evaluating a coupled phenology-surface energy balance model to understand stream-subsurface temperature dynamics in a mixed-use farmland catchment. *Water Resour. Res.*, 55(2), 1675–1697.
- Raoult, N., Ottlé, C., Peylin, P., Bastrikov, V., and Maugis, P. (2021). Evaluating and optimizing surface soil moisture drydowns in the orchidee land surface model at in situ locations. *J. Hydrometeorol.*, 22(4), 1025–1043.
- Reynolds, C. A., Jackson, T. J., and Rawls, W. J. (2000). Estimating soil water-holding capacities by linking the Food and Agriculture Organization Soil map of the world with global pedon databases and continuous pedotransfer functions. *Water Resour. Res.*, 36(12), 3653–3662.
- Rivière, A. (2012). *Influences du pergélisol sur les écoulements souterrains et leurs échanges avec une rivière : étude numérique et expérimentale*. PhD thesis, UPMC, Paris, <https://www.theses.fr/2012PA066457>. Accessed on 09/10/2021.
- Rivière, A., Ladet, D., Thomas, W., Le Breton, G., Ducharne, A., Oudin, L., and de France, D. I. (2021). *Projections des températures de l'eau de la seine à paris à l'horizon*. PIREN Seine. <https://hal-mines-paristech.archives-ouvertes.fr/hal-03533469>. Accessed on 21/02/2022.
- Saleh, F., Flipo, N., Habets, F., Ducharne, A., Oudin, L., Viennot, P., and Ledoux, E. (2011). Modeling the impact of in-stream water level fluctuations on stream-aquifer interactions at the regional scale. *J. Hydrol.*, 400(3–4), 490–500.
- Sauer, V. B. and Meyer, R. (1992). Determination of error in individual discharge measurements. Technical report, US Geological Survey; Books and Open-File Reports Section [distributor]. <https://pubs.er.usgs.gov/publication/ofr92144>. Accessed on 04/03/2022.
- Seyedhashemi, H., Vidal, J.-P., Diamond, J. S., Thiéry, D., Monteil, C., Hendrickx, F., Maire, A., and Moatar, F. (2022). Regional, multi-decadal analysis on the Loire river basin reveals that stream temperature increases faster than air temperature. *Hydrol.*

- Earth Syst. Sci.*, 26(9), 2583–2603.
- Singh, A. (2014). Groundwater resources management through the applications of simulation modeling: a review. *Sci. Total Environ.*, 499, 414–423.
- Tafasca, S., Ducharne, A., and Valentin, C. (2020). Weak sensitivity of the terrestrial water budget to global soil texture maps in the ORCHIDEE land surface model. *Hydrol. Earth Syst. Sci.*, 24, 3753–3774.
- Taylor, R. G., Scanlon, B., Doll, P., Rodell, M., van Beek, R., Wada, Y., Longuevergne, L., Leblanc, M., Famiglietti, J. S., Konikow, L., Green, T. R., Chen, J., Taniguchi, M., Bierkens, M. F. P., MacDonald, A., Fan, Y., Maxwell, R. M., Yechieli, Y., Gurdak, J. J., Allen, D. M., Shamsudduha, M., Hiscock, K., Holman, I., and Treidel, H. (2013). Ground water and climate change. *Nat. Clim. Change*, 3, 322–329.
- Van Vliet, M., Yearsley, J., Franssen, W., Ludwig, F., Haddeland, I., Lettenmaier, D., and Kabat, P. (2012). Coupled daily streamflow and water temperature modelling in large river basins. *Hydrol. Earth Syst. Sci.*, 16(11), 4303–4321.
- van Vliet, M. T., Franssen, W. H., Yearsley, J. R., Ludwig, F., Haddeland, I., Lettenmaier, D. P., and Kabat, P. (2013). Global river discharge and water temperature under climate change. *Glob. Environ. Change*, 23(2), 450–464.
- Vidal, J.-P., Martin, E., Franchistéguy, L., Baillon, M., and Soubeyroux, J.-M. (2010). A 50-year high-resolution atmospheric reanalysis over France with the Safran system. *Int. J. Climatol.*, 30(11), 1627–1644.
- Wagener, T., Sivapalan, M., Troch, P. A., McGlynn, B. L., Harman, C. J., Gupta, H. V., Kumar, P., Rao, P. S. C., Basu, N. B., and Wilson, J. S. (2010). The future of hydrology: an evolving science for a changing world. *Water Resour. Res.*, 46(5), article no. W05301.
- Wang, F., Cheruy, F., and Dufresne, J.-L. (2016). The improvement of soil thermodynamics and its effects on land surface meteorology in the ipsl climate model. *Geosci. Model Dev.*, 9(1), 363–381.
- Wang, S., Flipo, N., Bordet, A., Bernier, J., Merchat, M., and Rocher, V. (2021). Transport de chaleur par la seine à la traversée de l'agglomération parisienne. Rapport d'étude, PIREN Seine. [https://www.piren-seine.fr/rapports/rapports\\_annuels/rapports\\_dactivite\\_2021](https://www.piren-seine.fr/rapports/rapports_annuels/rapports_dactivite_2021). Last access: 08-01-2022.
- Wang, T., Oettle, C., Boone, A., Ciais, P., Brun, E., Morin, S., Krinner, G., Piao, S., and Peng, S. (2013). Evaluation of an improved intermediate complexity snow scheme in the orchidee land surface model. *J. Geophys. Res.: Atmos.*, 118(12), 6064–6079.
- Webb, B., Hannah, D., Moore, R., Brown, L., and Nobilis, F. (2008). Recent advances in stream and river temperature research. *Hydrol. Process.*, 22, 902–918.
- Webb, B. W. and Nobilis, F. (2007). Long-term changes in river temperature and the influence of climatic and hydrological factors. *Hydrol. Sci. J.*, 52(1), 74–85.
- Wondzell, S. M., Diabat, M., and Haggerty, R. (2019). What matters most: are future stream temperatures more sensitive to changing air temperatures, discharge, or riparian vegetation? *J. Am. Water Resour. Assoc.*, 55(1), 116–132.
- York, J. P., Person, M., Gutowski, W. J., and Winter, T. C. (2002). Putting aquifers into atmospheric simulation models: an example from the mill creek watershed, northeastern Kansas. *Adv. Water Resour.*, 25(2), 221–238.
- Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., Samanta, A., Piao, S., Nemani, R. R., and Myneni, R. B. (2013). Global data sets of vegetation leaf area index (lai) 3g and fraction of photosynthetically active radiation (fpar) 3g derived from global inventory modeling and mapping studies (gimms) normalized difference vegetation index (ndvi3g) for the period 1981 to 2011. *Remote Sens.*, 5(2), 927–948.





Research article

Geo-hydrological Data & Models

# Land drainage functioning and hydrological impacts in rural catchments: model development and field experiments

Daniel Zimmer<sup>Ⓢ, a</sup>, Julien Tournebize<sup>Ⓢ, \*, b</sup>, Sami Bouarfa<sup>Ⓢ, c</sup>, Cyril Kao <sup>d</sup> and Benoît Lesaffre <sup>e</sup>

<sup>a</sup> EIT Climate-KIC, France

<sup>b</sup> Université Paris-Saclay, INRAE – UR HYCAR, France

<sup>c</sup> Université de Montpellier, INRAE – UMR GEAU, France

<sup>d</sup> INRAE – Department for Academic Partnerships, Regional and European Affairs, France

<sup>e</sup> Member, French Academy of Agriculture, France

*E-mails:* daniel.zimmer@climate-kic.org (D. Zimmer), julien.tournebize@inrae.fr (J. Tournebize), sami.bouarfa@inrae.fr (S. Bouarfa), cyril.kao@inrae.fr (C. Kao), benoitlesaffre@sfr.fr (B. Lesaffre)

**Abstract.** The development of an integrated theory of subsurface drainage based on hydrology and hydrogeology concepts is presented. The historical context, the main hypothesis derived from the Boussinesq equation and the validation of the model predictions are discussed. Theoretical developments of this equation demonstrate that a single parameter ( $\sigma$ )—a combination of soil and drainage system properties—is sufficient for predicting the dynamics of subsurface drain flow rates. We also demonstrate that these drain flow rates are a function of the level of water replenishment in the system (classically the water table elevation), of the recharge intensity of the aquifer and of a buffer function related to the swelling or deflation of the water table shape during recharge events. For values of  $\sigma > 1$ , the buffer role of the water table is negligible. In that case approx. 13% of the water table recharge contributes to the flow rate, which is shown to explain the observed disconnection between water table elevations and peak flow rates at the outlet of classic agricultural drainage systems and to predict these peak flow rates accurately. A modelling approach based on this theory and validated experimentally (SIDRA model) allowed us to test the quality of the peak flow prediction. The SIDRA model also includes a surface runoff module and has been coupled to different modelling tools and used to analyse the impacts of subsurface drainage on water quality. The approach contributed towards the development of tools that helped to connect better the drainage systems to the hydrological functioning of watersheds.

**Keywords.** Subsurface drainage, Waterlogged soil, Modelling, Boussinesq equation, Peak flow, Surface runoff, Water quality.

**Quotation.** “Somewhere in the scientific community, over and above the modelling work guided by « useful » aims, work must also continue on models that serve no particular purpose” [de Marsily, 1994].  
*Manuscript received 19 September 2022, revised 9 November 2022, accepted 28 November 2022.*

## 1. Introduction

Since the late 1980s, research has been conducted in France to understand the influence of water man-

\* Corresponding author.

agement infrastructure, such as arterial and subsurface drainage in rural catchments, on floods and on surface water quality. This became an important concern during a period (1980–1990) in which a national program of subsurface drainage investments was supported by the French government [Lesaffre and Penel, 1990]. During this period, pollution of surface and groundwater progressively became an issue, linked to the intensification of agriculture, to which drainage contributed.

In France, many rural areas face seasonal waterlogging in the humid winter period [Lagacherie and Favrot, 1987]. This is particularly the case in shallow loamy soils where an impervious barrier has developed due to clay leaching and enrichment in the deeper layers. Waterlogging is detrimental to root growth and reduces crop yield, particularly of winter cereals. To increase the flexibility of their crop rotations, many farmers were at that time eager to implement subsurface drainage in response to the competitive and volatile market context in Europe due to intermittent surpluses of various commodities (butter, milk, meat, etc.) that impacted farmer revenues.

The combination of subsurface drainage installation (plastic pipes at 10–25-m spacing and an average depth of 1 m) and of arterial drainage recalibration (often following land consolidation measures) has had an obvious influence on the transfer of water from the field to the rivers: more rainfall water flows through the soil into the pipe network, and less through surface runoff. The transfer from arterial drainage channels to the rivers may also be accelerated [Lesaffre and Zimmer, 1988, and more recently King *et al.*, 2014 or Valayamkunnath *et al.*, 2022]. But how important are these changes, particularly in periods of floods or for the quality of the drainage water?

At that time, the scientific approach to subsurface drainage was based on a hydraulic method relating drain flow rates to the elevation of the water table mid-way between drains [Van Schilfhaarde, 1963, Dumm, 1964, Guyon, 1964]. The saturated zone of the soil only was taken into consideration and two extreme conditions were classically considered: (i) the tail recession phase, after a rainfall event when a univocal relationship between water elevation and drain flow rate was observed; and (ii) the steady-state situation, when the rainfall recharge of the water table is equal to the drain flow rate. However, this ap-

proach was unable to predict the dynamic process of transient rainfall events inducing a non-univocal relationship between water table elevation and drain flow rate [Hervé, 1980]. In particular, it was unable to predict the peaky response of drainage systems during rainfall events that were a growing concern regarding the influence of subsurface drainage on floods. As observed in several field experiments, the drain flow rates during rainfall events could be 3–10 times higher than predicted by classic equations linking drain flow rate to water table elevation [Lesaffre, 1989].

Since the late 1930s [Flodkvist in Russel, 1934], this observation has resulted in a debate about the water flow pattern generating these peak flows. If there is no simple correlation between water table elevation and drain flow rate, could it be that other mechanisms are at play? The debate continues today and several authors [Schneider *et al.*, 2022, Bjerre *et al.*, 2022, Williams and McAfee, 2021] argue that surface runoff or water flowing in the plough layer could infiltrate through macropores into the drainage trench and contribute to these peak flows.

The approach implemented in France combined detailed field experiments [Zimmer, 1989] with revisiting of the theoretical approach to subsurface drainage functioning based on the Boussinesq equation [Lesaffre, 1989]. This resulted in the development of the basic blocks of a drainage model (SIDRA), as explained below. This model was validated against experimental results and was then combined with a number of other processes to assess various impacts of the changing agricultural water management on solute and particle transport, on plant growth and development and on the hydrology of wetlands and rural catchments.

This paper follows the rationale of the scientific developments during the period 1986–2020, which can be presented as a general theory of land drainage systems translated into the SIDRA model development. It summarizes the initial results based on experimental observations and on an extension of the classic Boussinesq equation. It then describes in a synthetic way some of the applications obtained by the coupling of the SIDRA model with other tools.



## 2. Initial experimental results

Among the experimental sites used to analyse the field drainage design and its impacts on winter crop yields, the experimental site of Arrou (France, Eure-et-Loir) was intensively studied for hydrological purposes from 1978 to 1995 [Lesaffre and Zimmer, 1988, Tournebize *et al.*, 2004]. In particular, water table levels and drain flow rates were monitored at an hourly time step throughout this period. During the winter season of 1985–1986 monitoring was completed by tensiometers installed at different distances (0, 0.5, 1 and 5 m, i.e. until mid-drain spacing) from a drain pipe and between the soil surface and a depth of 1 m [Zimmer, 1989]. The purpose was to measure precisely the water potential inside the water table and in the unsaturated zone above it, as well as to analyse the water transfer patterns, with a particular focus on rainfall events.

The following observations were critical for the theoretical developments presented hereafter. During rainfall events, a unit hydraulic gradient is observed in the unsaturated zone of the soil as soon as a threshold pore water pressure of approx.  $-15$  hPa is reached, meaning that for pore water pressures in the range of  $(-15, 0)$  hPa, the flow is purely vertical, equal to the soil hydraulic conductivity, and driven by gravity. Detailed observations reveal that the pore water pressure in the unsaturated zone during rainfall events is not correlated with the rainfall intensity in a simple way. Lower pressure values are observed for a similar rainfall intensity when the event occurs after previous rainfall events. These empirical findings were theoretically studied and generalized by Kao *et al.* [2001] and confirmed through a laboratory experiment by Kao [2002].

Furthermore, as soon as the soil becomes saturated (water pressure becoming positive), the vertical hydraulic gradients decrease to 0, i.e. the water flow inside the water table is horizontal. This behaviour is well known as the Dupuit–Forchheimer (DF) assumption and is largely referred to in hydrogeology since it enables a simplification of the solution to the Boussinesq equation.

Similar observations were obtained in other field experiments, except in some heavy clay soils [Zimmer, 1989]. The interpretation of these observations is that, except in these soils, peak flows can be predicted by the water table behaviour only, without

specific horizontal water transfer disconnected from it. If the macroporosity of the soil played a role, it would seem to be through the establishment of a downward gravity flow (vertical gradient equal to 1) as soon as the pore water pressure reaches a threshold of approx.  $-15$ – $-20$  hPa. These findings were also recently reported in other contexts [Jacobsen and Kjær, 2007].

## 3. Model development

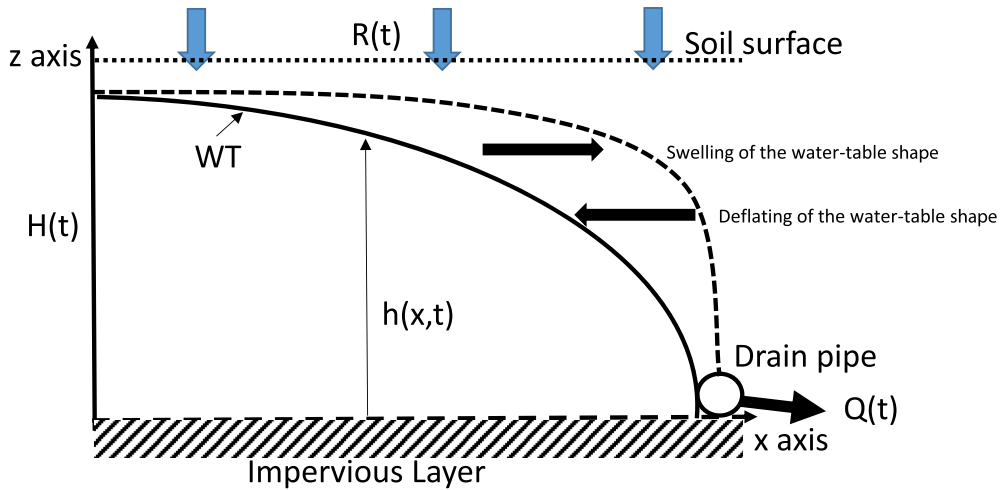
### 3.1. *The Boussinesq approach, associated assumptions and solutions developed*

Based on these observations, the classic Boussinesq approach [Boussinesq, 1904] that describes water transfers generated by a drained water table was revisited. In this approach, and for the specific case of shallow soils where the drain pipe rests on an impervious barrier, the system is represented (Figure 1) by a transect section of the soil bounded at the bottom by a shallow impervious barrier. A drain that can be a pipe or an open ditch imposes atmospheric pressure at the depth of the impervious barrier. At the soil surface, assumed to be horizontal, a transient rainfall intensity is applied and is assumed to be transferred in a homogeneous way to the water table in the case that the water table remains below soil surface. When the water table reaches the soil surface, a certain fraction of the rainfall generates surface runoff. Finally, at a certain distance to the drain, generally midway between drains, it is assumed that no lateral flow occurs, dividing the pipe space into a symmetric system. This latter assumption can easily be changed.

The Boussinesq approach rests on the combination of two basic equations. First, a motion equation (Equation (1)) describes the horizontal flow in the water table at a given abscissa  $x$ ,  $q(x, t)$ ; this equation results from Darcy's law integrated in the saturated zone between the drain level and water table surface. Second, a continuity equation (Equation (2)) describes the water balance, i.e. the variation of that horizontal flow when the water table elevation  $h(x, t)$  changes.

$$q(x, t) = \int_0^{h(x,t)} K_h(z) \frac{\partial \varphi(x, z, t)}{\partial x} dz \quad (1)$$

$$\frac{\partial q(x, t)}{\partial x} = R(t) - f \frac{\partial h(x, t)}{\partial t} \quad (2)$$



**Figure 1.** Sketch of a drained transept from pipe to mid-drain spacing with  $Q(t)$ , drain discharge,  $R(t)$  recharge generated by precipitations,  $h(x, t)$  and  $H(t)$  water table elevation above the impervious layer and WT water table surface.

where  $K_h(z)$  is the soil horizontal saturated conductivity ( $L \cdot T^{-1}$ ),  $\varphi(x, z, t)$  is the total hydraulic head inside the water table (L),  $f$  is drainable porosity (-), and  $R(t)$  ( $L \cdot T^{-1}$ ) is the recharge rate of the water table assumed to be independent of abscissa  $x$ .

Two important assumptions are associated with the Boussinesq approach.

The drainable porosity parameter ( $f$ ) has been the subject of considerable debate [see, e.g. in Vachaud, 1968]. Drainable porosity also called “specific water yield” corresponds to the amount of water which could be drained from the soil [Moriassi *et al.*, 2013]. The value depends on the soil–water dynamics and on the water table depth. The associated assumption is that this parameter is independent of the status (pressure, water content) of the unsaturated zone above the water table and that a similar value applies across the system and during recharge and drawdown of the water table. This assumption is a simplification of the reality, as it can be derived from more complete models such as those provided by the Richards equation including the concepts of dual porosity or dual permeability [Vauclin *et al.*, 1976, Akay *et al.*, 2008, Varvaris *et al.*, 2021]. But our experimental data highlight the importance of macropores in which gravity water transfers occur during rainfall events. We therefore assumed that these macropores played a role during both recharge and drawdown of

the water table and that the drainable porosity value was identical in these two phases.

The second important assumption is the Dupuit–Forchheimer (DF) assumption introduced earlier and derived from our field observations. This allows us to replace the total hydraulic head  $\varphi(x, z, t)$  by the water table elevation  $h(x, t)$  in Equation (1), which simplifies its analytical integration. As shown below, the assumption is only required at mid-drain spacing (L), and the errors associated with this assumption can be assessed and corrected [Lesaffre, 1989].

A third important assumption, introduced by Guyon [1964] on the basis of Boussinesq’s theoretical work, is that water table shapes during the recession and steady-state phases are very similar, so that the water table shape during the transient phase can be well approximated by the one under steady recharge. This leads to a relatively simple calculation of water table shapes.

The integration of Equations (1) and (2) was initially performed during tail recession, i.e. when the recharge rate of the water table  $R(t) = 0$ . The approach was later extended to the case of a positive recharge rate  $R(t)$  [Lesaffre, 1989]. The approach rests on two separate integrations of Equations (1) and (2). Equation (1) can be integrated against  $x$  using the discharge potential function introduced by Tcharny [1951], who noted that  $q(x, t)$  can be

expressed as:

$$q(x, t) = -\frac{\partial F(x, t)}{\partial x}, \quad (3)$$

where

$$F(x, t) = \int_0^{h(x, t)} K_h(z) [\varphi(x, z, t) - z] dz. \quad (4)$$

This yields:

$$\int_0^L q(x, t) dx = F(0, t) - F(L, t). \quad (5)$$

Calculating  $F$  in 0 and in  $L$  is then the next step. Assuming that the drain pipe is not surcharged yields  $F(L) = 0$ . Calculating  $F(0)$  is also easy using the  $DF$  assumption, which is only used in the case of  $x = 0$ , and assuming  $K_h(z)$  is constant (case of homogeneous soils). This directly yields:

$$F(0, t) = K_h \frac{H(t)^2}{2} \quad (6)$$

where  $H(t)$  is the water table elevation above the impervious layer at mid-drain spacing.

The approach can be extended to the general case of deep impervious barriers and to the case of heterogeneous soils.  $F(0, t) - F(L, t)$  can be written in this general case as:

$$F(0, t) - F(L, t) = J(t) * L^2/2 \quad (7)$$

where

$$J(H, t) = \frac{\tilde{K}H(t)^2 + 2\bar{K}d'H(t)}{L^2} \quad (8)$$

$J(H, t)$  is the classic Hooghoudt equation in homogeneous soils representing the steady-state drain flow rate [Hooghoudt, 1940]. In this equation,  $\tilde{K}$  is the equivalent horizontal conductivity as defined by Wolsack [1978], and  $\bar{K}d'$  represents the equivalent transmissivity of the layers situated below the drain level [see details in Lesaffre, 1989].

The next step is to integrate Equation (2) twice against  $x$  to obtain a second expression of  $\int_0^L q(x, t) dx$ . This is achieved through assumptions on the water table shape. Guyon [1964] initially assumed that  $h(x, t)$  can be written as  $W(X) \cdot H(t)$  (separation of variables), where  $W(X)$  represents the non-dimensional water table elevation ( $h(x, t)/H(t)$ ) at abscissa  $X = x/L$ .

Recognizing that the water table shape  $W(X)$  could change over time, particularly during recharge events (as occasionally observed in the field), Lesaffre [1989] extended the approach to the case of a time-dependent water table shape, i.e. where  $h(x, t)$  can

be written as  $W(X, t) \cdot H(t)$ , which was further developed by Bouarfa and Zimmer [2000].

This successively yields three equations:

$$\begin{aligned} \frac{\partial q(x, t)}{\partial x} &= R(t) - fW(X, t) \frac{dH(t)}{dt} \\ &\quad - fH(t) \frac{\partial W(X, t)}{\partial t} \end{aligned} \quad (9)$$

$$\begin{aligned} q(L, t)/L &= Q(t) = R(t) - fB(t) \frac{dH(t)}{dt} \\ &\quad - fH \frac{dB(t)}{dt}, \end{aligned} \quad (10)$$

obtained by integration of Equation (9) and replacement of  $x$  by  $L$ .

$$\begin{aligned} \int_0^L q(x, t) dx &= \frac{L^2}{2} \left[ R(t) - fC(t) \frac{dH(t)}{dt} - fH(t) \frac{dC(t)}{dt} \right] \end{aligned} \quad (11)$$

where  $B(t)$  and  $C(t)$  are water table shape coefficients obtained by the simple and double spatial integration of the non-dimensional water table shape  $W(X, t)$  against  $X$  [Lesaffre, 1989].

Combining Equations (5), (7) and (11) finally yields:

$$\begin{aligned} Q(t) = \frac{q(L, t)}{L} &= A(t)J(H, t) + (1 - A(t))R(t) \\ &\quad - fC(t) \frac{dA(t)}{dt} H(t) \end{aligned} \quad (12)$$

where  $B(t)$ ,  $C(t)$  and  $A(t) = B(t)/C(t)$  are water table coefficients derived from the integration of the non-dimensional water table shape:

$$\begin{aligned} B(t) &= \int_0^1 W(X, t) dX \quad \text{and} \\ C(t) &= 2 \int_0^1 dX \int_0^X W(X', t) dX'. \end{aligned} \quad (13)$$

When the pipe rests on the barrier in homogeneous soils, coefficients  $B$  and  $C$  are, respectively, close to 0.78 and 0.90 during tail recession.

The drain flow described by Equation (12) has three components:

- the first term on the right-hand side of the equation represents a water table contribution that is the classic expression of the flow during tail recession;
- the second term represents a proportion of the recharge rate reaching the water table which contributes directly to the drain flow; this proportion is determined by the water table coefficient  $A$ , of the order of 0.87 when the pipe rests on the barrier in homogeneous

soils, meaning that this proportion is of the order of 13% of the recharge rate;

- the third term is the contribution of water shape changes that occur mostly during recharge events; this contribution is due to the swelling or deflation of the water table (Figure 1); at the beginning of a rainfall event, coefficient  $A$  increases; the swelling of the water table shape mitigates the contribution of the recharge rate, thus creating a buffer.

The second and third terms are called “peak flow term” and “water table deformation”, respectively, by Lesaffre [1989]. According to Bouarfa and Zimmer [2000], these terms provide a physical explanation to the relative disconnection between drain flow and water table elevation during peak flow events and a possible answer to the debated issue of the origin of the peak flows observed during rainfall events. These terms also contribute towards explaining the preferential flow, as described by several authors [Akay *et al.*, 2008].

### 3.2. The critical driver of subsurface drain flows

Investigations on the sensitivity of the system behaviour to the different parameters showed that factors of water table shape are the most sensitive parameters [Zimmer *et al.*, 1995] and demonstrated that parameters  $K$  and  $f$  were interdependent for the prediction of drain flow rates [Favier *et al.*, 1990], i.e. that identical drain flow rates could be generated by several couples of  $K$  and  $f$  values. This observation led us to revisit the Boussinesq equation by introducing the variable  $g(X, t) = f \cdot h(X, t)$ , which represents the drainable water content stored inside the saturated zone of the soil at a given abscissa  $x$ . Combining Equations (1) and (2) and introducing the non-dimensional abscissa  $X = x/L$  yields the following in homogeneous soils:

$$\frac{\partial g(X, t)}{\partial t} = \frac{K}{f^2 L^2} \times \left[ \left( \frac{\partial g(X, t)}{\partial X} \right)^2 + g(X, t) \frac{\partial^2 g(X, t)}{\partial X^2} \right]. \quad (14)$$

This equation shows that  $g(X, t)$  can be fully predicted by a parameter combining the key features of the drainage system, named  $\sigma$ , and defined as:

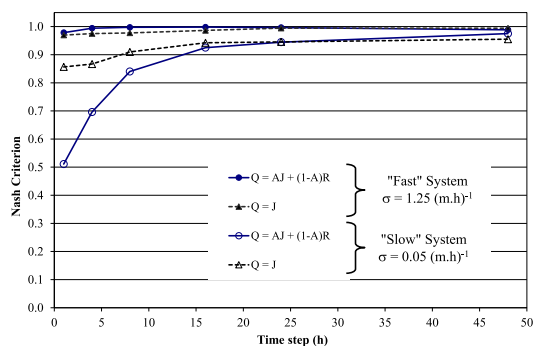
$$\sigma = \frac{K}{f^2 L^2}. \quad (15)$$

The prediction of  $g(x, t)$  suffices to fully predict the drain flow rates in shallow soils, as can be deduced by the continuity equation (Equation (2)), which implies that drain flow rates depend only on parameter  $\sigma$ . Another consequence is that  $\sigma$  also determines the dynamics of swelling and deflating of the water table shape  $W(X)$  and, in turn, the respective parts of the three terms in Equation (12), as shown by Bouarfa and Zimmer [2000]. The role and importance of the third term of Equation (12) in the prediction of drain flow rates are thus strongly influenced by this parameter.

Solving Equation (14) by a finite element scheme allowed us to examine in detail the water table shape changes for different values of  $\sigma$ . “Fast-response” systems ( $\sigma > 1 \text{ m}^{-1} \cdot \text{h}^{-1}$ ) generate high peak flows because of quick inflating/deflating processes of the water table shape  $W(X)$ . In such systems, peak flow values largely exceed the transient recession stage  $A(t)/J(t)$  component. Besides, at an hourly time step, the water table shape changes are fast enough to be neglected in the computations so that drain flow rates can be evaluated by the two first terms of Equation (12). In this case, the numerical solution of the Boussinesq equation can be simplified: the use of a simple Runge–Kutta method is sufficient for integrating the equations.

“Slow-response” systems ( $\sigma < 1 \text{ m}^{-1} \cdot \text{h}^{-1}$ ) do not generate high peak flow rates. The water table recharge contributes towards inflating the water table shape  $W(X)$ , which mitigates its effects on the drain flow rate. As a result, for these systems, accurate drain flow rate predictions require the computation of the three terms of Equation (12) through the complete solution to the Boussinesq equation.

For drainage simulation at a daily time step, water table shape changes have a lower impact on peak flow simulations. In the majority of French subsurface drainage systems,  $\sigma$  ranges between 0.05 and  $1.25 \text{ m}^{-1} \cdot \text{h}^{-1}$ . Tournebize *et al.* [2004] tested the errors due to the simplified approach that ignores the water table shape changes in the drain flow rate predictions. The authors showed that even for very slow systems ( $\sigma = 0.05 \text{ m}^{-1} \cdot \text{h}^{-1}$ ), the simplification error was negligible for daily flow predictions (Figure 2). This result was particularly useful when coupling the drainage simulation with other models running at a daily time step.



**Figure 2.** Evolution of Nash performance criterion versus time step to compare simplified equations of drainage modelling in the case of “fast” and “slow” drained soil systems [adapted from Tournebize *et al.*, 2004].

### 3.3. The SIDRA simulation tool

On the basis of the experimental and theoretical considerations above, a simulation tool of drainage systems, named “SIDRA” (Simulation of Drainage), was developed. The development of SIDRA, based on the above equations, followed different paths aimed at: (i) testing the quality of peak flow prediction; (ii) predicting surface runoff and its relative proportion versus subsurface drain flow; and (iii) coupling the drainage simulation with other tools simulating the effect of improved drainage or crop growth or the integrated hydrologic functioning of rural catchments. These different paths are summarized below together with the comparison against measured experimental results.

#### 3.3.1. Experimental sites

Several experiments have contributed to the validation of the simulation tool (Table 1). These experiments have been equipped since the 1980s and monitored for at least 10 years.

#### 3.3.2. Peak flow rate prediction in “fast” drainage systems

The Arrou site made it possible to analyse the seasonal functioning of drainage systems and to delineate precisely the periods of classic drainage functioning, termed “intense drainage seasons” [Lesaffre and Morel, 1986]. The initial validation work of

SIDRA was carried out for the intense drainage seasons of the period 1985–1995.

The value of  $\sigma$  in Arrou is equal to  $3.3 \text{ (m}\cdot\text{h)}^{-1}$  ( $K = 0.41 \text{ m/d}$  and  $f = 0.014$ ) for a drain spacing of 10 m, which makes this drainage system a “fast system” in which the water table shape changes are fleeting and can be neglected when simulating at a 1-h time step. The initial version of SIDRA did not consider these changes in water table shape. The management of the unsaturated zone was also very simple: a reservoir transfers the water table recharge instantly when full and is depleted by evapotranspiration between recharge events. During a rainy period, the reservoir is initially replenished before generating recharge. The simulation starts at the beginning of the intense drainage season during a rainfall event when the reservoir is assumed to be full.

In these conditions, and using the soil parameters  $K$  and  $f$  measured by Guyon’s pumping test [Guyon, 1976, Lesaffre, 1990], the drainage discharges, the water table elevations and the peak flow rates are well predicted [Lesaffre and Zimmer, 1988]. These results validated that, in the absence of significant surface runoff, the simple initial modelling of the system has a very good predictive capacity when the soil reaches its saturated status in winter. In such systems, an average of 13% of the rainfall intensity generates, together with the water table contribution, the peak flow rates observed. Outside of the intense drainage season, when the soil is not saturated, sporadic drainage discharges may be observed that cannot be predicted by the model.

### 3.4. Water balance of drained versus non-drained hydromorphic soils

A global understanding of hydrological processes occurring at the agricultural plot scale in subsurface-drained watersheds is essential for the improvement of management practices to control water pollution, soil erosion and flood genesis. The hydrological studies dealing with water balance in subsurface-drained areas received considerable attention in the second part of the 20th century [see Kao, 2008, for a review].

Among the various hydrological effects of subsurface drainage, a reduction in saturation excess surface runoff has been highlighted through the reduction of erosion [Skaggs *et al.*, 1982]. Subsurface drainage lowers the shallow water table, increases the

**Table 1.** Experimental sites with the corresponding SIDRA developments

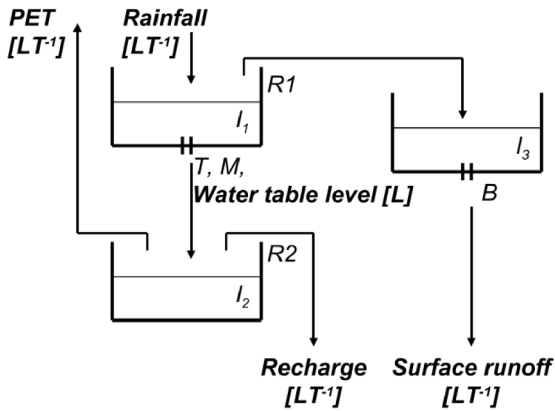
Experiment	Main purposes	Monitored data	References
Arrou (1 ha) 48° 05' 50.42 N 1° 07' 17.87 E Luvisol (Silty-Clayey)	Develop drainage theory, measure impacts on crop development and nitrate leaching	Precipitations, drain flow rates, water table elevations, water pressure, crop development, nitrate absorption and leaching	Zimmer [1989] Brisson <i>et al.</i> [2002] Tournebize <i>et al.</i> [2004]
La Jaillière (1 ha) 47° 27' 25.51 N 0° 51' 16.83 W Stagnic Luvisol (Silty-Clayey)	Impacts of drainage on crop development, on surface runoff and nitrate and pesticides transfers, impact and design of buffer zones	Precipitations, subsurface drain flow rates, surface runoff, water table elevations, solute and pesticides transfers	Kao <i>et al.</i> [1998] Henine <i>et al.</i> [2022] Branger <i>et al.</i> [2009] Jeantet <i>et al.</i> [2022] Kao [1994]
Chantemerle (36 ha) 48° 50' 32.01 N 3° 06' 41.82 E Luvisol (Silt)	Impact of drainage on water quality (nitrogen, carbon, pesticides). Hydrological question of calibration (using data assimilation)	Precipitations, discharges, high-frequency monitoring of nitrogen, seasonal exportation of pesticides	Chelil <i>et al.</i> [2022a]
Rampillon (355 ha) 48° 32' 19.66 N 3° 03' 43.91 E Luvisol (Silt)	Impact of drainage on water quality and remediation. Development of nitrate modelling	Precipitations, discharges, high-frequency monitoring of nitrogen, seasonal exportation of pesticides	Chelil <i>et al.</i> [2022b]
Merlachez (700 ha) 48° 52' 01.74 N 3° 11' 28.47 E Luvisol (Silt)	Impact of land drainage on flood dynamics	Precipitations, discharges	Henine <i>et al.</i> [2014]

storage volume prior to rainfall events and then reduces surface runoff [Lowery *et al.*, 1982, Enright and Madramootoo, 1994, Kao *et al.*, 1998].

Long-term experiments to record surface runoff rates in experimental subsurface-drained fields were performed in some locations in France [see, e.g. Kao *et al.*, 1998, Augeard *et al.*, 2005], in particular at the “La Jaillière” experimental site. The field experiment of La Jaillière is located in Maine et Loire [Henine *et al.*, 2022, Jeantet *et al.*, 2021, 2022] and was equipped since the early 1990s to monitor the

functioning of drainage systems as well as their impacts on water quality. For this purpose, it was equipped to measure both subsurface drainage and surface runoff, since their relative proportions have important consequences on the transport of: (i) solute products such as nitrates, mostly transferred by subsurface flows, and (ii) non-solute products associated with soil particles such as most pesticides, mostly transferred by surface runoff.

In these experimental fields, observed surface runoff volumes are usually small (from 5% to 10%



**Figure 3.** SIRUP module structure.  $I_1$ ,  $I_2$  and  $I_3$  are the water levels in the reservoirs;  $R1$  and  $R2$  are the volumes of Reservoirs 1 and Reservoir 2;  $T$ ,  $M$  and  $B$  are the parameters of the emptying laws for  $R1$  and  $R2$ ;  $PET$  is the potential evapotranspiration, adapted from Branger et al. [2009], and Kao et al. [1998].

of the rainfall amount), which confirms the high infiltration capacity of such soils. After autumn harrowing, surface runoff is mainly generated midway between drains when rainfall intensity has exceeded the soil infiltration capacity controlled by the water table depth. Surface runoff propagation over the area above the drain depends on the drainage efficiency (depth, spacing, possibly plugging).

Modelling the distribution of surface runoff and subsurface drain discharges was managed by coupling the SIDRA model with a reservoir-based model (SIRUP) aiming to generate three types of flows: (i) classic Hortonian surface runoff generated by heavy rainfall intensity before complete water replenishment of the upper soil horizon; (ii) Dunne surface runoff generated by the saturation of the soil mid-way between drains; and (iii) water table recharge rate [see a review and the development of the SIRUP model by Kao et al., 1998 and Branger et al., 2009].

The SIRUP model consists of three separate conceptual reservoirs, respectively, accounting for (Figure 3):

- water storage in the superficial soil layer, and infiltration/runoff distribution depending on the water table depth (Reservoir 1 with three parameters);

- storage of infiltrated water and moisture distribution in deeper soil layers, evapotranspiration and recharge to the water table (Reservoir 2 with one parameter);
- lamination of surface runoff (Reservoir 3 with one parameter).

Reservoir 1 has a maximum water level  $R1$  (L) and represents the rainfall water storage on the soil surface. Water flows from Reservoir 1 to 2 according to the emptying equation:

$$\varphi 1(t) = I1(t)[T(d - H(t)) + M] \quad (16)$$

where  $\varphi 1$  ( $L \cdot T^{-1}$ ) is the emptying flow,  $d$  (L) is the depth of the impervious layer,  $I1(t)$  (L) is the water level in Reservoir 1, and  $T$  ( $L^{-1} \cdot T^{-1}$ ) and  $M$  ( $T^{-1}$ ) are empirical parameters.

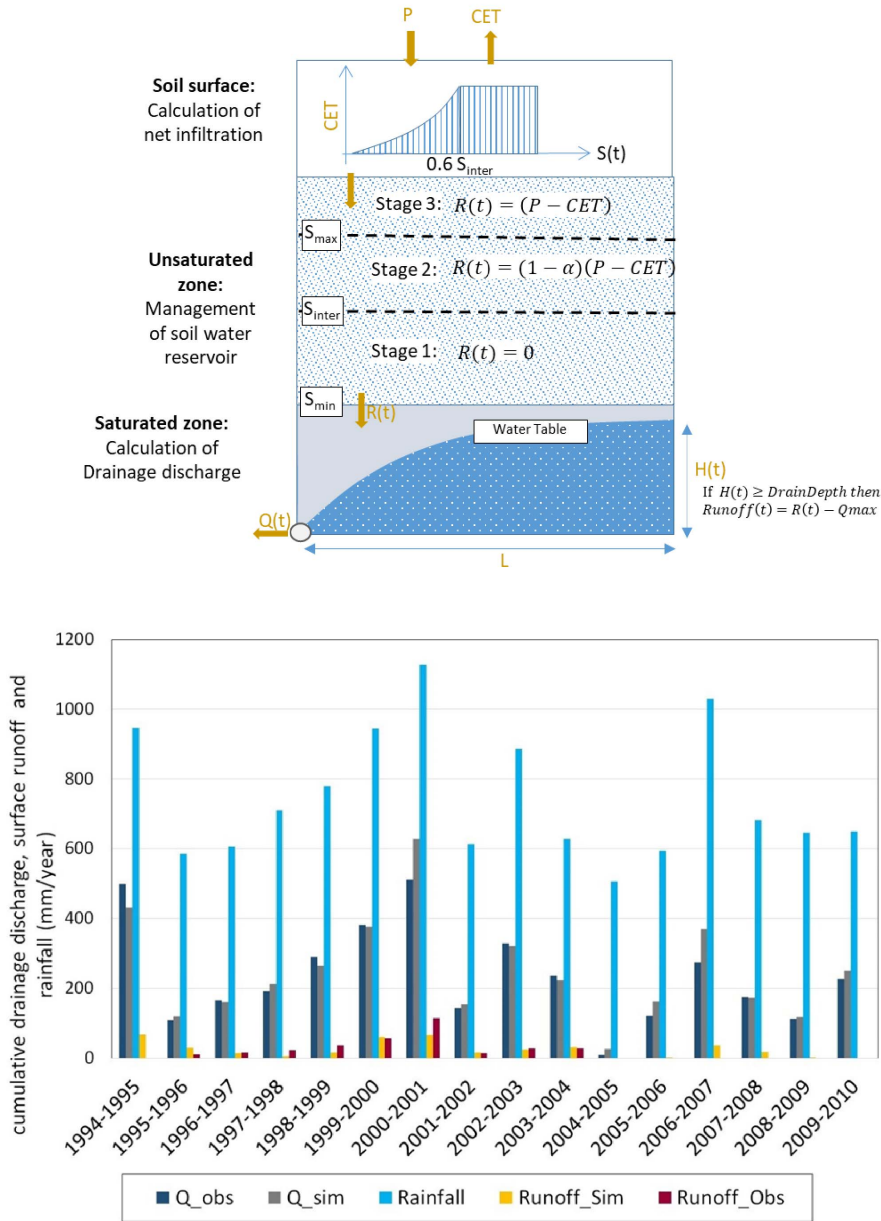
Equation (16) accounts very simply for the influence of the water table level on soil infiltration: a high water table level implies a reduced infiltration flow. When Reservoir 1 overflows, excess water flows to Reservoir 3 and is changed into surface runoff, according to:

$$\varphi 3(t) = I3(t)B \quad (17)$$

where  $\varphi 3$  ( $L \cdot T^{-1}$ ) is the emptying flow of Reservoir 3,  $I3(t)$  (L) is the water level and  $B$  ( $T^{-1}$ ) is an empirical parameter.

Reservoir 2 receives infiltration flow from Reservoir 1 and is emptied by evapotranspiration. It has a maximum water level  $R2$  (L). The overflowing water constitutes the recharge to the water table.

Based on the knowledge gained in La Jaillière, and in order to extend the use of SIDRA to other sites and to the different stages of replenishment of the soil water reserves across the year, a simplified version of the surface module SIRUP was developed and named “SIDRA-RU” [Henine et al., 2022; Figure 4]. Three stages were defined: (i) when the soil is unsaturated during late spring and summer, the water table recharge term is 0; (ii) during the fall, as water reserves are being replenished, one-third of the net rainfall replenishes the water table and two-thirds replenish the soil water reserves until soil saturation is reached; this required that the three parameters of SIRUP be replaced by two new parameters,  $Sinter$  and  $Smax$ ; (iii) after saturation is reached, during winter, water table recharge  $R(t)$  is the full net rainfall. The  $Sinter$  and  $Smax$  parameters represent two



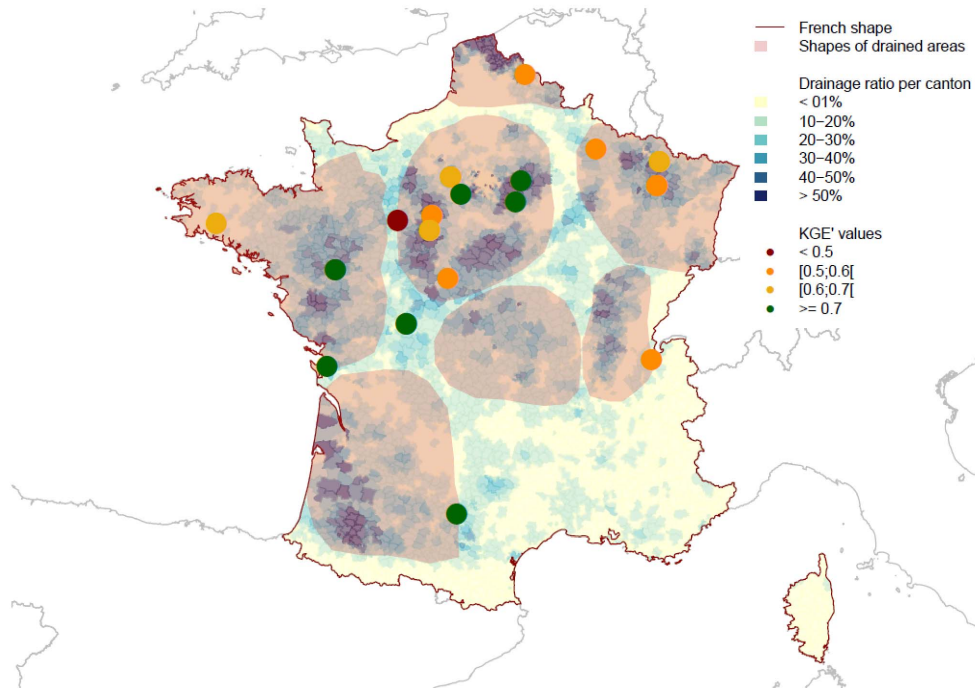
**Figure 4.** Coupling of SIDRA-RU and comparison between yearly cumulative observation and simulation results with SIDRA-RU for plot T04, using the parameter set of calibration period P1 (T04\_P1): rainfall, simulated and observed drainage discharge and surface runoff (mm/year) [adapted from Henine et al., 2022].

soil water depth thresholds. Sinter (L) is an intermediate level of replenishment allowing for initial discharges in subsurface-drained soils. It is of the order of 85% of the total water holding capacity. S<sub>max</sub> (L) corresponds to the full water replenishment of the

soil profile, allowing for a conventional water table to be present in the soils.

The simulated total runoff during the intense drainage season (i.e. when the soil water reserves have been replenished) was comparable to





**Figure 5.** Experimental sites used for assessment of the SIDRA-RU model in France and related performance with the Kling–Gupta efficiency (KGE) criterion. The map also displays the main regions where subsurface drainage has been implemented [adapted from Jeantet et al., 2021].

observations (35 mm/year) and represented 5% of total annual rainfall, congruent with the results of Moriasi et al. [2013]. Note that subsurface-drained discharges were estimated at 40% of annual rainfall, as reported by Moriasi et al. [2013], and surface runoff volumes in drained plots remain below 10% of the drained volume [Kao et al., 1998, Henine et al., 2022].

### 3.5. Robustness of SIDRA approach

The SIDRA model, initially developed and validated on two specific sites (Arrou and La Jaillière), was subsequently shown to predict drain flow rates and water table elevations on other sites. Its robustness was tested on 23 experimental sites in France with contrasting soil and climatic contexts [Jeantet et al., 2021]. These sites were monitored from 1969 to 2017. Their areas range from plot scale of ca. 1 ha to catchments of ca. 700 ha. The modelling was performed for a total of 170 years. SIDRA-RU requires few parameters as compared to other drainage models such as MACRO [Jarvis et al., 1997, Larsbo and Jarvis, 2003]

or DRAINMOD [Skaggs et al., 2012]. To obtain these parameters, two methods were used: regional delineation from calibrated values of the observed discharges on the 23 sites and using the split-sample test on nine of these 23 sites [Jeantet et al., 2021], or directly from on-site measurements through pumping tests such as Guyon's test [Lesaffre, 1990].

The performance criterion used in Figure 5 is the Kling–Gupta efficiency (KGE) [Gupta et al., 2009], defined as the combination of linear correlation, relative variability and bias between observed and simulated values. Values above 0.5–0.6–0.7 thresholds for hydrological models showed, respectively, acceptable, good and very good simulation quality [Moriasi et al., 2015]. The results showed congruent comparisons in simulations versus observations ( $KG2 > 0.5$ ) especially in loamy soils, representing 80% of drained soils in France [Lagacherie and Favrot, 1987]. The worst simulations were encountered in swelling clay soils ( $KGE < 0.5$ ), due to probable mis-alignment of the assumptions used in the model (depth to the impervious layer, important role of soil cracks

and macropores) such as Courcival (only point with  $KGE < 0.5$  in Figure 5). Some regions with significant drained areas in the centre of France (Auvergne, Burgundy) were not tested due to a lack of experimental data.

controlled by the number and position of hydraulic

#### **4. Impact of land drainage on flood generation and control**

Research aiming to understand the consequences of subsurface drainage expanded from this experimental and modelling approach. This was achieved from plot scale to catchments or marshlands with large field drainage intensities such as the Melarchez, part of the Orgeval catchment on the Seine basin (upstream of Paris), or the Moëze marshland, a 22-km<sup>2</sup> polder in western France. The SIDRA model was used to generate the distributed production functions of hydraulic models solving the classic Saint-Venant equations [Giraud *et al.*, 1997].

The effect of subsurface drainage on floods is a matter of debate. Does subsurface drainage increase or decrease flooding? A modelling approach was used by Henine *et al.* [2012, 2014] to represent the complex interactions between pipe networks and groundwater flows during flood events. In this case, free-surface flow conditions in pipe drains are no longer respected and may flow under pressurized conditions due to high water levels in open channels at the drainage outlets. The approach consisted in coupling subsurface drainage processes through a 2D Boussinesq equation with a 1D Saint-Venant network model to take into account the interactions between subsurface drainage and open-channel flows. This model was used to investigate the effects of pressurization during flood events as compared with non-pressurization at the pipe outlet and compared with experimental observations. The simulation results were in good agreement with experimental observations [Henine *et al.*, 2014]. The contribution of subsurface drainage flow to floods could be limited during peak events by redesigning the pipe outlet diameter (reducing the diameter of the collector into the ditch). This would be equivalent to an increase in the field buffer capacity of a few millimetres of rainfall.

In general, the arterial drainage function is the most influential and its storage and laminating capacity is essential. This capacity is to a great extent

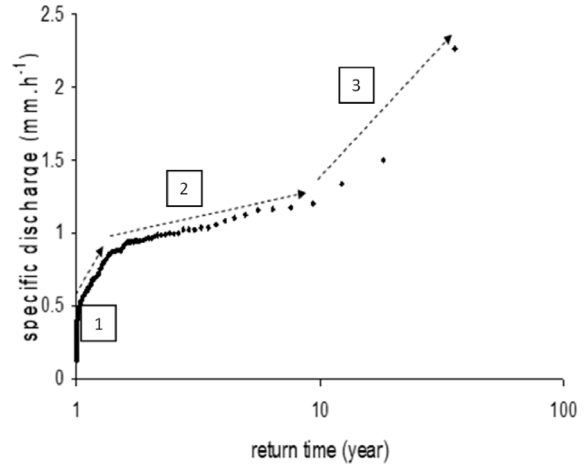
structures: reduced cross sections (e.g. culverts in road crossovers) or of a network of small gullies that has been built over time in wet areas.

In large marshlands (e.g. polders), the key buffer capacity is related to the large number of small gullies that store water in a decentralized way. Reducing this number to create large field plots, as is often implemented in the intensification process of agriculture, has been shown to have the greatest impact on wetland hydrology [Giraud *et al.*, 1997]. By comparison, the influence of field drainage on the buffering capacity of the marshland is much smaller. In this application, controlled drainage, as largely studied in the United States, could be an option [DRAINMOD model application of drainage water management such as controlled drainage, *i.e.* Sloan *et al.*, 2016].

Similar results can be deduced from observations and simulations carried out in the Orgeval basin (east of Paris). In this catchment that is representative of the French context, the retention capacity of the ditches determined by the number of culverts and the capacity to store water upstream of these is critical for the flood laminating capacity. More precisely, the number of retention structures along the network is more critical than their total storage capacity, which is explained by the dynamic storage (*i.e.* the fact that a given quantity of water is slowed down several times successively during its downstream transfer) provided by these structures.

Beyond these generic conclusions, the consolidated impacts of land drainage (*i.e.* the combination of field and arterial drainage) on hydrological regimes of rural catchments depend on the return periods of floods (Figure 6):

- Stage 1: increase in flood intensity for return periods aligned with the design of the subsurface drainage network (1–2 years); in that case, subsurface drainage has a negative impact on hydrological regimes by accelerating the propagation of the flood;
- Stage 2: self-limitation and storage in the network or in the field, for floods with a return period between 5 and 10 years; in that case, drainage has a positive impact by attenuating the propagation of the flood;
- Stage 3: beyond a return period of 10 years, in the hydro-system being saturated, drainage no longer shows any impact.



**Figure 6.** Flood return periods of hourly specific discharges measured at Melarchez Station (Orgeval Observatory <https://gisoracle.inrae.fr/>) during a 20-year period (1995–2015). Stages 1, 2, 3 correspond to increasing impacts, self-limitation, and no additional impacts of drainage on flooding, respectively [adapted from Henine *et al.*, 2012].

## 5. Solute transfer highlights hydrological functioning

Drain flow dynamics were studied by several authors [Jacobsen and Kjær, 2007, Dairon *et al.*, 2017, Kao, 2002, Gatel *et al.*, 2019] using in situ tracer experiments or model results. Since nitrate and some pesticides are highly soluble, the specific studies of their transfer dynamics at the drainage outlet stressed the preferential flow above the pipe and the matrix contribution from the area between the pipe and the mid-drain space. This duality of behaviour corresponds to the fast and slow transfers. The C-Q relationships shift during peak events exhibiting hysteresis patterns due to a high variability of the flushing/dilution effect [Liu *et al.*, 2020]. From this state, a conceptual approach based on compartments was developed and coupled with the SIDRA model. Branger *et al.* [2009] proposed PESTDRAIN based on coupling of the SIDRA–SIRUP–SILASOL modules (Figure 7). The originality of this approach is the simplified choice of hydrological process for solute transfer by using transfer functions [Jury and Roth, 1990]. An exponential formula [Magesan *et al.*, 1994] is applied for two distinct arti-

ficial compartments (fast and slow). The calibration procedure led to quantification of the volume of soil reservoirs at 2 mm for fast transfers above the drain and at 200 mm for slow transfers between the drain and mid-drain space, thereby successfully reproducing pesticide transfer at the plot scale. The interesting point is the surface contribution of the two distinct compartments to the exported fluxes: 13% for fast transfers (area above the drain) and 87% for slow transfers (between the drain and mid-drain space), corresponding to the hydrological contribution (1-A) and A from SIDRA Equation (12).

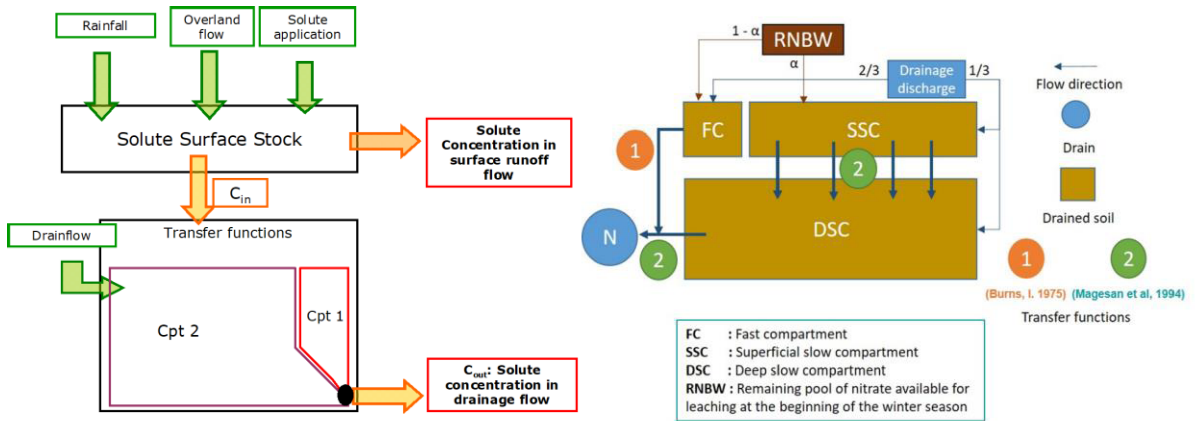
High-frequency monitoring of nitrate concentrations (hourly observation at the drainage outlet) provides clues to improve the conceptual approach of solute modelling based on hydrological drainage functioning. Thus, a new model for nitrate, NITDRAIN [Chelil *et al.*, 2022a,b], considers three compartments and two distinct transfer functions (Figures 7 and 8): (1) the Burns equation for preferential vertical transfer [Burns, 1975], and (2) the Magesan exponential equation for matrix lateral flow for upper and deeper slow reservoirs. The model succeeds in reproducing the flushing and dilution typical of nitrate concentration behaviour in subsurface-drained soil (Figure 8a and b). The mixing water at the outlet of the drained plot is composed of fast and slow flows, as presented in Figure 8. The fast compartment (FC) contributes directly to the flow outlet above the drain pipe. The surface slow compartment (SSC) transfers solute to the deep slow compartment (DSC), which is connected to the flow outlet (Figure 8c). The model provides a temporal evolution of solute stock in all compartments (Figure 8d). The results obtained with this model also help to improve our understanding of hydrological drainage processes. The switch between nitrate flushing and nitrate dilution concentration behaviour at the pipe outlet (Figure 8d) occurred specifically when the hydrological status of the drained soil switches from the beginning stage (unsaturated soil profile) to the intensive drainage season (saturated soil profile).

Both examples of solute modelling confirm the dual processes between preferential flow above the drain pipe and the water table slow contribution from the rest of the transect, as the hydraulic approach showed. The understanding of hydrological processes in subsurface-drained soils is thus supported by the analysis of solute transfer.

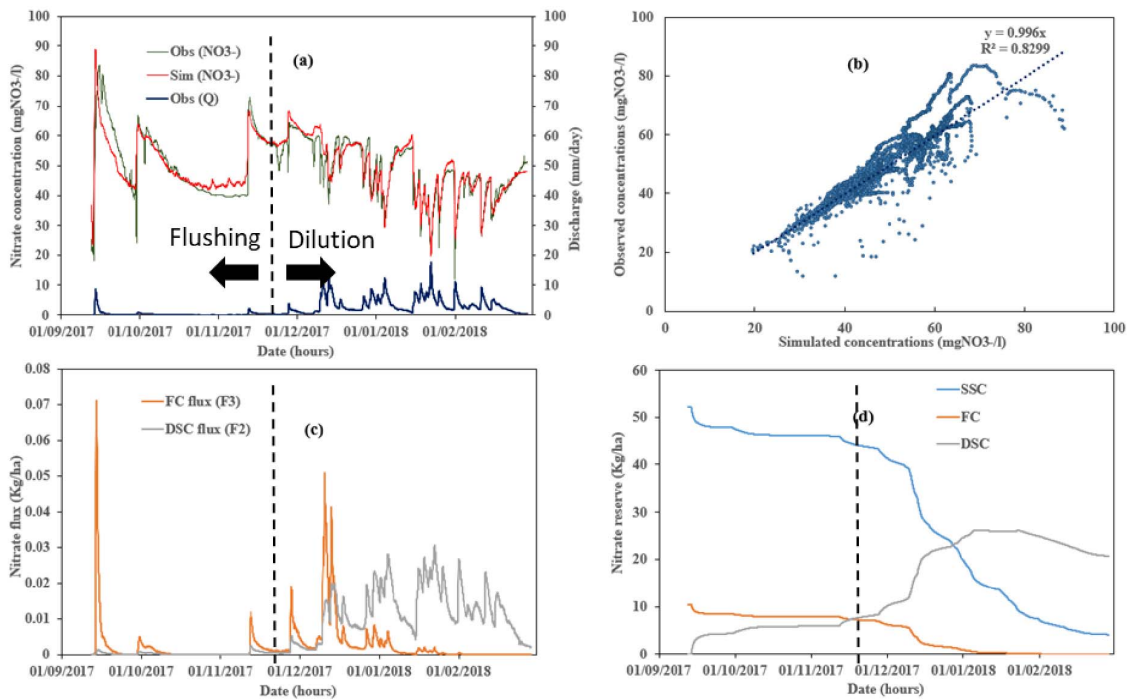
## 6. Conclusions

During 30 years of research on the functioning of subsurface drainage, we could demonstrate that relatively simple modelling tools based on a few equations and a limited number of parameters provide important insights into the functioning of drainage systems. In particular, the semi-analytical solution to the Boussinesq equation offers a simple way to describe the functioning of a majority of drainage systems in the French context as well as very important insights explaining their functioning. Specifically:

- It provides an explanation of the experimental results and confirms that, in most subsurface-drained soils, there is no need to advocate preferential flow in the plough layer or in the trench backfill to explain the peak drain flow rates at the outlet of these systems. This result was shown to be relatively robust in many soils except for heavy clay soils where the modelling approach did not reproduce the drain flow rates accurately.
- It provides a correct description of the relative disconnection between water table elevations and subsurface drain flow rates observed in experimental fields. Drain flow rates depend not only on the water table position in the soil but also on the recharge rate reaching that water table. The theory provides a good order of magnitude of the proportion (around 13%) of that recharge rate that contributes to the drain flow.
- It goes even further and adds to the water table elevation and the recharge rate a third component, i.e. the drain flow rate related to the swelling or deflating of the water table shape during recharge events. The reparameterization of the equation also demonstrates that drain flow rates can be fully predicted by a single parameter ( $\sigma$ ), a combination of two soil parameters (saturated hydraulic conductivity and drainable porosity) and of the drainage system dimension (drain spacing). With high values of  $\sigma$ , the swelling–deflating process is very brief and can be ignored in the modelling process. With low  $\sigma$  values, the process cannot be ignored and has a mitigation effect on the peak flow rate.



**Figure 7.** Solute transfer in drained soil: PESTDRAIN version [left side, Branger *et al.*, 2009] considering two reservoirs and the Magesan transfer function; NITDRAIN version [right side, Chelil *et al.*, 2022a,b] considering three reservoirs and the Burns and Magesan transfer functions. Drain flow rates are simulated by the SIDRA model.



**Figure 8.** Simulation of nitrate concentrations with NITDRAIN on Rampillon watershed (Seine et Marne) showing the time of switching between flushing and dilution processed (dashed line): (a) time series of simulated and observed nitrate concentrations, (b) simulated vs observed concentrations, (c) simulated contribution fluxes of fast compartment (FC) and deep slow compartment (DSC), (d) evolution of nitrate storage in FC, DSC and SSC (surface slow compartment) [adapted from Chelil *et al.*, 2022b].

These theoretical results developed at field level have predictive capacity for other drainage systems such as water tables drained by water channels or rivers. They could also be extended to the generic case of deep impervious barriers. They explain in particular the steep increase of the flow rates during recharge periods in such systems and as such may have practical importance for the prediction of peak flows in such systems.

On the basis of these results, a series of modelling tools were developed, starting with the simple model SIDRA, based on a small number of parameters (4), which has demonstrated its capacity to predict subsurface drain flows—and particularly peak flows—in various regions of France during the winter drainage season. Papers dealing with the SIDRA model from 1987 comprised about 48 documents: 18 publications in peer-reviewed journals, 20 as conference presentations, and 10 as technical publications. In 2022, Scopus identified more than 235 citations of the SIDRA model in publications. As SIDRA is modular and can run independently, its insertion in different modelling tools is easier. For instance, the agronomic model STICS [Brisson *et al.*, 2002] simulating crop production as well as environmental data below the soil root zone (water balance and nitrogen leaching) proposes a “subsurface drainage” option including SIDRA’s concept.

SIDRA was subsequently enriched and coupled with several other modelling tools to analyse the impact of subsurface drainage on the hydrology of rural catchments.

- A surface runoff module (SIRUP) was developed and validated against experimental data to predict the share of surface runoff and subsurface drainage.
- Coupling SIDRA with hydraulic models solving the Saint-Venant equations in the ditches and channels of the arterial drainage network was also used to analyse the impacts of land drainage at catchment level [Branger *et al.*, 2010]. Results from a polder and from a more classic catchment demonstrated the predominant role of the retention in the channel system often reduced by the removal of the network of small gullies (e.g. after land consolidation programmes), and the recalibration of the ditch network combined with culverts and cross-overs designed for 10-year

return flows (to avoid road degradation during floods). The field subsurface drainage instead has a mitigating effect on floods for average return periods (>10 years).

- Finally, SIDRA was coupled with solute transfer models to understand the impacts of drainage on pesticide and nitrate transfers in relation to surface and subsurface flow patterns. The hydrological understanding and modelling conceptualization of SIDRA helped and supported the compartmental approach leading to a simplified modelling structure (PESTDRAIN and NITDRAIN). Results confirmed the duality of solute behaviour in drained soil with a fast transfer from soil surface above the pipe area during peak flow and a slow transfer for the rest of the transect mid-drain space, especially during the water table tail recession.

### Conflicts of interest

Authors have no conflict of interest to declare.

### Dedication

The authors would like to dedicate the SIDRA model story to Ghislain de Marsily from Sorbonne University and a member of the *Académie des Sciences*, for his scientific support (participating directly in the PhD supervision or as a PhD jury member).

### Acknowledgments

The authors would like to thank Bernard Vincent (1952–2022) for his relentless contribution to the improvement of drainage practices in France.

### References

- Akay, O., Fox, G. A., and Šimůnek, J. (2008). Numerical simulation of flow dynamics during macropore–subsurface drain interactions using HYDRUS. *Vadose Zone J.*, 7, 909–918.
- Augeard, B., Kao, C., Chaumont, C., and Vauclin, M. (2005). Mechanisms of surface runoff genesis on a subsurface drained soil affected by surface crusting: a field investigation. *Phys. Chem. Earth*, 30, 598–610.

- Bjerre, E., Fienen, M. N., Schneider, R., Koch, J., and Højberg, A. L. (2022). Assessing spatial transferability of a random forest metamodel for predicting drainage fraction. *J. Hydrol.*, 612, article no. 128177.
- Bouarfa, S. and Zimmer, D. (2000). Water table shapes and drain flow rates in shallow drainage systems. *J. Hydrol.*, 235(3–4), 264–275.
- Boussinesq, J. (1904). Recherches théoriques sur l'écoulement des nappes infiltrées dans le sol ; compléments. *J. Math. Pures Appl.*, 10(1), 5–78. et 10(4), 63–394.
- Branger, F., Braud, I., Debionne, S., Viallet, P., Dehotin, J., Henine, H., Nédélec, Y., and Anquetin, S. (2010). Towards multi-scale integrated hydrological models using the LIQUID<sup>®</sup> framework. Overview of the concepts and first application examples. *Environ. Model. Softw.*, 25(12), 1672–1681.
- Branger, F., Tournebize, J., Carluier, N., Kao, C., Braud, I., and Vaucelin, M. (2009). A simplified modelling approach for pesticide transport in a tile-drained field: The PESTDRAIN model. *Agric. Water Manage.*, 96, 415–428.
- Brisson, N., Rebière, B., Zimmer, D., and Renault, P. (2002). Response of the root system of a winter wheat crop to waterlogging. *Plant Soil*, 243(1), 43–55.
- Burns, I. G. (1975). An equation to predict the leaching of surface-applied nitrate. *J. Agric. Sci.*, 85, 443–454.
- Chelil, S., Oubanas, H., Henine, H., Igor, G., Malaterre, P. O., and Tournebize, J. (2022a). Variational data assimilation to improve subsurface drainage model parameters. *J. Hydrol.*, 610, article no. 128006.
- Chelil, S., Tournebize, J., Henine, H., and Chaumont, C. (2022b). NIT-DRAIN model to simulate nitrate concentrations and leaching in a tile-drained agricultural field. *Agric. Water Manage.*, 271, article no. 107798.
- Dairon, R., Dutertre, A., Tournebize, J., Marks-Perreau, J., and Carluier, N. (2017). Long-term impact of reduced tillage on water and pesticide flow in a drained context. *Environ. Sci. Pollut. Res.*, 24(8), 6866–6877.
- de Marsily, G. (1994). On the use of models in hydrology [free opinion]. *Rev. des Sciences de l'eau/J. Water Sci.*, 7(3), 219–323. (in abstract).
- Dumm, L. D. (1964). Transient flow concept in subsurface drainage: its validity and use. *Trans. ASAE*, 7(2), 142–146.
- Enright, P. and Madramootoo, C. A. (1994). Hydrologic response of surface and subsurface drained agricultural fields. *Can. Agric. Eng.*, 36(1), 15–24.
- Favier, M., Lesaffre, B., Leviandier, T., Penel, M., and Zimmer, D. (1990). Deterministic versus conceptual models to evaluate the influence of drainage on water regime. In *Proceedings of Hydrological Research Basins and the Environment, Wageningen*, volume 44, pages 123–130.
- Gatel, L., Lauvernet, C., Carluier, N., Weill, S., Tournebize, J., and Paniconi, C. (2019). Global evaluation and sensitivity analysis of a physically based flow and reactive transport model on a laboratory experiment. *Environ. Model. Softw.*, 113, 73–83.
- Giraud, F., Faure, J. B., Zimmer, D., Lefeuvre, J. C., and Skaggs, R. W. (1997). Hydrologic modeling of a complex wetland. *J. Irrig. Drain. Eng.*, 123(5), 344–353.
- Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F. (2009). Decomposition of the mean squared error and nse performance criteria: implications for improving hydrological modelling. *J. Hydrol.*, 377(1–2), 80–91.
- Guyon, G. (1964). Considérations sur l'hydraulique du drainage des nappes par canalisations enterrées. *Bull. Tech. Génie Rural*, 65. 45 pages + annexes.
- Guyon, G. (1976). Détermination de la porosité de drainage sur le terrain. *Bull. Int. CIID*, 25, 112–117.
- Henine, H., Chaumont, C., Tournebize, J., Augéard, B., Kao, C., and Nédélec, Y. (2012). Le rôle des réseaux de drainage agricole dans le ralentissement dynamique des crues : interprétation des données de l'observatoire Orgeval (image). *Sci. Eaux Territoires Cah. Spec.*, 3, 16–23.
- Henine, H., Jeantet, A., Chaumont, C., Chelil, S., Lauvernet, C., and Tournebize, J. (2022). Coupling of a subsurface drainage model with a soil reservoir model to simulate drainage discharge and drain flow start. *Agric. Water Manage.*, 262, article no. 107318.
- Henine, H., Nédélec, Y., and Ribstein, P. (2014). Coupled modelling of the effect of overpressure on water discharge in a tile drainage system. *J. Hydrol.*, 511, 39–48.
- Hervé, J. (1980). Limites et validité des modèles

- hydrodynamiques. Document Cemagref, Antony, 20 p.
- Hooghoudt, S. B. (1940). Bijdragen tot de kennis van enige natuurkundige grootheden van de grond. No. 7. Versl. Landbouwk. Onderz. (Contributions to the knowledge of some physical constants of the soil. No. 7. Report Agric. Res.). 46, 515–707.
- Jacobsen, O. H. and Kjær, J. (2007). Is tile drainage water representative of root zone leaching of pesticides? *Pest Manage. Sci.*, 63(5), 417–428.
- Jarvis, N. J., Hollis, J. M., Nicholls, P. H., Mayr, T., and Evans, S. P. (1997). MACRO-DB: a decision-support tool for assessing pesticide fate and mobility in soils. *Environ. Model. Softw.*, 12, 251–265.
- Jeantet, A., Henine, H., Chaumont, C., Collet, L., Thirel, G., and Tournebize, J. (2021). Robustness of a parsimonious subsurface drainage model at the French national scale. *Hydrol. Earth Syst. Sci. Discuss.*, 2021, 1–33.
- Jeantet, A., Thirel, G., Jeliakov, A., Martin, P., and Tournebize, J. (2022). Effects of climate change on hydrological indicators of subsurface drainage for a representative french drainage site. *Front. Environ. Sci.*, 10, article no. 899226.
- Jury, W. A. and Roth, K. (1990). *Transfer Functions and Solute Movement through Soil; Theory and Applications*. Birkhäuser Verlag, Basel.
- Kao, C. (1994). Le ruissellement en parcelle agricole en situation hivernale. Mécanismes de genèse et modélisation. Parcelles drainées et non drainées, site expérimental de la Jaillière (Maine et Loire). Mémoire de Stage de Recherche. DEA National d'Hydrologie.
- Kao, C. (2002). *Fonctionnement hydraulique des nappes superficielles de fonds de vallées en interaction avec le réseau hydrographique*. Thèse de doctorat, Université Pierre et Marie Curie, Paris. 266 pages + annexes.
- Kao, C. (2008). *Une hydrologie des versants à nappes superficielles drainées : approches multi-processus et multi-échelles*. Mémoire d'habilitation à diriger des recherches, Université Pierre et Marie Curie - Paris VI. 98 pages + annexes.
- Kao, C., Bouarfa, S., and Zimmer, D. (2001). Steady state analysis of unsaturated flow above a shallow water-table aquifer drained by ditches. *J. Hydrol.*, 250, 122–133.
- Kao, C., Nedelec, Y., and Zimmer, D. (1998). Surface runoff mechanisms and modelling in subsurface drained fields. In Brown, L., editor, *Proceedings of the 7th International Drainage Symposium "Drainage in the 21st Century: Food Production and the Environment"*, Orlando, USA, 8–10 March 1998, volume 7, pages 258–266. ASAE.
- King, K. W., Fausey, N. R., and Williams, M. R. (2014). Effect of subsurface drainage on streamflow in an agricultural headwater watershed. *J. Hydrol.*, 519(PA), 438–445.
- Lagacherie, P. and Favrot, J. C. (1987). Synthèse générale sur les études de secteurs de référence drainage (période 1981–1986). SES Montpellier, INRA, Montpellier, n° 591, 146 pages.
- Larsbo, M. and Jarvis, N. J. (2003). MACRO 5.0. A model of water flow and solute transport in macroporous soil. Technical report, Swedish University of Agricultural Sciences, Uppsala, Sweden.
- Lesaffre, B. (1989). Fonctionnement hydrologique et hydraulique du drainage souterrain des sols temporairement engorgés : débits de pointe et modèle SIDRA. Etudes du Cemagref, Série Hydraulique Agricole, 4, 334 pages.
- Lesaffre, B. (1990). Field measurement of saturated hydraulic conductivity and drainable porosity using Guyon's pumping test. *Trans. ASAE*, 1(33), 173–178.
- Lesaffre, B. and Morel, R. (1986). Use of hydrographs to survey subsurface drainage networks ageing and hydraulic operating. In *Agric. Water Management Seminar*, pages 175–189. Balkema, Arnhem, The Netherlands.
- Lesaffre, B. and Penel, M. (1990). Drainage project preparation in France. In *4th International Drainage Workshop, Cairo, Egypt*, pages 253–256. Cemagref.
- Lesaffre, B. and Zimmer, D. (1988). Subsurface drainage peak flows in shallow soil. *J. Irrig. Drain. Eng.*, 114(3), 387–406.
- Liu, W., Youssef, M. A., Birgand, F. P., Chescheir, G. M., Tian, S., and Maxwell, B. M. (2020). Processes and mechanisms controlling nitrate dynamics in an artificially drained field: Insights from high-frequency water quality measurements. *Agric. Water Manage.*, 232, article no. 106032.
- Lowery, B., Kling, G. E., and Vomocil, J. A. (1982). Overland flow from sloping land: effects of perched water tables and subsurface drains. *Soil Sci. Soc. Am. J.*, 46(1), 93–99.
- Magesan, G. N., Scotter, D. R., and White, R. E.



- (1994). A transfer function approach to modelling the leaching of solutes to subsurface drains I. Non-reactive solutes. *Aust. Soil Res.*, 32(1), 69–83.
- Moriasi, D., Zeckoski, R. W., Arnold, J., Baffaut, C., Malone, R. W., Daggupati, P., Guzman, J. A., Saraswat, D., Yongping Yuan, Y., Wilson, B. N., Shirmohammadi, A., and Mankin, K. R. (2015). Hydrologic and water quality models: key calibration and validation topics. *Trans. ASABE*, 58(6), 1609–1618.
- Moriasi, D. N., Gowda, P. H., Arnold, J. G., Mulla, D. J., Ale, S., Steiner, J. L., and Tomer, M. D. (2013). Evaluation of the hooghoudt and kirkham tile drain equations in the soil and water assessment tool to simulate tile flow and nitrate-nitrogen. *J. Environ. Qual.*, 42, 1699–1710.
- Russel, J. L. (1934). Scientific research in soil drainage. *J. Agric. Sci.*, 24(4), 544–573.
- Schneider, R., Stisen, S., and Højberg, A. L. (2022). Hunting for information in streamflow signatures to improve modelled drainage. *Water (Switzerland)*, 14(1), article no. 110.
- Skaggs, R. W., Nassehzadeh-Tabrizi, A., and Foster, G. R. (1982). Subsurface drainage effects on erosion. *J. Soil Water Conserv.*, 37(3), 167–172.
- Skaggs, R. W., Youssef, M., and Chescheir, G. M. (2012). DRAINMOD: model use, calibration, and validation. *Trans. ASABE*, 55, 1509–1522.
- Sloan, B. P., Basu, N. B., and Mantilla, R. (2016). Hydrologic impacts of subsurface drainage at the field scale: Climate, landscape and anthropogenic controls. *Agric. Water Manage.*, 165, 1–10.
- Tcharny, I. A. (1951). Démonstration rigoureuse de la formule de Dupuit pour un écoulement à surface libre avec surface de suintement. *C. R. Acad. Sci. d'URSS*, 79(6), 937–940.
- Tournebize, J., Kao, C., Nikolic, N., and Zimmer, D. (2004). Adaptation of the STICS model to subsurface drained soils. *Agronomie*, 24(6–7), 305–313.
- Vachaud, G. (1968). Etude du régime transitoire de l'écoulement de l'eau pendant le drainage des nappes à surface libre. *Houille Blanche*, 5, 391–406.
- Valayamkunnath, P., Gochis, D. J., Chen, F., Barlage, M., and Franz, K. J. (2022). Modeling the hydrologic influence of subsurface tile drainage using the national water model. *Water Resour. Res.*, 58(4), 1–24.
- Van Schilfgaarde, J. (1963). Design of tile drainage for falling water table. *J. Irrig. Drain. Division, ASCE*, 89(2), 1–11.
- Varvaris, I., Pittaki-Chrysodonta, Z., Børgesen, C. D., and Iversen, B. V. (2021). Parameterization of two-dimensional approaches in HYDRUS-2D: Part 1. Simulating water flow dynamics at the field scale. *Soil Sci. Soc. Am. J.*, 85, 1578–1599.
- Vauclin, M., Khanji, D., and Vachaud, G. (1976). Etude expérimentale et numérique du drainage et de la recharge des nappes à surface libre, avec prise en compte de la zone non saturée. *J. Méc.*, 15(2), 307–348.
- Williams, M. R. and McAfee, S. J. (2021). Water storage, mixing, and fluxes in tile-drained agricultural fields inferred from stable water isotopes. *J. Hydrol.*, 599, article no. 126347.
- Wolsack, J. (1978). Quelques extensions de la théorie du drainage au cas des sols hétérogènes et anisotropes. *Bull. du BRGM*, 3(2), 101–113.
- Zimmer, D. (1989). Transferts hydriques des sols drainés par tuyaux enterrés: compréhension des débits de pointe et essai de typologie des schémas d'écoulement. Etudes du Cemagref, Série Hydraulique Agricole, 5, 321 pages.
- Zimmer, D., Lorre, E., and Lesaffre, B. (1995). Parameter sensitivity and field evaluation of SIDRA model. *Irrig. Drain Syst.*, 9(3), 279–296.



# **Groundwater Hydrology**





Research article

Geo-hydrological Data & Models

# Hard rock aquifer architecture and water circulation levels in the Strengbach critical zone observatory (France)

François Chabaux <sup>\*,a</sup>, Philippe Négrel <sup>\*,b</sup>, Frederick Gal <sup>\*,b</sup>, Daniel Viville <sup>a</sup>, Coralie Ranchoux <sup>a</sup>, Jérôme Van der Woerd <sup>a</sup>, Frederick Delay <sup>a</sup>, Julien Ackerer <sup>a</sup>, Luc Aquilina <sup>c</sup>, Marc Diraison <sup>a</sup>, Frédéric Masson <sup>a</sup>, Anne-Sophie Mériaux <sup>a</sup>, Thierry Labasque <sup>c</sup>, Yann Lucas <sup>a</sup>, Philippe Adrien Pezard <sup>d</sup>, Thierry Reuschlé <sup>a</sup>, Anne-Désirée Schmitt <sup>a</sup>, Peter Stille <sup>a</sup>, Robert Wyns <sup>b</sup>, Catherine Lerouge <sup>\*,b</sup> and Chrystel Dezayes <sup>\*,b</sup>

<sup>a</sup> Institut Terre et Environnement de Strasbourg, Université de Strasbourg, CNRS, ENGEES, 67000 Strasbourg, France

<sup>b</sup> BRGM, 45060 Orléans, France

<sup>c</sup> Géosciences Rennes, Université Rennes, CNRS, 35000 Rennes, France

<sup>d</sup> Géosciences Montpellier, CNRS, Université de Montpellier, Place Eugène Bataillon, 34000 Montpellier, France

*E-mails:* fchabaux@unistra.fr (F. Chabaux), p.négrel@brgm.fr (P. Négrel), f.gal@brgm.fr (F. Gal), c.lerouge@brgm.fr (C. Lerouge), c.dezayes@brgm.fr (C. Dezayes)

**Abstract.** We present an integrated petrological, petrophysical, and hydrogeological study of the critical zone (CZ) developed in the Hercynian granitic basement of the Strengbach watershed (Vosges Massif, France) to characterize its deep architecture and water circulation levels. For this purpose, six boreholes (50–120 m depth), from which three are cored, and three piezometers (10–15 m depth) were drilled to define the vertical extension and lateral variability of the main CZ horizons.

The Strengbach watershed is composed of a topsoil horizon of limited vertical extension (0.8–1.2 m), a mobile saprolite level, and an in-place fractured bedrock. The latter is subdivided into a few meters thick saprock horizon, defined by open sub-horizontal fractures and a deeper fractured bedrock horizon with steeply dipping fractures (>50°). In the north-facing slope, the vertical extension of the mobile saprolite horizon increases from ≈1–2 m at the top of the slope to ≈9 m downstream, close to the valley bottom. In contrast, the south-facing and more easterly slope shows a mobile saprolite horizon with limited vertical extension (≈2–3 m thick). Such a difference is associated with the existence of a knickpoint in the river bed, separating a downstream zone marked by currently active erosion from an upstream one, less prone to erosion, with preserved reliefs formed around 20 ka ago.

The water circulation scheme within the Strengbach watershed involves two different systems: a subsurface circulation within the shallow aquifer, corresponding to the mobile saprolite horizon and the saprock, and a deeper circulation in the fractured bedrock. The water circulation in the fractured bedrock is controlled by fractures of regional orientations, linked to the Vosges massif and the Rhine Graben Tertiary tectonics, and partly to reactivated Hercynian fracture zones. The unaltered bedrock was not reached by any of the three cores. These results from the Strengbach CZ demonstrate the

\* Corresponding authors.

importance of integrating geological history of the watershed, either the long-term geological bedrock evolution or the Quaternary erosion patterns, to better understand and model the CZ hydrological functioning at the watershed scale.

**Keywords.** Deep critical zone architecture, Hard rock aquifer, Water circulation levels, Geological heritage, Strengbach Critical Zone Observatory.

**Funding.** French ANR program (Grant agreement no. ANR-15-CE06-0014), REALISE program (CPER Etat Alsace 2003–2014), INSU-CNRS and French Equipex program CRITEX.

*Manuscript received 6 September 2022, revised 21 July 2023, accepted 26 September 2023.*

*“Hydrogeology is a discipline in which everything happens, by definition, hidden from view in the crevices of the underground environment. What is required of the hydrogeologist is to explain what happens underground...” [de Marsily, 2004].*

## 1. Introduction

The Critical Zone (CZ) is classically defined as “the near-surface layer where water, rock, air, and life meet in a dynamic interplay that sculpts landscapes, generates soils, and builds foundation for Earth’s terrestrial ecosystems” (National Research Council NRC 2001). Its vertical extension, beyond the few meters classically studied through the analysis of weathering profiles, is however still an open question [e.g., Riebe *et al.*, 2017]. This lack of knowledge is probably associated with difficulties in sampling and gathering information from the deep CZ shielded by several meters of weathered materials [Riebe *et al.*, 2017, Holbrook *et al.*, 2019]. A commonly accepted view—but still debated [see Riebe *et al.*, 2017]—is to assimilate the CZ lower boundary to the level where subsurface materials grades into compact rocks no longer affected by meteoric fluids [Riebe *et al.*, 2017]. Thus, defining the deep CZ means characterizing the zones of groundwater flow connected with surface circulation and distinguishing them from deeper levels without evidence of water circulation or with circulations unrelated to surface processes. Parallel to the rise of the CZ concept, hydrogeological studies on continental water resources led to the definition of a conceptual structural scheme for Hard Rock Aquifers (HRA), *i.e.*, aquifers in granitic and silicate metamorphic bedrocks [review in Lachassagne *et al.*, 2021]. The theoretical HRA scheme is a stratiform structure beneath ground level and above non-weathered rocks, consisting of an unconsolidated saprolite (or regolith) level (a clay-rich zone formed by the weathering of silicate rocks) superimposed on a permeable fissured layer. The latter is characterized by a secondary fracture network with fracture density decreasing over depth until reaching

the unweathered bedrock [e.g., Dewandel *et al.*, 2011, Lachassagne *et al.*, 2021]. In this type of structure, the hydraulic conductivity of the fractured level is mainly controlled by the permeability of sub-horizontal fractures, which is higher than that of other fracture families, and which also goes with the usual development of sub-horizontal groundwater flow at the regional scale. For its part, the existence of a secondary fracture network is interpreted as a consequence of alteration processes [e.g., Lachassagne *et al.*, 2021, Worthington *et al.*, 2016]. The HRA scheme has been built on the basis of a large context of field observations issued from weathering profiles [review in Lachassagne *et al.*, 2021]. It typically concerns the first tens of meters beneath ground level [Wyns *et al.*, 1999] and is undoubtedly a relevant model of CZ architecture in silicate rock context. However, the HRA model, similar to the CZ concept, is not very detailed regarding the exact location of their lower bound, *i.e.*, the transition between the fissured horizon housing groundwater flow and the fractured bedrock underneath. Characterizing this boundary involves comparing, at the same location, the horizons where flow occurs and the structural and petrological data of the host rocks. This idea motivated in recent years diverse works to equip observatory sites of the CZ with deep boreholes of several tens of meters. Those should increase the cross-knowledge between the structure and architecture of the CZ over depth and the nature and types of water circulating within the CZ [e.g., Ayraud *et al.*, 2008, Dewandel *et al.*, 2011, 2017, Brantley *et al.*, 2013, Buss *et al.*, 2013, Orlando *et al.*, 2016, Chabaux *et al.*, 2017, Riebe *et al.*, 2017, Hahm *et al.*, 2019, Holbrook *et al.*, 2019, Ackerer *et al.*, 2021]. Such cross-information has also proved to be important for calibrating relatively recently developed hydro-

geophysical investigation methods to better understand the structure and hydrological functioning of the CZ [e.g., Grelle and Guadagno, 2009, Sailhac *et al.*, 2009, Befus *et al.*, 2011, Johnson *et al.*, 2012, Masson *et al.*, 2012, Holbrook *et al.*, 2014, Parsekian *et al.*, 2015, Orlando *et al.*, 2016, Lesparre *et al.*, 2020, Chaffaut *et al.*, 2023].

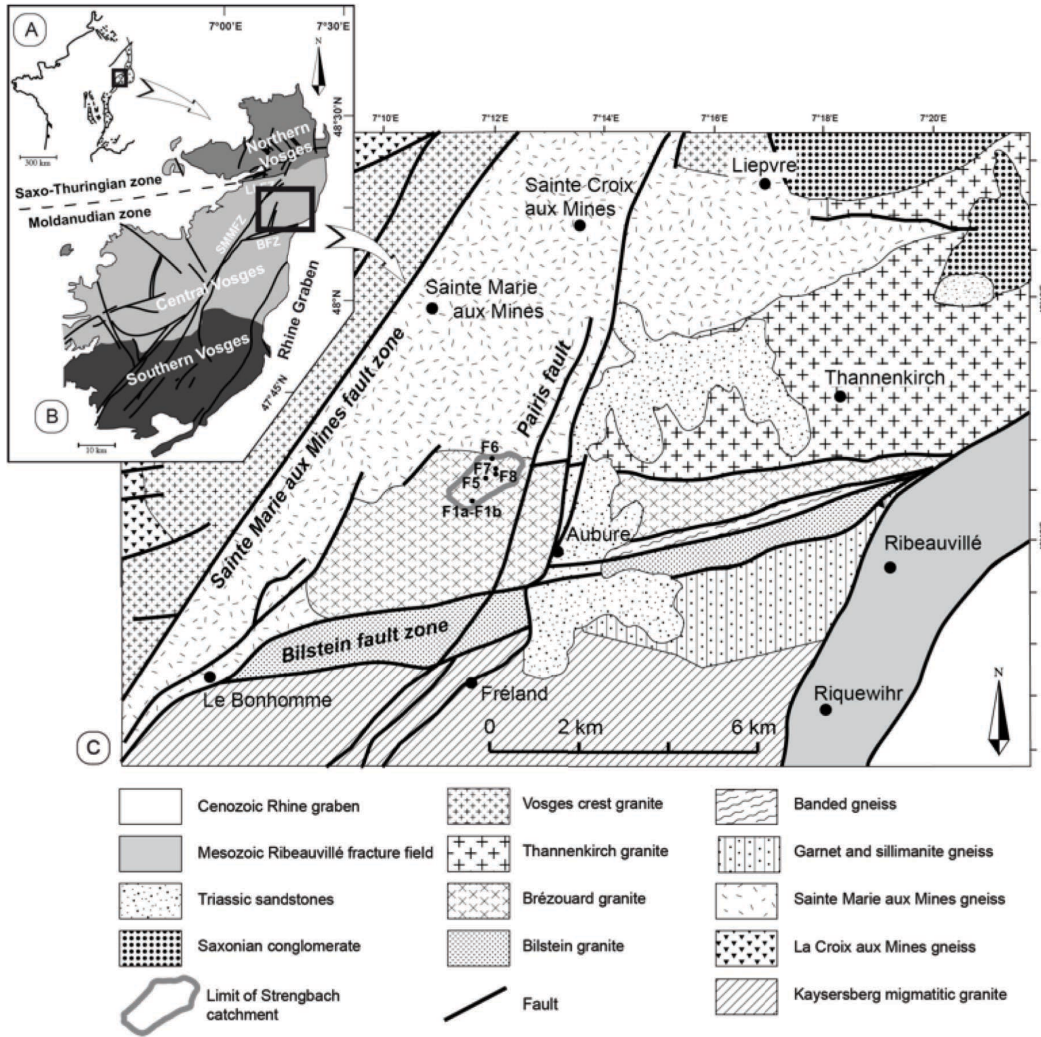
The Strengbach Critical Zone Observatory (CZO) is a fully equipped environmental observatory with meteorological, hydrological, and geochemical data recorded since 1986 (<https://ohge.unistra.fr>). It is one of the reference sites of the French OZCAR network (Critical Zone Observatories: Research and Application—<https://www.ozcar-ri.org>). Specifically, the Strengbach observatory partly targets the study of weathering processes on the basis of stream water, spring water, and soil analysis [Viville *et al.*, 2012, 2017, Pierret *et al.*, 2018, Chabaux *et al.*, 2019, and references therein]. It has been equipped with six 50–100 m-deep boreholes and three shallower 10–15 m-deep boreholes (hereinafter referred to as piezometers) in order to better understand the structure and circulation of water in the bedrock of the watershed. In the present study, we propose a synthesis of the main structural, petrological, and petrophysical data collected from the three deep-cored boreholes and the hydro-geophysical investigations acquired through the boreholes. These data, combined with those previously published on water chemistry and age [Chabaux *et al.*, 2017, Ranchoux *et al.*, 2021, Ackerer *et al.*, 2021], allow us to propose a structural architecture of the deeper zone developed in the granitic basement of the Strengbach watershed. A consequence would be to give a first conceptual scheme for the localization and nature of water circulations at a site representative of HRA in landscapes of mid-altitude mountains.

## 2. The Strengbach critical zone observatory

The Strengbach CZO, hereinafter referred to as the Strengbach watershed, relates to the uppermost part of the Strengbach catchment, a watershed covering 0.8 km<sup>2</sup>. Its location is upstream of Aubure village, ≈15 km from the western border of the Rhine valley on the western side of the Upper Rhine Graben in the Vosges Mountains (Figure 1).

The geological context of the Vosges Mountains consists of an arrangement of elongated major tectonic terranes that collided during the Hercynian orogeny between –340 Ma and –320 Ma [e.g., Fluck *et al.*, 1991, Lardeaux *et al.*, 2014, Skrzypek *et al.*, 2014, and references therein]. These terranes form three NE- to NNE-oriented main lithotectonic domains, known as the Northern, Central, and Southern Vosges (Figure 1). From a structural point of view, the Strengbach watershed belongs to the Central Vosges, which mainly exposes migmatites, gneisses, and granites with a complex multiphase tectonic history from the Hercynian orogeny to the formation of the current Tertiary–Quaternary reliefs. During the Hercynian collision, compressive structures induced crustal thickening associated with strike-slip faulting, followed by gravitational collapse and extensional tectonics. After a subsidence phase and Mesozoic sedimentation, the Rhine Graben opening during the Oligocene uplifted the basement in the rift shoulders. In response to the more recent Mio-Pliocene phases of alpine convergence, the Hercynian basement has been further exhumed in the Southern Vosges. As a result of this geological history, present-day regional bedrock structures trend ENE, N to NNE, and SW (Figure 1). During the Quaternary, the Vosges massif underwent erosion and weathering processes and was affected by glaciations, which decreased in impact northward together with topography [Ackerer *et al.*, 2022].

The Strengbach watershed is located at altitudes ranging between 880–1150 m asl. The watershed is marked by significant slope asymmetry, with a gentler north-facing slope (8–9°) and a south-facing slope around 15–16° (Figure 2). The climate is temperate, oceanic and mountainous with an average rainfall of 1400 mm yr<sup>-1</sup>. Forest covers 90% of the area with 80% spruce (*Picea abies* L.) and 20% beeches (*Fagus sylvatica*). The bedrock is mainly composed of Hercynian Ca-poor granite, fairly hydrothermally altered, with small microgranite and gneiss inclusions along the southern and northern slopes, respectively [Pierret *et al.*, 2018, Chabaux *et al.*, 2019, and references therein]. Compared with granite, gneiss is Mg-rich due to the higher proportion of biotite and the presence of chlorite [El Gh'Mari, 1995, Fichter *et al.*, 1998a,b]. The granitic bedrock has undergone late hydrothermal alterations, which are described as more important on the



**Figure 1.** Geological setting of Strengbach watershed. (A) European Cenozoic rift system and location of Vosges Mountains in the western shoulder of Rhine graben. (B) Simplified terrain map of Vosges Mountains. (C) Geological map of Strengbach watershed and surroundings [after Blanalt *et al.*, 1972a,b, Von Eller *et al.*, 1975, Ménillet *et al.*, 1978, Skrzypek *et al.*, 2008]. SMMZ: Saintes Marie aux Mines fault zone; BFZ: Bilstein fault zone.

northern slope of the watershed than on the southern slope [Fichter *et al.*, 1998a,b]. The Strengbach watershed was most probably marked by periglacial events and not fully covered by glaciers or ice caps [Ackerer *et al.*, 2016, Dehnert *et al.*, 2010, Heyman *et al.*, 2013]. Several meters of stripping at the summit area and colluvium deposits over the slopes occurred approximately 20 ky ago [Ackerer *et al.*, 2016, 2022]. The bedrock is covered by  $\approx 80\text{--}100$  cm-

thick soils ranging from Dystric Cambisol to Albic Podzol [brown acidic to ochreous brown podzolic, WRB, 2014]. The soils usually cover a sandy saprolite, reaching several meters in thickness in some places [Pierret *et al.*, 2018, Chabaux *et al.*, 2019]. Bedrock outcrops are scarce, except for a few bouldery tors.



### 3. Sampling and analytical methods

The drilling of six deep boreholes (50–120 m depth; F1a, F1b, F5, F6, F7, F8 in Figure 2b) and three piezometers ( $\approx 15$  m depth; pz3, pz5, pz7 in Figure 2b) was performed in 2012–2014. The nine boreholes are distributed along two transects, one along the south-facing slope and the other along the north-facing slope (Figure 2). F1b, F5, and F6 were cored and F6, F7, and F8 were entirely cased and equipped with multi-level screened sections. Acoustic and optical borehole wall images identify fracture zones at depth, leading to the installation of well screens (1.5 mm slots) in front of these fracture zones, considered as potential water circulation levels (five screened sections in F6, three sections in F7, two sections in F8). Sectional drawings of the six boreholes drilled in the Strengbach watershed and their main construction features are provided in Supplementary Section 1. A gravel pack (gravel diameter: 3–5 mm) was set up outside the screened sections to ensure proper water flow, and clay plugs of bentonite were used to isolate the screened sections. F5 and F1 remained uncased, except for the first 8 m, where a PVC casing and cement sheath were installed.

Physical–chemical loggings were regularly performed in the different wells with an OTT KL 010 TCM probe over the period 2015–2019 and an Idronaut Ocean Seven 303 multiparameter CTD probe in 2018. Electrical conductivity calculated at 25 °C ( $\mu\text{S}/\text{cm}$ ) and temperature (°C) were always measured, with the latter probe providing pH, redox potential (expressed in mV in reference to the potential of the Standard Hydrogen Electrode), and dissolved oxygen concentration (% of saturation) profiles in the water column of the wells. The presentation of the downhole logging approaches used in the study is given in Supplementary Section 2.1. We also determined the fracture number and orientation in boreholes F1, F5, F6, F7, and F8 from the analysis of acoustic and optical borehole wall images acquired in November 2012 and June 2014. Fracture orientations were also measured along scanlines in the two pits Fo1 and Fo3 (Figure 2), previously studied by Ackerer *et al.* [2016], Ackerer [2017]. The method is detailed in Supplementary Section 2.2. Rock samples representative of the principal rock types observed in the cores were collected along cores F1b, F5, and F6 for porosity and

permeability determination. Methods are detailed in Supplementary Section 2.3.

## 4. Results

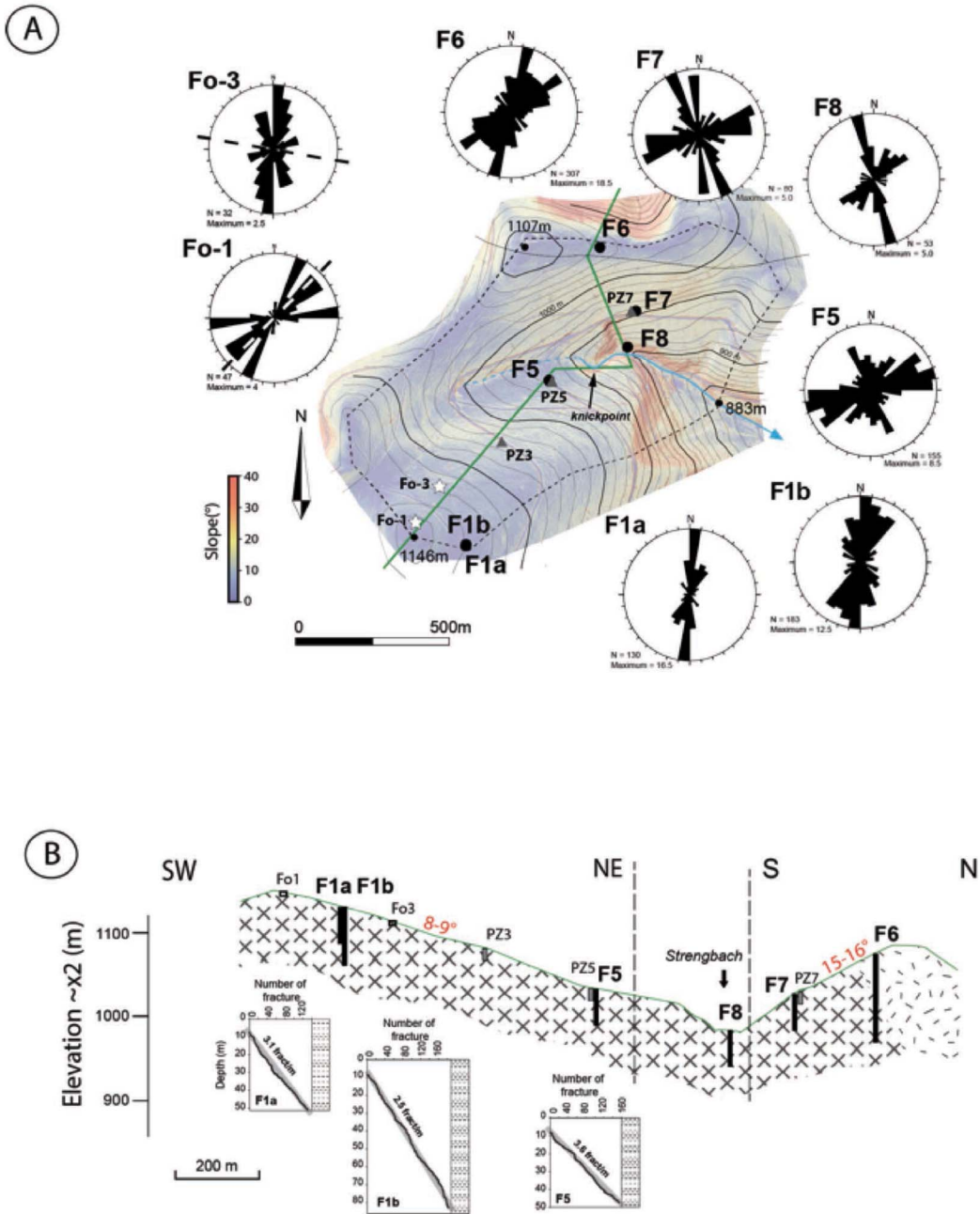
### 4.1. Petrological units

The three cores' macroscopic petrological characteristics are given in Supplementary Section 3. The principal information is summarized below and in Figure 3 and Supplementary Figure S1 (Supplementary Section 3).

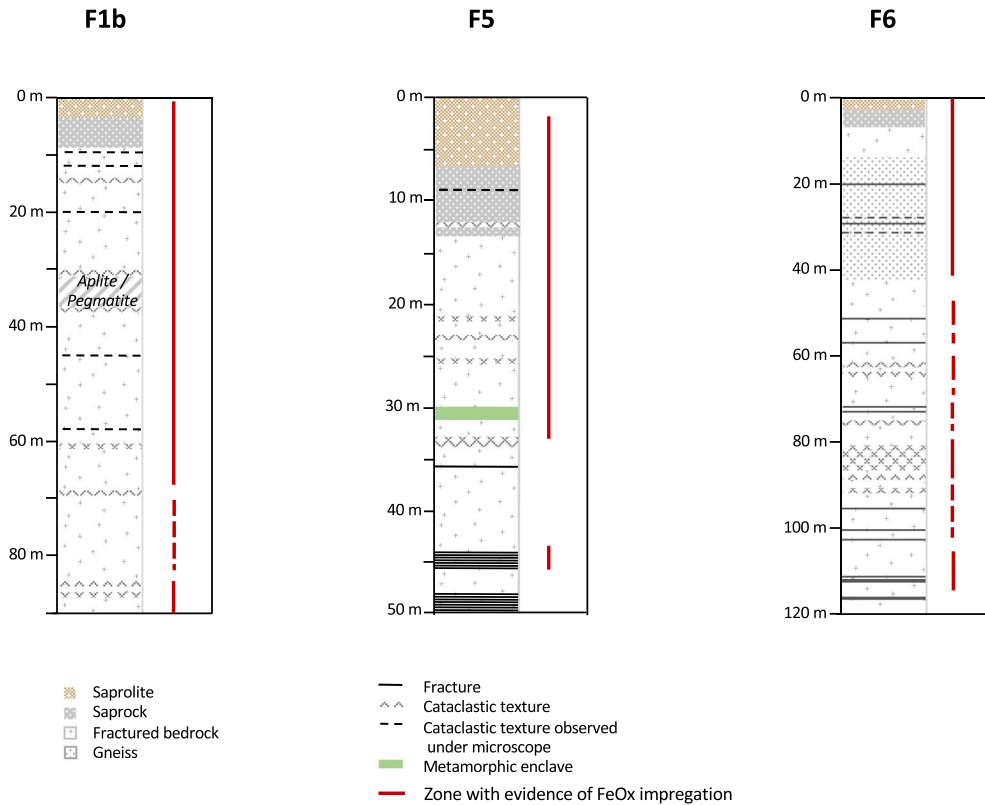
The data, associated with those previously published for pits Fo1 and Fo3 [Ackerer, 2017, Ackerer *et al.*, 2016], show that in pits Fo1 and Fo3 and boreholes F1b and F5, the bedrock is porphyry granite. In F1b, the granite is crosscut by decimetric aplite and pegmatite veins at 31–35 m and 71–72 m. Core F5 is marked by the presence of a metamorphic enclave at 30.66 m. F6, located on the northern slope, first crosscut the gneiss injected by granite down to 42 m depth and then a porphyry granite similar to that observed on the southern slope.

Based on the petrological description of the three cores given in Supplementary Section 3, the following main features can be drawn:

- The three cores can be divided into a soil horizon with limited vertical extension (0.8–1.2 m), an unconsolidated saprolite level, and an in-place fractured bedrock. The latter is subdivided into an upper level of a few meters with open sub-horizontal fractures that pass with depth to a more massive bedrock with much less or no visible open sub-horizontal fracturing. The soil and unconsolidated saprolite horizons constitute the so-called mobile regolith of the profile. The upper part of the in-place granite with open sub-horizontal fractures is usually named the saprock level and is identified as such in this study. The fractured bedrock term has thus been limited to the bottom section of the core below the saprock level.
- The fractured bedrock is more massive than the saprock with the presence of macroscopic fracture zones or non-horizontal cataclase zones of limited vertical extension. The zones of fractures and cataclases identified in the cores along with their main direction are



**Figure 2.** Strengbach watershed and location of the deep wells (Fx), piezometers (pzx) and pits (Fox). (A) Fracture direction measured in boreholes (Fx) and pits (Fox) located on simplified topographic map (contours every 10 m). The background is a slope map derived from an aerial 2018 Lidar survey (OHGE data). Knickpoint in valley bottom is indicated separating the western preserved upper part of the lower incised part of the watershed. Symmetrical rose diagram of direction represented with 10° classes, radius corresponds to 100% of maximum class, N: number of data; dashed line: direction of scanlines in the pits. (B) Geological cross-section with location of boreholes and pits, and fracture distribution along depth for F1a, F1b, and F5 (cumulative number of fractures vs depth). Bad image quality in other boreholes prevents relevant fracture distribution analysis.



**Figure 3.** Geological log of the three cored boreholes (F1b, F5, and F6) with localization of the main fractures and indications of the main zones of FeOx impregnation.

located in Figure 3. These zones evidenced current or recent water circulation. Outside such zones, the deep bedrock is slightly altered in appearance.

- Saprock and fractured bedrock horizons are characterized by mesoscale fractures clogged by secondary mineralogical infills. In both levels, these sealed fractures are intersected by the presently open fractures.
- The mobile regolith thickness increases along the southern slope of the basin, from  $\approx 1\text{--}2$  m upslope (Fo1) to  $\approx 9\text{--}10$  m downslope, close to the Strengbach stream. The saprock zone is less developed in core F1b (2–3.3 m) than in core F5 (9–12.5 m).
- In borehole F6, located at the top of the northern slope, the transition between mobile regolith and in-place gneiss of the saprock horizon is at 2.5 m depth, and the transition between saprock and fractured

gneiss horizons is at 3.3–3.5 m. F7, F8, and pz7, located along the same hillside, were not cored. However, as detailed in Supplementary Section 3, some observations suggested a mobile regolith thickness of  $\approx 2$  m, varying only slightly from top to bottom along the south-facing slope as opposed to observations from the north-facing slope.

- Colluvium horizons are observed in pit Fo3 [Ackerer, 2017, Ackerer et al., 2022] and, as detailed in Supplementary Section 3, in cores F5 and likely F1b, indicating a relatively large lateral extension of colluviums on the north-facing slope. The absence of pits or non-destructive coring along the south-facing slope made it impossible to assess whether or not the regolith of this hillside is also dominated by colluvium deposits.

#### 4.2. Fracture network: description and orientations

Borehole wall images are only available below 3–8 m depths, depending on the wells, essentially within the saprock and fractured bedrock zone. Borehole imagery combined with core observations shows that the basement of the Strengbach watershed is affected by a large fracture network, irrespective of depth (Figure 2; Supplementary Table S1). Two types of fractures are distinguished: the meso-fractures, which are thin fractures with an aperture less than 1 mm, and major fractures, which present thick or associated thin fractures in fractured zones with a thickness over 1 cm.

Most of the meso-fractures have a steep dip ( $>50^\circ$ ). They are present both in saprock and fractured bedrock and are the major ones in the fractured bedrock. Others with a  $<20^\circ$  dip are considered as sub-horizontal fractures. They are often sealed by iron hydroxides and occur as sub-parallel  $<4$ – $5$  cm long and  $<1$  mm-thick cracks. Open sub-horizontal fractures are mainly observed in the saprock horizon and only a few in the fractured bedrock. The largest fractures show mineral infilling, cataclastic texture, and associated wall-rock alteration (Figure 3). Orientations of the largest fractures and of steep-dipped meso-fractures give evidence of one to two main fracture sets, depending on the considered borehole (Figure 2). The main set S1 is oriented N–S to NNE–SSW, and the secondary set S2 is oriented NE–SW to E–W. S1 has an orientation similar to the direction of the Upper Rhine Graben (Rhenan direction), whereas S2 has an orientation similar to the Hercynian terrane directions (Figure 1c). These two main fracture directions are thus those of the regional tectonic orientations inherited from the geological history of the massif.

#### 4.3. Porosity and permeability

Porosity and permeability were measured in core samples from the boreholes F1b, F5, and F6 (Supplementary Table S2; Figure 4). The data pointed out a general decrease in porosity with depth. Porosity decreases from 20% to 0.6% in F5 when depth increases from 6 to 47 m. Even if the feature is less pronounced and assigned larger data dispersions, porosity decreases from 19% to 1.6% in F6 when the depth

increases from 14 m to 118 m. The range of porosity variation was smaller in F1b, dropping from 5% to 1% as depth increased from 2 m to 86 m, except at 30 m depth, where a singular high porosity value was measured. The relationship between permeability and depth is similar to that of porosity (Figure 4). Permeability decreases by three orders of magnitude as depth increases from 6 m to 47 m in F5. There is a similar decrease for F6 but over the whole depth range (ca. 100 m) and if the data corresponding to the fracture zones ( $\approx 60$  m) are not considered. Finally, there is almost uniform permeability over the investigated depth range for F1b, with two exceptions at 20 m and 86 m. The fracture measurements on the pit outcrops Fo1 and Fo3 indicate that the dominant fracture sets are roughly NE–SW ( $N20^\circ E$ – $N60^\circ E$ ) and E–W ( $N80^\circ E$ – $N90^\circ E$ ) oriented in Fo1 and NNE–SSW ( $N0^\circ E$ – $N30^\circ E$ ) in Fo3 (Figure 2b).

#### 4.4. Physical–chemical borehole fluid logging

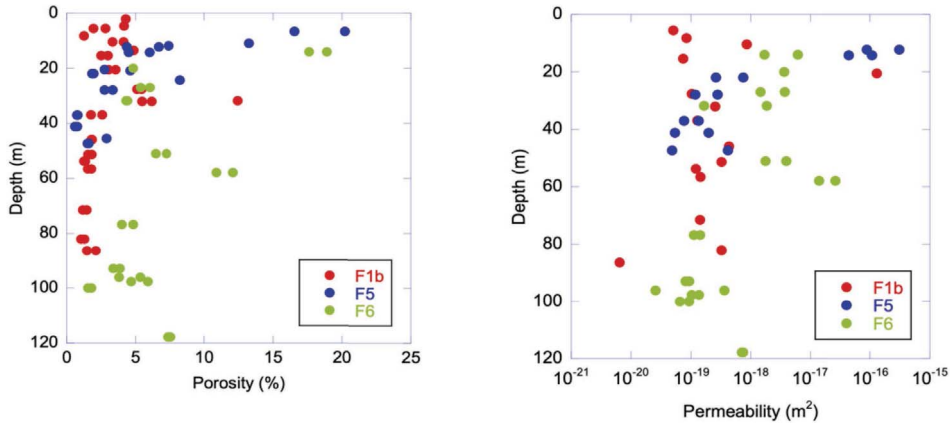
Several physical and chemical waterlogging campaigns have been performed since 2016, following the protocols given in Supplementary Section 2.2. The majority of the loggings were carried out under natural flow conditions and only a few after water extraction by pumping. The results of these logs are described below for borehole fluid temperature and electrical conductivity under natural flow conditions and after pumping.

##### 4.4.1. Temperature

Temperature profiles measured under natural flow conditions in boreholes and piezometers (Figure 5) show seasonal variations down to 20–25 m depth, whether the well is cased or not. Borehole F6, with its deeper water table level, is an exception as it was only slightly influenced by seasonal changes. Below 20–25 m depth (40 m depth for F6), temperatures were much more stable over time and increased with depth according to the regional geothermal gradient. These profiles reflect water columns in thermal equilibrium with their local environment and therefore indicate that water inflows are not sufficiently large to disrupt these temperature profiles at depth.

##### 4.4.2. Fluid electrical conductivity

The borehole fluid electrical conductivity profiles are given in Figure 6. They are more contrasted than



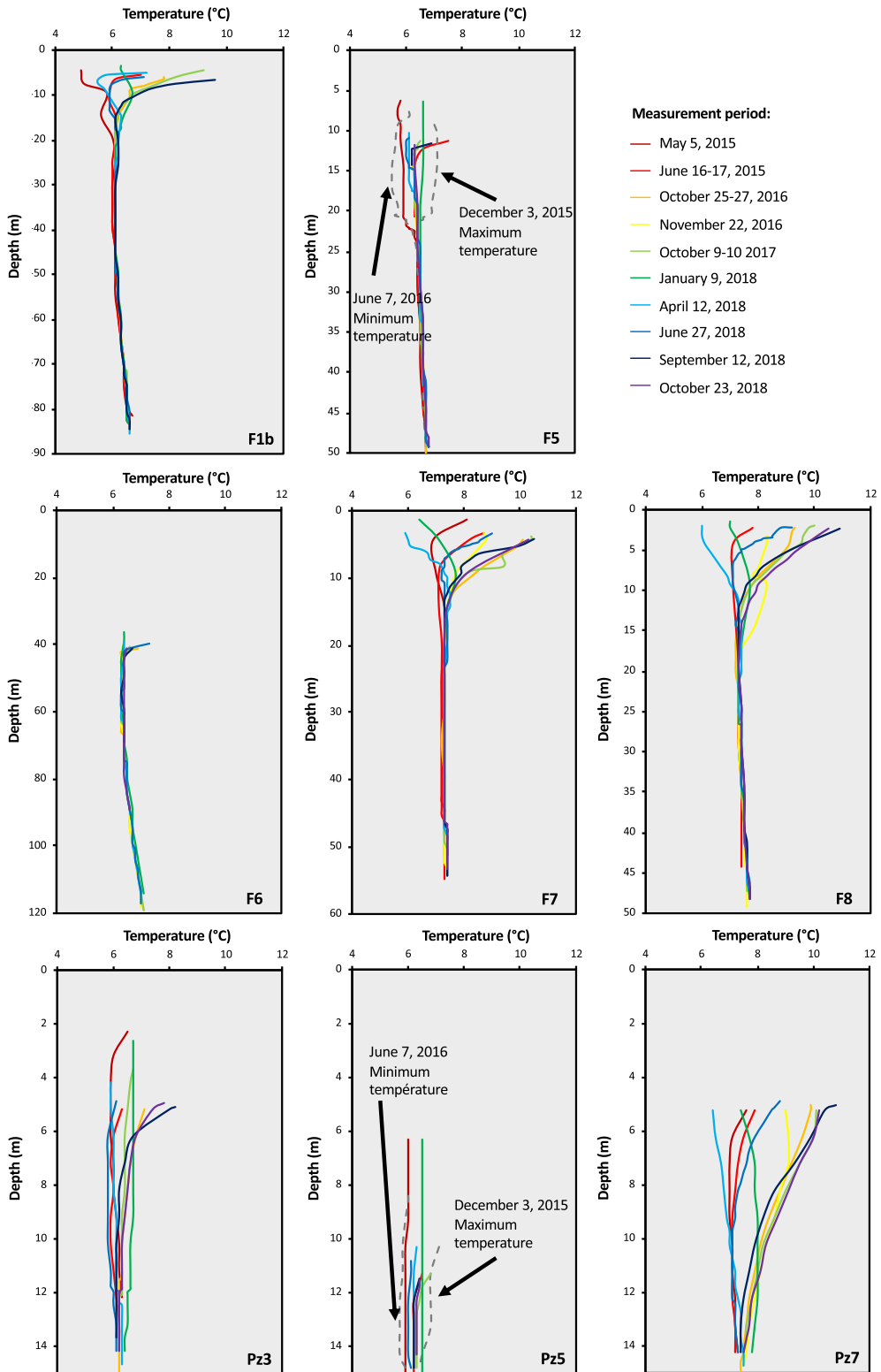
**Figure 4.** Depth variation of porosity and permeability. For F5, the porosities decrease progressively down to 14 m depth ( $\approx$  bottom of the saprock horizon) to reach low values at 20 m depth and beyond, except in the vicinity of macroscopic fracture or cataclastic fractured zones. In F1, the low saprock extension logically results in low porosity and permeability from 3–4 m depth. In F6, high porosity values are measured well below the saprock level, i.e., up to 31 m depth, which is in line with the high degree of fracturation of the gneiss bedrock horizon.

temperature profiles. A first group of boreholes includes F6, pz5 and pz7, where the electrical conductivity does not vary with depth, or slightly at the well as a result of the dissolution of accumulated material [Ranchoux, 2020]. Changes over time in electrical conductivity are limited everywhere. A second group includes all the other boreholes and pz3. Conductivity profiles show one or several increases with depth, e.g., from values near 30  $\mu\text{S}/\text{cm}$  (at 25  $^{\circ}\text{C}$ ) from 2 to 12 m depth up to  $>200$   $\mu\text{S}/\text{cm}$  from 12 m to 15 m in pz3. Changes of similar amplitude were observed in F5, with sudden increases varying in depth as a function of time, shallower when the static level was shallow (6.3 m and 12 m respectively), and deeper when the static level was deeper (11.6 m and 23 m depth, respectively). However, the interface depth is not a direct function of the water table level evolution. As F5 is largely uncased, such conductivity profiles suggest the existence of at least two contrasted water inflows along the borehole, with low-conductivity water similar to that measured in spring water at the surface and higher-conductivity water in the deeper part of the borehole.

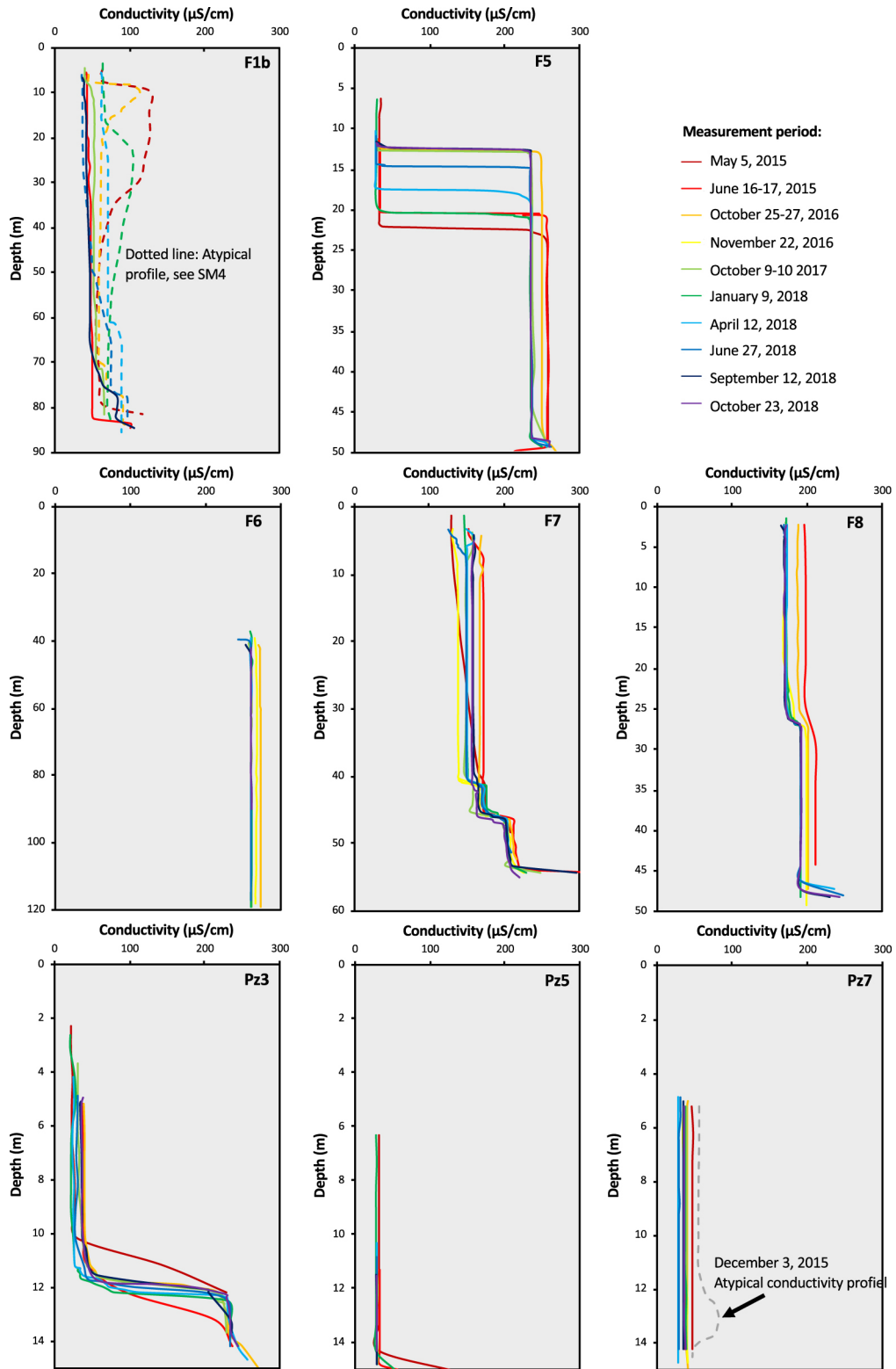
For F1b, when the static water level is high, an interval of higher conductivity often appears at the top of the water column. It is attributed to concrete leachate (see Supplementary Section 4). The conductivity profiles in F1b also suggested that this higher

level progressively migrated downhole over time and was therefore not diluted by lateral flow. This implies a low rate of water production in this borehole. This is confirmed by the weak local hydrological connection between F1b and F1a, 6 m away from each other, indicated by the absence of any significant variation in the piezometric level in F1b when F1a was purged. The conductivity in F1b is low compared to that measured in the deep F5 levels, but systematically higher than that of spring water or piezometer water. Moreover, two to three different conductivity zones were observed during some investigations, with conductivity increasing with depth (Supplementary Figure S2 in Supplementary Section 4), suggesting the occurrence of different deep-water inflows in F1b.

Contrary to F1b and F5, F6, F7, and F8 were equipped with a succession of cased and multi-level screened sections (Supplementary Section 1). Their water conductivity was high, e.g.  $>120$   $\mu\text{S}/\text{cm}$  for the lowest values of F7 to approximately 260–270  $\mu\text{S}/\text{cm}$  in F6 or the highest water conductivities in F7 (Figure 6). There are no low conductivity horizons and thus no surface waters infiltration in these boreholes due to the casing of the upper part. Only F6, equipped with four screened sections, has relatively uniform conductivity profiles throughout the water column. Such an observation can indicate that this borehole is mainly supplied by the water flow



**Figure 5.** Borehole fluid temperature loggings using the OTT KL 010 TCM probe.



**Figure 6.** Borehole fluid conductivity loggings using the OTT KL 010 TCM probe.

identified at 63 m depth during drilling and with no other significant contribution from other horizons. Borehole F7, equipped with three screened sections, has a conductivity profile marked by several steps, suggesting the existence of several water-productive horizons along the borehole, with at least two different production zones between 40 m and 47 m. Borehole F8, with two screened sections, shows small conductivity changes with depth that occurred at a depth between 25 m and 30 m (Figure 6), suggesting the occurrence of at least one specific water supply in each screened section.

#### 4.4.3. Influence of pumping on physical–chemical logs

The influence of pumping on the temperature and chemical profiles of the water column was evaluated by performing physical–chemical logs just after a pumping session. The results are compared to the logging results before pumping (Figure 7), the latter being considered representative of equilibrium conditions under natural flow. In addition to a decrease in the water table level, indicative of relatively low-productive water horizons (pumping rate  $<1.5 \text{ m}^3/\text{h}$ ), the temperature and chemical profiles change in all boreholes and depart from equilibrium conditions, with the occurrence of some anomalies. When the temperature peak is at the pump level, an artifact induced by the local pumping stress cannot be avoided (e.g., heating induced by the pump rotor). However, when located at different depths or associated with other chemical changes, the temperature anomalies are better related to water inflow from productive zones. In our study, temperature anomalies were often associated with conductivity anomalies, as shown in Figure 7. For example, this is the case for F6 at 62–63 m depth, where an increase of more than  $0.5 \text{ }^\circ\text{C}$  is measured. Second-order changes are observed from pH or dissolved oxygen concentrations. Based on such anomalies, a limited amount of water circulation zones are identified. These zones are systematically associated with fracture zones identified on the cores and borehole wall images. These zones are listed in Table 1.

Physico-chemical logging, performed sequentially after pumping, provides data on the return time to equilibrium. Figure 8 shows such acquisitions, performed only in F5, which is open below 9 m (i.e., above the static water level). All the water

column data correspond to flow conditions in open holes, with lateral flows coming from productive levels and eventually from vertical flows induced by temperature and/or salinity gradients. Three reference logs acquired under undisturbed conditions (natural flow conditions before pumping) are presented for April 2018 and June/July 2019 (Figure 8).

Pumping actions performed without packers severely disturbed the natural flow conditions, as shown by the conductivity log performed shortly after pumping on April 24, 2018. This perturbation was not rapidly resorbed, as highlighted by the conductivity log recorded two days later (April 26). Thus, several days were needed to recover steady-state conditions, as also evidenced by Ranchoux [2020] for similar monitoring made in 2015. A similar statement can be proposed regarding pH and even dissolved oxygen. The situation is different with temperature. The effect of pumping is shortly transient, probably because the fluid temperature is quickly buffered by the rock temperature. Perturbations lasting several days also suggest quite a reduced water production from the different production levels.

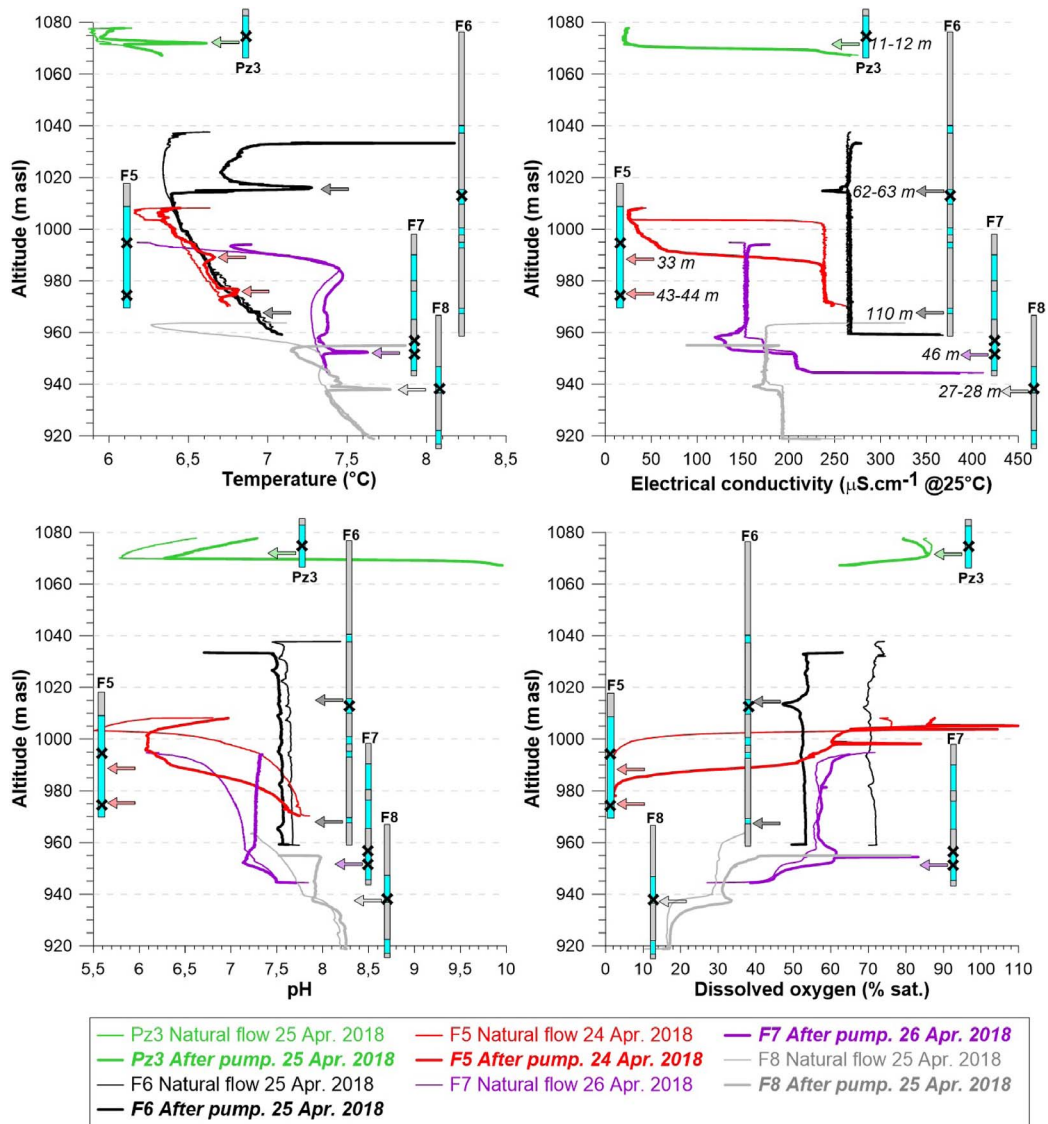
## 5. Discussion

The results from the Strengbach CZO lead to a comparison of the architecture of the watershed bedrock with the nature of the water flowing through it. Such a combined dataset, still scarce at this scale, is relevant to discuss the aquifer and water circulation level pattern in this watershed, representative of mid-mountain granitic basins.

### 5.1. Nature and vertical extension of the surface aquifer

Structural and petrological observations of the cores and boreholes combined with observations of the pits give a first view of the underground vertical structure in the Strengbach watershed. It consists of a loose or mobile saprolite, on top of which the soil develops and under which an in-place fractured bedrock zone extends over depth. The in-place bedrock is subdivided into a few meters of extended saprock level characterized by open sub-horizontal fractures that pass with depth to a more massive bedrock (the fractured bedrock level) with few or no visible open sub-horizontal fracturing. Nowhere is





**Figure 7.** Comparison of borehole fluid physico-chemical logs (temperature, electrical conductivity, pH, and dissolved oxygen using the Idronaut probe) prior to and after pumping. The geometry of the boreholes is indicated by the technical sections. The main productive water horizons are indicated by arrows.

a level of fresh bedrock reached without evidence of water circulation.

Water conductivity data evidence that the piezometer and surface waters from F5 have much lower conductivities than the F5 deep waters. Moreover, their conductivities are close, not to say identical, to those of spring waters. Such an observation indicates that both water types circulate in the same

aquifer system. The very similar geochemical and isotopic data of these waters [Chabaux *et al.*, 2017, Ackerer *et al.*, 2018, 2020, Ranchoux *et al.*, 2021], as well as the results of the modeling of the geochemical composition of the spring waters and the piezometers of the south-facing slope [Ackerer *et al.*, 2020], lead to the same conclusions. All these waters circulate within the same hydrological compartment

**Table 1.** Water flow levels determined from the physico-chemical loggings in the Strengbach watershed boreholes and relation with geological features deduced from borehole imagery and core observations

Borehole	Water flow depth interval (m)*	Used physico-chemical criteria	Geological observation
Pz3	11–12	Temperature and redox	No available observation
F5	15–16	Electrical conductivity	NE–SW and E–W fractures
F5	20–24	Dissolved oxygen	E–W fracture zone
F5	30–33	Temperature	Lithological discontinuity at 33 m
F5	43–44	Electrical conductivity	E–W fracture zone at 45 m
F6	62–63	Temperature, electrical conductivity and dissolved oxygen	Major fracture zone at 63.5–64.3 m (N72° E direction)
F6	77–80	Dissolved oxygen	Large N–S fracture at 79 m
F6	110	Temperature	NW–SE to E–W fracture zone between 113.3–114.4 m
F7	41–42	Electrical conductivity	NE–SW single fractures
F7	46	Temperature, pH and electrical conductivity	NE–SW single fractures
F8	21	Dissolved oxygen and electrical conductivity	E–W (N72° E) fracture at 20.7 m depth
F8	27–28	Electrical conductivity and temperature	NE–SW fracture zone between 25–29 m
F8	43–47	Dissolved oxygen	N55° E single fracture

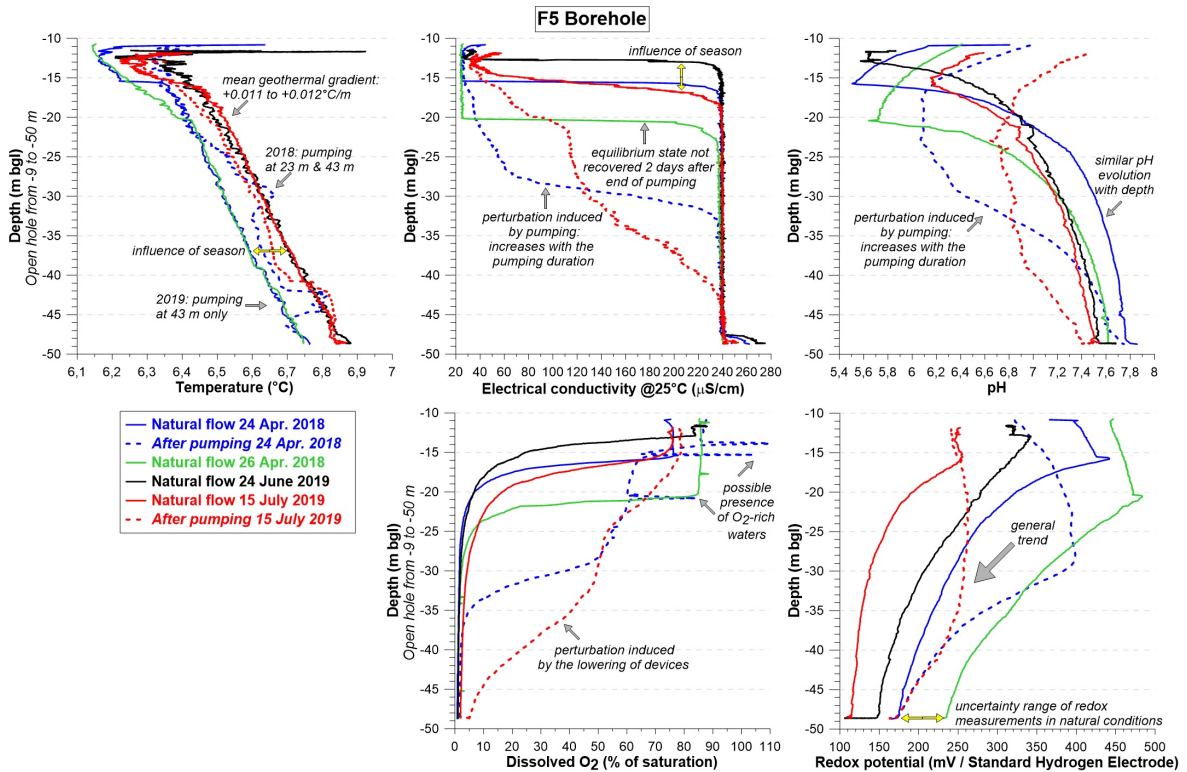
\* Depth interval as suggested by changes noticed on physico-chemical logs when comparing log under natural flow conditions and log shortly after pumping.

corresponding to the surface aquifer controlling the spring waters' chemical composition. Comparison of hydrological data with core and well data helps clarify the nature and vertical extent of this aquifer.

Borehole F5 was drilled in a 9 m-thick mobile regolith zone and a 3–4 m-thick saprock zone. It was sealed with PVC and cement pipe to a depth of 8–9 m. Borehole F5 can, therefore, only be fed by water circulating in the lower horizons of the mobile regolith or deeper. Observation of water of low conductivity in the upper part of the water column to depths up to 13 m depth or more, depending on the season and in the absence of wellbore flow, suggests that the water in the so-called shallow subsurface zone may still circulate below a depth of 10 m, that is, in the saprock. The measurement of very high porosity and permeability at 10 m depth and still high at 12–15 m depth in core F5 is consistent with this conclusion. The conductivity profiles measured in pz5, always with low values of the spring water type, are also consistent with this scheme. The pz5 was drilled

close to F5 down to a depth of 15 m, with a sealed casing over the first two meters, thus a priori only intersecting the mobile regolith–saprock level. In the core from F5 below these depths and particularly below 20 m, the bedrock is characterized by low porosity and permeability measurements, apart in the close vicinity of fracture zones, suggesting a different water circulation scheme in the fractured bedrock zone, as detailed in Section 5.3. Therefore, based on these different observations, the entire saprolite–saprock system can be interpreted as representing the watershed's shallow aquifer in the Strengbach critical zone observatory.

Field data show a thickening of the saprolite–mobile saprock system along the north-facing slope of the watershed and an a priori thinner thickness on the south-facing slope, at least in the vicinity of F7 and F8. Assimilation of this zone to the subsurface aquifer would, therefore, imply similar characteristics for the aquifer, which hydrogeochemical data confirmed. Pz3, set up at the mid-slope of the



**Figure 8.** Borehole fluid physico-chemical logs from the Idronaut probe in F5, presenting temperature, electrical conductivity, pH, dissolved oxygen, and redox changes prior to and after pumping.

north-facing slope, is located in a thinner saprolite-saprock zone than F5. Sealed over its first two meters, pz3 collects water flowing between 2 and 14 m depths. The conductivity profiles show low values ( $\approx 30 \mu\text{S}/\text{cm}$ ) of spring water-type up to  $\approx 10$  m depth and more conductive water deeper down. This suggests that this piezometer might intercept the deep circulation system in its lower part. The electrical conductivity profiles carried out after pumping in the piezometer pz3 confirm this characteristic as they show a fracture-type water inlet at 11–12 m depth (Figure 6). Borehole F1b is drilled further up the slope where the mobile regolith is thinner than in pz3 and is watertight over its first 8 m depth, i.e., over the entire zone of mobile regolith and saprock section of the core (Figure 3). Therefore, this well would only drain water circulating in the fractured bedrock. The higher water conductivities in F1b ( $\approx 50 \mu\text{S}/\text{cm}$  or above) than in spring water and piezometers (20–30  $\mu\text{S}/\text{cm}$ ) certainly validate that the waters in F1b are not part of the shallow circulation system but of

one with a deeper circulation. The observation of low porosity and permeability in samples from F1b (Figure 4) is consistent with this pattern since the samples used for these analyses were taken at depths of 2 m and more, thus essentially outside the shallow aquifer, explaining their low porosity and permeability values.

On the south-facing hillside, the casing equipment of F6 and F8 (see Supplementary Section 1) only allows for sampling waters below the saprock. Only pz7, sealed over the first 2 m depth, allows for surface water sampling. Low water conductivity measurements in pz7, similar to the conductivity of water from other piezometers and spring water, indicate that this well is mainly supplied by shallow subsurface water. The flow rate through this well seems much lower than in pz3 and pz5, based on the very rapid emptying of the well compared to the other two during pumping operations. This would imply a relatively weak circulation for the saprolite not occluded by the casing, consistent with the suggestion that on

this hillside, at least at the location of F7 and F8, the saprock's vertical extension is lower than on the opposite slope. Such a configuration would also imply that F7, uncased between 8 m and 18 m depth, is not supplied by surface water, as confirmed by the absence of any low conductivity water level measurement in this well, unlike F5.

All these observations confirm a pattern of meteoric water flow in a shallow aquifer, formed by mobile saprolite and saprock, with reasonably strong lateral variability in aquifer thickness at the watershed level. A more accurate geometry of the subsurface aquifer at the watershed scale will certainly be necessary to model hydrological flows at this scale. This will undoubtedly be achieved through geophysical investigation methods. Our results are interesting in providing a relatively simple and sound scheme for the vertical structure of the subsurface aquifer for the Strengbach watershed. This scheme implicitly assumes or confirms that the saprock level marked by subhorizontal open fractures certainly represents the propagation zone of the surface weathering front in the bedrock and thus marks the lower limit of the shallow aquifer of the Strengbach watershed.

### 5.2. *Spatial variation of the surface aquifer thickness*

Our data suggest a difference in the regolith thickness depending on slope orientation, as well as along the north-facing slope. Watersheds under geomorphological equilibrium or steady-state forcing conditions usually have greater mobile regolith thicknesses near the summit than in the talweg [Rempe and Dietrich, 2014]. This was not observed on the north-facing slope and may suggest that the watershed is far from equilibrium. The Lidar Digital Terrain Model (DTM) of the Strengbach watershed (Figure 2) and, especially, the derived slope map highlight the recent V-shaped stream carving of the valley bottom, incising the gentler slopes of the upper parts of the catchment. The well-defined slope breaks surrounding the deepest part of the valley suggest a regressive erosion of the Strengbach stream and that the watershed is not in a geomorphological steady state but on the way to re-equilibration.  $^{10}\text{Be}$  data validate such an interpretation [Ackerer *et al.*, 2022]. This interpretation implies that the watershed can be subdivided

into two sub-zones. First, a zone above the knick-point, preserving the topography inherited from erosion processes that occurred during the last glacial maximum and stabilized afterward. Second, a zone of active erosion below the slope break where the V-valley morphology is observed. Upstream and west of the system, where pits and boreholes F1–F5 are located, the watershed is a large valley with gentle slopes (8–9°) with little headward erosion. Downstream and further north, primarily where F7 and F8 are located, the valley is narrower with steeper slopes (15–16°) in response to stream carving. Based on these different observations, we propose to explain the differences in slope geometry and structure between the two hillsides by different erosion rates or processes between the two watershed slopes based on the fact that the Strengbach watershed area has not yet reached a geomorphological steady state.

### 5.3. *Deep water circulations and spatial variations of water conductivity*

The loggings in the fractured bedrock below the saprock horizon, especially those after pumping stimulations, identify specific and well-localized water inflows along the cores. When comparing the location of the water inflows to the description of the available cores (F1–F5–F6) or wall images, they are found to be systematically associated with or close to well-identified fracture zones or discontinuities (Table 1). These data, therefore, identify in the deep part of the bedrock a specific water circulation in the macroscopic fracture zones riddling the bedrock. Nevertheless, at this stage, it is not possible to eliminate a matrix circulation in the deep zone outside the macroscopic fracture zones. If such a matrix circulation exists, our data point out, however, that it would be much weaker than the circulations in the fracture network and undetectable based on the measurements presented in this work.

Our data also suggest a systematic variation of deep-water conductivity with water circulation depth in granitic bedrock. In F7 and F8, cored on the south-facing slope and only collecting water circulating in the fractured granitic bedrock, an increase with depth of water conductivity (150–205  $\mu\text{S}/\text{cm}$  in F7 and 170–200  $\mu\text{S}/\text{cm}$  in F8) was observed. Such a systematic variation can suggest that

flow paths or water travel times are longer when water circulates deeper in the bedrock. On the north-facing slope, F1b and F5, also drilled in the granitic bedrock, are open along their whole length except the upper 8 m. The equilibrium conductivity profiles of F5 imply feeding by a highly conductive deep water (Figure 6). Pumping experiments in F5 suggest circulation levels at  $\approx 23$  and  $\approx 32$  m depth (Figure 7), with intermediate conductivities between those of deep water and surface water, even if it is difficult to precisely characterize them. In F1b, for periods assumed to be free of pollution problems (details in Supplementary Section 4), two or three circulation levels with different and increasing conductivities from surface to depth can also be considered. Thus, these data follow the logic observed on the opposite slope, indicating that for the entire granitic basin, the deeper the fractures, the longer the path and/or the longer the water transit time in the fractures. With that kind of scheme, the lower water conductivities in F1b compared to the deep water in F5 would indicate that borehole F1b at the top of the slope is supplied with water of relatively short transit times compared to the F5 deep-water borehole. This would be logical in a context fed with local water, i.e., waters simply percolating from the catchment area and not imported by subsurface flow from farther areas. Only the water from borehole F6, cored at the top of the south-facing slope and mainly fed by a fracture at 63 m depth, does not follow the above pattern. Deep water in F6 is more conductive than that in F7 and F8, located further down the slope. It is tempting to link this difference to the lithological variability of the watershed, particularly as F6 intersects a significant thickness of gneiss, unlike the other wells in the watershed. Nevertheless, it cannot be excluded that this well, marked by a vast cataclase zone, could drain water from outside the watershed, unlike the other wells, even if the hydrogeochemical modeling of the deep water does not support such a scenario [Ackerer *et al.*, 2021].

Global water fluxes transiting through deep circulations cannot be quantified on the basis of our data. However, qualitative information can be derived from the partial emptying experiments of F5 (Figure 8). An extension of the low conductivity level toward the lower part of the water column during the borehole emptying, whatever the pumping depth, indicates that systematically, the water produced by

the deep horizon was preferentially abstracted during pumping and was not renewed at a rate corresponding to that of the pumping. This suggests that the production of more mineralized water from the deeper part of the well is low and that pumping induces the invasion of bottom-holes by less mineralized waters from shallower horizons. Such an observation is easily explained in the context of surface circulation marked by more intense water fluxes than in deep-water circulation systems. Similarly, the fact that it took a few days to bring the disturbed chemical profiles to an equilibrium profile, marked by a high electrical conductivity level rising to its pre-pumping level, indicates a lower intensity of the deeper water circulation than the subsurface one. A low circulation intensity in the deep fracture zones compared with the surface circulation is also consistent with the difficulty of detecting these flows by flowmetry for all wells. These methods could only identify a flow in the fractured zone associated with cataclastic textures in F6 at a depth of 63 m by identifying a low and not quantifiable intensity flow just slightly above the background noise of the flowmeter. Based on these data, we propose a circulation in the deep bedrock mainly through the fracture network that structures it, with lower intensities (this study) and much slower circulation times [Ranchoux *et al.*, 2021, Ackerer *et al.*, 2021] than those estimated for waters in subsurface aquifers. Our data show that the fracture network is developed along regional tectonic orientations inherited from the region's geologic history. Thus, it corresponds to a pre-existing fracture network, partially reopened. These results highlight the need to consider the inherited geological structures to understand the water circulation pathways in the deep CZ and, more generally, the structure and functioning of the CZ. The data also question the cause of the fracture opening. Is it an opening of clogged fractures that progressively propagate via the action of water infiltrating them? Or an opening prior to the circulations that secondarily offer a path to the shallower waters infiltrating them? Our data are inconclusive, but the hypothesis of a reopening of the fractures in response to rock decompression during the development of the Tertiary and Quaternary Vosges reliefs, or simply in connection with erosion processes, is entirely conceivable. In such a scenario, the vertical extension of these open fractures and thus of the deep circulations they control in

the bedrock could be much larger than the few tens of meters investigated by the boreholes. It may imply a water outlet much further downstream in the valley than the Strengbach CZO outlet. Independent of this point, our data also question the significance of this deep water with respect to the critical zone: is the deep water still influenced by surface circulations and therefore does it still belong to the critical zone? Or is ancient water that is decoupled from current surface processes and/or has another origin (marine, glacial...) as seen in deep hard rocks [Sanjuan *et al.*, 2016, Négrel *et al.*, 2001, 2018] or in more shallow waters as observed in the Armorican massif for instance [Aquilina *et al.*, 2015]. In the case of the Strengbach deep waters, our results and interpretation suggest that deep waters are fed by waters simply percolating from the catchment area. Water dating using CFC methods [Ranchoux *et al.*, 2021] shows ancient ages for the Strengbach borehole deep waters. However, these initial results do not allow us to choose between a pattern of ancient waters disconnected from surface waters and a pattern of mixing between ancient waters and present-day surface waters. Further investigations are needed to constrain the precise degree of connectivity of the deep waters with the surface waters, and therefore whether they should be considered as belonging to the critical zone or not.

## 6. Conclusion

This study integrates geological, petrophysical, and hydrogeological information from six 50–120 m-depth boreholes, complemented by three 15 m-depth piezometers drilled at the Strengbach Observatory site. The observation of the boreholes, in addition to the analysis of cores recovered for three of the six deep boreholes, renders the first detailed picture of the structure/architecture of the first  $\approx 100$  m of the CZ at the scale of the Strengbach CZO. The main features of hydrogeological functioning are characterized, and a simple conceptual model of water circulation can be proposed. It includes shallow water circulation within an aquifer corresponding to the geomorphological mobile saprolite horizon and the upper part of the saprock. This shallow circulation superimposes a deeper circulation in the fractured bedrock via the fracture networks riddling

this horizon. The data show that the vertical extension of the saprolite and saprock horizons, and thus of the shallow aquifer, differs for the two hillsides. We propose to relate these differences to an evolving erosion regime between the upstream and downstream portions on each side of a knickpoint in the main stream bed. The data also show that a deep bedrock level without evidence of water circulation is never reached even below 100 m depth. They also provide evidence that deep water circulations in the fractured bedrock horizon are controlled by fractures of regional orientation inherited from the geological history of the Vosges massif, integrating the Hercynian orogeny and the Tertiary and Quaternary tectonics. All these results show the importance of accounting for the geological heritage of the rocks. Correctly modeling the hydrogeochemical functioning of the CZ at the scale of a watershed requires an understanding of the bedrock emplacement's long-term geological history. It is also necessary to consider in detail the more recent history of the Quaternary erosion and alteration, incorporating it into the overall view of the system.

## Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

## Acknowledgments

The authors warmly thank the Guest Editors, especially Valérie Plagnes, for their advices and suggestions, which greatly helped to improve the article. They also thank the three anonymous reviewers for their constructive comments and suggestions. This study was financially supported by the French ANR program under grant agreement ANR-15-CE06-0014 (Project CANTARE-Alsace). Core drillings were funded by the REALISE program (CPER Etat Alsace 2003–2014) along with a funding contribution from INSU-CNRS and another one from the French Equipex program CRITEX. The contributions of the two OHGE responsables (M. C. Pierret and S. Cotel) are acknowledged.

## Supplementary data

Supporting information for this article is available on the journal's website under <https://doi.org/10.5802/crgeos.239> or from the author.

## References

- Ackerer, J. (2017). *Mécanismes et taux de dénudation d'un bassin versant élémentaire (Strengbach, France) : apport de l'étude couplée des méthodes de datation isotopique (déséquilibres U-Th-Ra,  $^{10}\text{Be}$  in situ) et des méthodes de modélisation hydrogéochimique (KIRMAT)*. Phd thesis, Université de Strasbourg, France.
- Ackerer, J., Chabaux, F., Lucas, Y., Clément, A., Fritz, B., Beaulieu, E., Viville, D., Pierret, M. C., Gangloff, S., and Négrel, Ph. (2018). Monitoring and reactive-transport modeling of the spatial and temporal variations of the Strengbach spring hydrochemistry. *Geochim. Cosmochim. Acta*, 225, 17–35.
- Ackerer, J., Chabaux, F., Van der Woerd, J., Viville, D., Pelt, E., Kali, E., Lerouge, C., Ackerer, P., Di Chiara, R., and Négrel, P. (2016). Regolith evolution on the millennial timescale from combined U-Th-Ra isotopes and in situ cosmogenic  $^{10}\text{Be}$  analysis in a weathering profile (Strengbach catchment, France). *Earth Planet. Sci. Lett.*, 453, 33–43.
- Ackerer, J., Jeannot, B., Delay, F., Weill, S., Lucas, Y., Fritz, B., Viville, D., and Chabaux, F. (2020). Crossing hydrological and geochemical modeling to understand the spatiotemporal variability of water chemistry in a headwater catchment (Strengbach, France). *Hydrol. Earth Syst. Sci.*, 24, 3111–3133.
- Ackerer, J., Ranchoux, C., Lucas, Y., Viville, D., Clément, A., Fritz, B., Lerouge, C., Schäfer, G., and Chabaux, F. (2021). Investigating the role of deep weathering in critical zone evolution by reactive transport modeling of the geochemical composition of deep fracture water. *Geochim. Cosmochim. Acta*, 312, 257–278.
- Ackerer, J., Van der Woerd, J., Meriaux, A. S., Ranchoux, C., Schäfer, G., Delay, F., and Chabaux, F. (2022). Quantifying geomorphological evolution from  $^{10}\text{Be}$  denudation rates: Insights from high-resolution depth profiles, topsoils, and stream sediments (Strengbach CZO, France). *Earth Surf. Process. Landf.*, 47(14), 3239–3258.
- Aquilina, L., Vergnaud-Ayraud, V., Les Landes, A. A., Pauwels, H., Davy, P., Pételet-Giraud, E., Labasque, T., Roques, C., Chatton, E., Bour, O., Ben Maamar, S., Dufresne, A., Khaska, M., Le Gal La Salle, C., and Barbécot, F. (2015). Impact of climate changes during the last 5 million years on groundwater in basement aquifers. *Sci. Rep.*, 5(1), article no. 14132.
- Ayraud, V., Aquilina, L., Labasque, T., Pauwels, H., Molenat, J., Pierson-Wickmann, A. C., Durand, V., Bour, O., Tarits, C., Le Corre, P., Fourre, E., Merot, P., and Davy, P. (2008). Compartmentalization of physical and chemical properties in hard-rock aquifers deduced from chemical and groundwater age analyses. *Appl. Geochem.*, 23(9), 2686–2707.
- Befus, K. M., Sheehan, A. E., Leopold, M., Anderson, S. P., and Anderson, R. S. (2011). Seismic constraints on critical zone architecture, Boulder Creek watershed, Front Range, Colorado. *Vadose Zone J.*, 10(4), article no. 1342.
- Blanalt, J. G., Von Eller, J. P., Fluck, P., Geffroy, J., Hirlemann, G., Jehl, M., Bonnet, Ch., Schreiner, J., Sittler, C., Tricart, J., Tricart, C., Hirth, C., Voght, H., and Weil, R. (1972a). *Colmar Artolsheim, Notice explicative de la Carte géologique de la France à 1/50 000, 342*. Bur. Rech. Géol. Min., Orléans. 55 p.
- Blanalt, J. G., Von Eller, J. P., Fluck, P., Geffroy, J., Jeanette, D., Ruhland, M., Schwoerer, P., Thévenin, A., Stieber, A., and Vogt, H. (1972b). *Sélestat, Notice explicative de la Carte géologique de la France à 1/50 000, 307*. Bur. Rech. Géol. Min., Orléans. 47 p.
- Brantley, S. L., Holleran, M. E., Jin, L., and Bazilevskaya, E. (2013). Probing deep weathering in the Shale Hills Critical Zone Observatory, Pennsylvania (USA): the hypothesis of nested chemical reaction fronts in the subsurface. *Earth Surf. Process. Landf.*, 38(11), 1280–1298.
- Buss, H. L., Brantley, S. L., Scatena, F. N., Bazilevskaya, E. A., Blum, A., Schulz, M., Jiménez, R., White, A. F., Rother, G., and Cole, D. (2013). Probing the deep critical zone beneath the Luquillo Experimental Forest, Puerto Rico. *Earth Surf. Process. Landf.*, 38, 1170–1186.
- Chabaux, F., Stille, P., Prunier, J., Gangloff, S., Lemarchand, D., Morvan, G., Négrel, J., Pelt, E., Pierret, M. C., Rihs, S., Schmitt, A. D., Trémolières, M., and Viville, D. (2019). Plant-soil-water interactions: Implications from U-Th-Ra isotope analysis in soils, soil solutions and vegetation (Strengbach CZO, France). *Geochim. Cosmochim. Acta*, 259, 188–210.
- Chabaux, F., Viville, D., Lucas, Y., Ackerer, J., Ran-

- choux, C., Bosia, C., Pierret, M. C., Labasque, T., Aquilina, L., Wyns, R., Lerouge, C., Dezaye, C., and Negrel, Ph. (2017). Geochemical tracing and modeling of surface and deep water–rock interactions in elementary granitic watersheds (Strengbach and Ringelbach CZOs, France). *Acta Geochim.*, 36, 363–366.
- Chaffaut, Q., Hinderer, J., Masson, F., Viville, D., Bernard, J.-D., Cotel, S., Pierret, M.-C., Lesparre, N., and Jeannot, B. (2023). Continuous monitoring with a superconducting gravimeter as a proxy for water storage changes in a mountain catchment. In *Beyond 100: The Next Century in Geodesy*, volume 152 of *International Association of Geodesy Symposia*, pages 261–267. Springer, Cham.
- de Marsily, G. (2004). *Cours d'hydrologie*. Université Pierre et Marie Curie, France. 220p.
- Dehnert, A., Preusser, F., Kramers, J. D., Akcar, N., Kubik, P. W., Reber, R., and Schluechter, C. (2010). A multi-dating approach applied to proglacial sediments attributed to the Most Extensive Glaciation of the Swiss Alps. *Boreas*, 39, 620–632.
- Dewandel, B., Jeanpert, J., Ladouche, B., Join, J. L., and Maréchal, J. C. (2017). Inferring the heterogeneity, transmissivity and hydraulic conductivity of crystalline aquifers from a detailed water-table map. *J. Hydrol.*, 550, 118–129.
- Dewandel, B., Lachassagne, P., Zaidi, F. K., and Chandra, S. (2011). A conceptual hydrodynamic model of a geological discontinuity in hard rock aquifers: example of a quartz reef in granitic terrain in South India. *J. Hydrol.*, 405(3–4), 474–487.
- El Gh'Mari, A. (1995). *Etude minéralogique, pétrophysique et géochimique de la dynamique d'altération d'un granite soumis aux dépôts atmosphériques acides (bassin versant du Strengbach, Vosges, France): mécanismes, bilans et modélisations*. Thèse, Université Louis Pasteur de Strasbourg. 199 p.
- Fichter, J., Turpault, M.-P., Dambrine, E., and Ranger, J. (1998a). Localization of base cations in particle size fractions of acid forest soils\_Vosges Mountains, N-E France. *Geoderma*, 82, 295–314.
- Fichter, J., Turpault, M.-P., Dambrine, E., and Ranger, J. (1998b). Mineral evolution of acid forest soils in the Strengbach catchment (Vosges mountains, N-E France). *Geoderma*, 82, 315–340.
- Fluck, P., Piqué, A., Schneider, J. L., and Whitechurch, H. (1991). Le socle Vosgien. *Sci. Géol. Bull.*, 44(3/4), 207–235.
- Grelle, G. and Guadagno, F. M. (2009). Seismic refraction methodology for groundwater level determination: Water seismic index. *J. Appl. Geophys.*, 68, 301–320.
- Hahm, W. J., Rempe, D. M., Dralle, D. N., Dawson, T. E., Lovill, S. M., Bryk, A. B., Bish, D. L., Schieber, J., and Dietrich, W. E. (2019). Lithologically controlled subsurface critical zone thickness and water storage capacity determine regional plant community composition. *Water Resour. Res.*, 55(4), 3028–3055.
- Heyman, B. M., Heyman, J., Fickert, T., and Harbor, J. M. (2013). Paleo-climate of the central European uplands during the last glacial maximum based on glacier mass-balance modeling. *Quat. Res.*, 79(1), 49–54.
- Holbrook, W. S., Marcon, V., Bacon, A. R., Brantley, S. L., Carr, B. J., Flinchum, B. A., Richter, D. D., and Riebe, C. S. (2019). Links between physical and chemical weathering inferred from a 65-m-deep borehole through Earth's critical zone. *Sci. Rep.*, 9, 1–11.
- Holbrook, W. S., Riebe, C. S., Elwaseif, M. L., Hayes, J., Basler-Reeder, K. L., Harry, D., Basler-Reeder, K., Malazian, A., Dosseto, A., Hartsough, P. C. W., and Hopmans, J. (2014). Geophysical constraints on deep weathering and water storage potential in the Southern Sierra Critical Zone Observatory. *Earth Surf. Process. Landf.*, 39(3), 366–380.
- Johnson, T. C., Slater, L. D., Ntarlagiannis, D., Day-Lewis, F. D., and Elwaseif, M. (2012). Monitoring groundwater-surface water interaction using time-series and time-frequency analysis of transient three-dimensional electrical resistivity changes. *Water Resour. Res.*, 48, article no. W07506.
- Lachassagne, P., Dewandel, B., and Wyns, R. (2021). Hydrogeology of weathered crystalline/hard-rock aquifers—guidelines for the operational survey and management of their groundwater resources. *Hydrogeol. J.*, 29(8), 2561–2594.
- Lardeaux, J. M., Schulmann, K., Faure, M., Janousek, V., Lexa, O., Skrzypek, E., Edel, J. B., and Stipska, P. (2014). The Moldanubian Zone in the French Massif Central, Vosges/Schwarzwald and Bohemian Massif revisited: differences and similarities. In Schulmann, K., Martínez Catalán, J. R., Lardeaux, J. M., Janousek, V., and Oggiano, G., editors, *The Variscan Orogeny: Extent, Timescale and the For-*



- mation of the European Crust*, Geological Society, London, Special Publications 405, pages 7–44. Geological Society of London.
- Lesparre, N., Girard, J. F., Jeannot, B., Weill, S., Dumont, M., Boucher, M., Viville, D., Pierret, M. C., and Delay, F. (2020). Magnetic resonance sounding measurements as posterior information to condition hydrological model parameters: Application to a hard-rock headwater catchment. *J. Hydrol.*, 587, article no. 124941.
- Masson, F., Viville, D., Pierret, M.-C., Mouyen, M., and Chabaux, F. (2012). Time-lapse microgravity study of the Strengbach catchment (Vosges mountains, France). *C. R. Géosci.*, 344, 357–365.
- Ménillet, F., Fluck, P., Flageollet, J.-C., Maïaux, C., and Lougnon, J. (1978). *Gérardmer. Notice explicative de la Carte géologique de la France à 1/50 000*, 341. Bur. Rech. Géol. Min., Orléans. 73 p.
- Négrel, Ph., Casanova, J., and Aranyossy, J. F. (2001). Strontium isotope systematics used to decipher the origin of groundwaters sampled from granitoids: the Vienne case (France). *Chem. Geol.*, 177, 287–308.
- Négrel, Ph., Pauwels, H., and Chabaux, F. (2018). Characterizing multiple water-rock interactions in the critical zone through Sr-isotope tracing of surface and groundwater. *Appl. Geochem.*, 93, 102–112.
- Orlando, J., Comas, X., Hynek, S. A., Buss, H. L., and Brantley, S. L. (2016). Architecture of the deep critical zone in the Rio Icacos watershed (Luquillo Critical Zone Observatory, Puerto Rico) inferred from drilling and ground penetrating radar (GPR). *Earth Surf. Process. Landf.*, 41(13), 1826–1840.
- Parsekian, A. D., Singha, K., Minsley, B. J., Holbrook, W. S., and Slater, L. (2015). Multiscale geophysical imaging of the critical zone. *Rev. Geophys.*, 53(1), 1–26.
- Pierret, M.-C., Cotel, S., Ackerer, P., Beaulieu, E., Benarioumlil, S., Boucher, M., Boutin, R., Chabaux, F., Delay, F., Fournet, C., Friedmann, P., Fritz, B., Gangloff, S., Girard, J.-F., Legtchenko, A., Viville, D., Weill, S., and Probst, A. (2018). The Strengbach catchment: A multidisciplinary environmental sentry for 30 years. *Vadose Zone J.*, 17, article no. 180090.
- Ranchoux, C. (2020). *Caractérisation géochimique et datation des circulations d'eaux profondes dans la zone critique : cas du Bassin versant du Strengbach*. Phd thesis, Université de Strasbourg, France.
- Ranchoux, C., Chabaux, F., Viville, D., Labasque, T., Lucas, Y., Van der Woerd, J., Ackerer, J., and Aquilina, L. (2021). Characterization of groundwater circulations in a headwater catchment from an analysis of chemical concentrations, Sr-Nd-U isotope ratios, and CFC, SF6 gas tracers (Strengbach CZO, France). *Appl. Geochem.*, 131, article no. 105030.
- Rempe, D. M. and Dietrich, W. E. (2014). A bottom-up control on fresh-bedrock topography under landscapes. *Proc. Nat. Acad. Sci. USA*, 111(18), 6576–6581.
- Riebe, C. S., Hahm, W. J., and Brantley, S. L. (2017). Controls on deep critical zone architecture: A historical review and four testable hypotheses. *Earth Surf. Process. Landf.*, 42(1), 128–156.
- Sailhac, P., Bano, M., Behaegel, M., Girard, J. F., Para, E. F., Ledo, J., Marquis, G., Matthey, P. D., and Ortega-Ramrez, J. (2009). Characterizing the vadose zone and a perched aquifer near the Vosges ridge at the La Soutte experimental site, Obernai, France. *C. R. Géosci.*, 341(10–11), 818–830.
- Sanjuan, B., Millot, R., Innocent, C., Dezayes, C., Scheiber, J., and Brach, M. (2016). Major geochemical characteristics of geothermal brines from the Upper Rhine Graben granitic basement with constraints on temperature and circulation. *Chem. Geol.*, 428, 27–47.
- Skrzypek, E., Cruz-Mermy, D., Ménillet, F., and Chèvremont, P. (2008). Unified geological map of Haut-Rhin (68)-Explanatory notes. Rapport BRGM/RP-56029-FR, 325 pages, in French.
- Skrzypek, E., Schulmann, K., Tabaud, A.-S., and Edel, J.-B. (2014). Palaeozoic evolution of the Variscan Vosges Mountains. In Schulmann, K., Martínez Catalán, J. R., Lardeaux, J. M., Janoušek, V., and Oggiano, G., editors, *The Variscan Orogeny: Extent, Timescale and the Formation of the European Crust*, Geological Society, London, Special Publications 405, pages 45–75. Geological Society of London.
- Viville, D., Chabaux, F., Pierret, M. C., Gangloff, S., Cotel, S., Probst, A., Dambrine, E., Fritz, B., and Ambroise, B. (2017). L'observatoire Hydro-Géochimique de l'Environnement (OHGE): 30 années d'observations scientifiques sur le bassin versant du Strengbach (massif des Vosges – France). *Géologues*, 195, 120–125.

- Viville, D., Chabaux, F., Stille, P., Pierret, M. C., and Gangloff, S. (2012). Erosion and weathering fluxes in granitic basins: example of the Strengbach catchment (Vosges massif, Eastern France). *Catena*, 92, 122–129.
- Von Eller, J. P., Ménillet, F., Hollinger, J., Fluck, P., and Maïaux, C. (1975). *St. Dié. Notice explicative de la Carte géologique de la France à 1/50 000*. 306. Bur. Rech. Géol. Min., Orléans. 55 p.
- Worthington, S. R. H., Davies, G. J., and Alexander Jr., E. C. (2016). Enhancement of bedrock permeability by weathering. *Earth-Sci. Rev.*, 160, 188–202.
- WRB (2014). World reference base for soil resources 2014. International soil classification system for naming soils and creating legends for soil maps. Update 2015. <http://www.fao.org/3/a-i3794e.pdf>.
- Wyns, R., Gourry, J.-C., Baltassat, J.-M., and Lebert, F. (1999). Caractérisation multiparamètres des horizons de subsurface (0–100 m) en contexte de socle altéré [Multiparameter characterization of subsurface horizons (0–100 m) in a weathered basement context]. *PANGEA*, 31(32), 51–54.



Research article

Geo-hydrological Data & Models

# Water cycle modelling strengthened by probabilistic integration of field data for groundwater management of a quite unknown tropical volcanic hydrosystem

Marc Dumont<sup>\*,a,b</sup>, Valérie Plagnes<sup>\*,a</sup>, Patrick Lachassagne<sup>\*,c</sup>, Roger Guérin<sup>\*,a</sup>, Bayu Nugraha<sup>d</sup>, Febriwan Mohamad<sup>\*,d</sup>, Ludovic Oudin<sup>\*,a</sup>, Arif Fadillah<sup>e</sup>, Danièle Valdès<sup>\*,a</sup>, Gilles Brocard<sup>\*,f</sup>, Jean-Luc Bonjour<sup>g</sup>, Mohamed Saadi<sup>\*,a,h</sup>, Anne-Sophie Esneu<sup>a</sup>, Aswar Muhammad<sup>e</sup>, Hendarmawan<sup>d</sup> and Nathalie Dörfliger<sup>\*,g</sup>

<sup>a</sup> Sorbonne Université, CNRS, EPHE, UMR 7619 METIS, F-75005 Paris, France

<sup>b</sup> Urban and Environmental Engineering, Applied Geophysics, University of Liege, B-4000 Liege, Belgium

<sup>c</sup> HSM, Univ. Montpellier, CNRS, IRD, IMT Mines Alès, Montpellier, France

<sup>d</sup> Faculty of Geological Engineering, Universitas Padjadjaran, Jatinangor, Sumedang 45363, Indonesia

<sup>e</sup> Danone Aqua group, Department of Water Resources, Jakarta, Indonesia

<sup>f</sup> Archéorient, Maison de l'Orient et de la Méditerranée, Université Louis Lumière-Lyon 2, France

<sup>g</sup> Water Institute by Evian, Water Resources and Sustainability Division, Danone Waters, Evian-les-Bains, France

<sup>h</sup> Institut de Mécanique des Fluides de Toulouse (IMFT)—Université de Toulouse, CNRS-INPT-UPS, Toulouse, France

*E-mails:* mdumont@uliege.be (M. Dumont), valerie.plagnes@upmc.fr (V. Plagnes), patrick.lachassagne@umontpellier.fr (P. Lachassagne), roger.guerin@sorbonne-universite.fr (R. Guérin), geo.bayunugraha@gmail.com (B. Nugraha), febrivan@unpad.ac.id (F. Mohamad), ludovic.oudin@sorbonne-universite.fr (L. Oudin), arif.fadillah@danone.com (A. Fadillah), danièle.valdes\_lao@upmc.fr (D. Valdès), gilles.brocard@mom.fr (G. Brocard), jeanluc.bonjour@danone.com (J.-L. Bonjour), mohamed.saadi@toulouse-inp.fr (M. Saadi), anne-sophie.esneu@ifp.fr (A.-S. Esneu), azwar.muhammad@danone.com (A. Muhammad), hendarmawan@unpad.ac.id (Hendarmawan), nathalie.dorfliger@danone.com (N. Dörfliger)

\* Corresponding author.

**Abstract.** Andesitic volcanic hydrosystems in Indonesia are mostly hydrogeologically unknown despite their socio-economic importance. The development of robust and easy-to-implement methodologies to conceptualize and quantify the water cycle components becomes a prerequisite for their sustainable management.

We developed a lumped hydrological model to mimic the structure and functioning of a previously unknown hydrosystem located on the flanks of the Salak volcano (West Java). The structure of the aquifers was revealed with electrical resistivity tomography. The distinction between springs fed by the extensive artesian aquifer and others fed by shallow perched aquifers was obtained mostly using hydrochemistry. The elevation of the recharge area was identified using isotopic analysis of spring water.

After designing the hydrological model structure, we carried out a probabilistic parameters exploration using the multiple-try differential evolution adaptive Metropolis algorithm to calibrate aquifer discharge. Multiple Markov chains allow a better exploration of the parameter values. The Bayesian approach provides the best water cycle simulation with a parameter uncertainty analysis, improving the accuracy and representation of the water cycle appropriate for previously unknown hydrosystems.

**Keywords.** Groundwater, Groundwater management, Multidisciplinary approach, Lumped model, Probabilistic, Andesitic setting.

*Manuscript received 6 April 2022, revised 13 July 2022 and 17 October 2022, accepted 28 November 2022.*

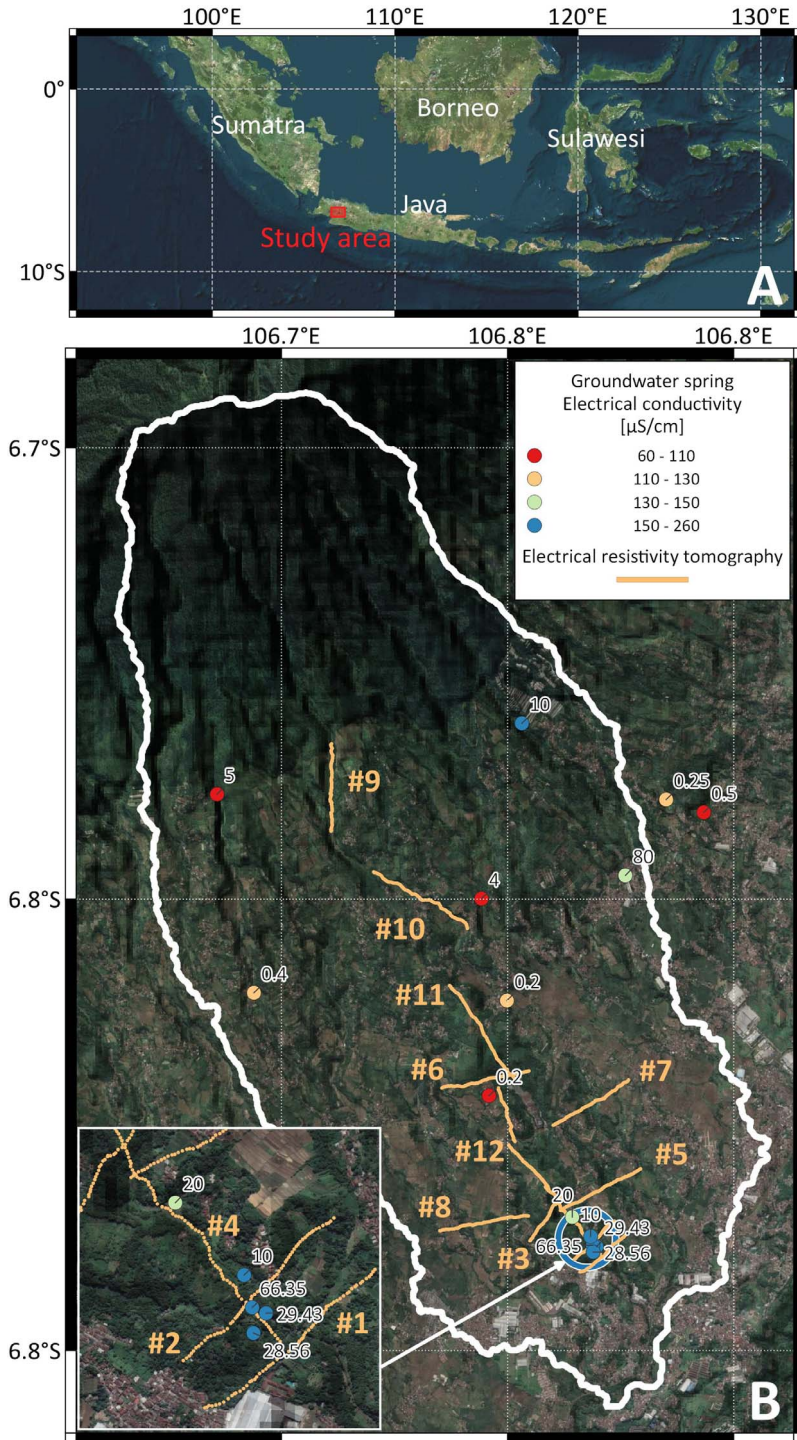
## 1. Introduction

Freshwater is an indispensable component of ecosystems and human activities' essential needs. The sustainability of human societies depends on the constant supply of clean fresh water. Currently, only 10% of the Earth's renewable freshwater is used by humans, while the other 90% sustains natural ecosystems and biodiversity [de Marsily, 2020]. To improve the human management of renewable water and sustain societal development, we need to increase our understanding of water interaction with the other components of the water cycle. Nevertheless, the consideration of water fluxes in all natural compartments is hampered by the availability of relevant data and/or conceptual models. This problem is all the more important in countries where observation networks of the water cycle have only recently been developed, and often sparsely. In such conditions, the first challenge is to accurately monitor and quantify the water fluxes from the atmosphere to the subsurface. While the dichotomy of the water cycle between the water reservoirs may be difficult to grasp in complex geological settings, our study presents a methodology combining multidisciplinary field measurements and probabilistic simulations to (i) quantify the water cycle within the different compartments of a previously unknown watershed, and (ii) define field monitoring needs to increase the accuracy of the water cycle representation.

Our study focuses on a watershed located along the flank of Mount Salak, an andesitic volcano at the western part of Java island, Indonesia (Figure 1A).

In this tropical region, the humid tropical climate generates significant rainfall (from 2 to 5 m per year) year-round accompanied by high evapotranspiration (from 1 to 1.4 m per year). Java Island climate is highly impacted by Asian and Australian monsoon leading to intense rainfall during the wet season [from November to March; Aldrian and Djamil, 2008]. Conversely, the dry season is characterized by low rainfall inducing notably a decrease in surface water and spring discharge. Population growth and urban sprawl lead to a constant increase in anthropogenic pressure on water resources. Upstream, surface water is largely used for the irrigation of rice fields [Khasanah *et al.*, 2021]. Downstream, industrial and urban wastewater discharges to the rivers have led to a degradation of surface water quality [Buwono *et al.*, 2021, Suriadikusumah *et al.*, 2021]. This expansion of human activities has led to a significant increase in anthropogenic pressure on groundwater resources throughout the watersheds [Khasanah *et al.*, 2021]. This situation could lead to water supply shortages during the dry season and degradation of groundwater quality.

In this context, it becomes crucial to understand the distribution of renewable water flows through the different compartments of the watershed. However, such quantification is hampered by the heterogeneity of the subsurface, which is composed of volcanic rocks and/or detrital formations that host spatially variable heterogeneous aquifers [Lachasagne *et al.*, 2014, Selles *et al.*, 2015]. The summit of the volcanoes, characterized by steep slopes above 20° and up to 40°, is usually mainly com-



**Figure 1.** (A) Location map of the study area in Indonesia. (B) Map of the investigated surface watershed (delineated in white) with location of surveyed groundwater springs (solid circles), and of the twelve numbered ERT profiles. The springs are colored as a function of the water electrical conductivity and labeled with spring discharge (in L/s).

posed of andesitic lava flows and pyroclastic projections. At the foot of these steep slopes, lava flows transition to alternations of detrital and distal ash fallout layers that form peripheral detrital aprons. Lower down, the aprons become alluvial floodplains, with alternating laharic formations and fine-grained soils, which can prevent the upwelling of underlying groundwater, creating artesian aquifers [Toulier *et al.*, 2022]. This layering of geological formations related to the distribution of volcanic formations spatially disconnects the surface watersheds from underground ones [Charlier *et al.*, 2011, Dumont *et al.*, 2021, Vittecoq *et al.*, 2014]. At the same time, geological discontinuities along slopes may allow the release of groundwater at spatially-restricted emergence points, which may substantially contribute to surface water flows [Vittecoq *et al.*, 2019]. These andesitic hydrosystems are complex due to their lateral and vertical heterogeneity and are poorly or not instrumented.

Hydrogeological models are useful tools to better assess water fluxes between the compartments of the water cycle. The existing modelling strategies cover a large spectrum organized along a two-dimensional continuum according to spatial resolutions and process complexities [Hrachowitz and Clark, 2017], from lumped, conceptual models to process-based models describing the system from a microscale perspective. Both approaches were applied in andesitic settings. Lumped or semi-distributed holistic reservoir models were developed to accurately capture water exchanges between the atmosphere, the soil, runoff and the subsurface [Cerbelaud *et al.*, 2022, Charlier *et al.*, 2008, Gómez-Delgado *et al.*, 2011, Mottes *et al.*, 2015]. Generally calibrated on river flows, they simplify the complexity of the aquifers that is induced by the andesitic volcano structure. Process-based models solving Richards' equation and Darcy's law have also been developed to represent physical processes [Chang *et al.*, 2020, He *et al.*, 2008, Lo *et al.*, 2021, Satapona *et al.*, 2018]. In this case, the finite element mesh is constructed by incorporating geological and hydrogeological features to simulate the water table with a calibration on monitored boreholes. Nevertheless, the surface/underground exchanges are significantly simplified limiting the representativeness of the aquifers in the watershed water cycle. Both approaches rely heavily on data, either to calibrate model parameters or to prescribe the geometry and

hydrodynamic properties of geological and hydrogeological features. To calibrate and/or to validate the models, continuous river discharge or piezometric data are required over long record periods to represent annual and interannual variability. This traditional calibration framework does not allow for the integration of non-continuous data or expert knowledge that could improve the performance of the simulations [Gharari *et al.*, 2014].

To model andesitic hydrosystems, we argue that representing the surface/underground exchanges is mandatory while at the same time keeping the representativeness of the aquifers consistent with the heterogeneity of the environment. In addition, the management of weakly instrumented watersheds would benefit from the development of an iterative approach between field measurements and/or monitoring and modelling approach to incrementally improve water cycle quantification. Such an approach may provide early-stage estimations of water fluxes and highlight which data are missing for more accurate estimation. The development of such water cycle simulation is strongly needed in order to sustainably manage water resources and to quantify the impact of climate on water availability.

To achieve this, our study develops a methodology to structure the progressive characterization of a minimally instrumented hydrosystem. The objective is to obtain a first simulation usable for water use management while guiding future surveys and monitoring. This approach was designed for a previously hydrogeologically unknown mountainous watershed in West Java (Figure 1A). Located on the volcano chain crossing the island, this mountainous area is strategic for a wide variety of water use, such as for population supply, irrigation, and industry. However, these hydrosystems are considered hydrogeologically unknown because their geological structures, extensions and dynamics have not yet been determined, which is necessary to improve their management. The only hydrological data available in the watershed is the discharge of the artesian springs used to supply the local population and/or industrial uses. Consequently, a better understanding of the hydrosystems and their functioning is necessary to define sustainable water management. Our study is divided into three steps: (i) hydrosystem characterization through a multidisciplinary study to set up a conceptual model; (ii) the transformation of the

conceptual model into a numerical one; and (iii) the calibration of this latter using field data.

## 2. Material and method

### 2.1. Methodological approach

To set up the conceptual model, we need to define the structure, characterize the compartments of the water cycle and estimate their behaviors to construct a meaningful model. First, the volcanic structure and aquifer(s)/impervious bodie(s) geometries have to be determined. For decades, geophysical methods demonstrated their strength to image volcanic structures and identify aquifer geometries [Descloitres *et al.*, 1997, Dumont *et al.*, 2019, Gresse *et al.*, 2021, Vittecoq *et al.*, 2019]. We thus imaged the hydrogeological structure of the watershed using 2D electrical resistivity tomography (ERT) combined with a geological review of surface outcrops and of artesian borehole lithological data. Second, the different aquifers need to be identified and characterized between shallow, perched, and deep reservoirs [Join *et al.*, 2016]. Geochemistry has proven its efficiency in identifying the recharge areas and the main characteristics of the aquifer [Bénard *et al.*, 2019, Join *et al.*, 2016, Toulhier *et al.*, 2019]. To do so, we used geochemical analyses and discharge measurements to identify both perched and deep artesian aquifers. This further allowed defining the extent of each identified aquifer type. In addition, the elevation of aquifer recharge was assessed using water stable isotope of rainfall within the watershed and deep artesian aquifers. Lastly, hydrological monitoring of river or spring discharges are an important data to compute water balance between the different water reservoirs of the watershed [Charlier *et al.*, 2011, Vittecoq *et al.*, 2019]. To overcome the lack of such dataset, a rough surface water balance was obtained from point measurements of river flow, which are used to estimate surface water/groundwater flows.

As the model relies on limited information while integrating the whole water cycle, we used a lumped model. The production function adopts the equations of Thornthwaite [1948], while the model storage discharge equation is derived from Perrin *et al.* [2003]. We applied a Bayesian approach, the multiple-try differential evolution adaptive Metropolis (MT-DREAM<sub>ZS</sub>) to account for the uncertainty in

model parameters calibration [Laloy *et al.*, 2015]. The starting probability distribution of the parameters is set as uniform within a threshold estimated from field measurements. Multiple runs of Markov chains then allow transforming the a priori distribution functions into a posteriori probability distribution functions [de Marsily *et al.*, 1999]. This probabilistic exploration of the model parameters space, framed by field interpretations, allows (i) to improve the representability of the water cycle simulation, and (ii) to guide future field characterization by estimating the influence of each parameter on the simulation uncertainty. Thus future field surveys or instrumentation can be optimized using the simulation guidelines.

### 2.2. Study area

The study site is located on the southern flanks of the Salak volcano, which culminates at 2900 m a.s.l. (Figure 1). This volcano is part of the mountain belt that stretches along the center of the island of Java from west to east. The study area straddles the surface drainage limit between the northern and southern sides of the island. Water resources of the watershed supply local populations as well as agricultural and industrial activities. While the main agricultural use consists of irrigating rice fields by river diversion, other activities are supplied by spring captation or pumping into wells. It is therefore essential to characterize renewable water volumes between the different components of this hydrosystem to maintain a base flow sufficient to sustain local biodiversity and local communities.

### 2.3. Field measurements

#### 2.3.1. Geomorphology

The geological description of the Salak volcano is still limited. Therefore, we applied a geomorphological analysis to define its main geological units. In particular, slopes and land roughness allow to distinguish areas underlain by lava flows from those underlain by detrital deposits, as described by Selles *et al.* [2015]. Surface topography characterization was performed on a digital elevation model (DEM) calculated from the Advanced Land Observing Satellite dataset [ALOS—Tadono *et al.*, 2014]. With a resolution of 5 m, the ALOS DEM provides an accurate

depiction of the slopes morphology of Salak volcano. The DEM geodesic reference is EPSG 32748—WGS 84/UTM zone 48S. The DEM was used to extract slope maps and delineate surface watersheds using the QGIS software (v3.10.11—<https://www.qgis.org/>). A pit-filling algorithm (`r.fill.dir` from the GRASS v7.8.4 package) was used to remove local depressions and other artifacts. The filled DEM was then used to define the watersheds and the hydrographic network using the “`r.watershed`” function from the GRASS v7.8.4 package. The slope map was calculated using the slope function of the GDAL v3.1.4 package.

### 2.3.2. Geophysics

To confirm the geomorphological classification of land surfaces into geological units and to provide some insights on the underlying geology, geophysical and geological campaigns were conducted during two field surveys. The first, conducted in December 2020, focused initially on the immediate surroundings of the artesian springs (Figure 1), before proceeding upstream in order to explore the artesian structure extension. Artesian spring water is collected into shallow boreholes, which logging data provide deep geological information. This initial survey was meant to understand the geological parameters controlling the location of the artesian springs, and to collect the geological information at depth needed to assist geophysical imaging interpretations. A second survey was conducted in June 2021 in order to complete geophysical imaging, such as to obtain a nearly continuous longitudinal geophysical section from the artesian springs area up to the break-in-slope marking the transition from the core of the volcano to the surrounding detrital accumulation zone.

During these two surveys, a total of 12 ERT profiles were acquired using an IRIS Instruments Syscal Pro resistivity meter (<http://www.iris-instruments.com/syscal-pro.html>). All 2D lines were acquired using 64 electrodes with a 10 m spacing and some roll-along specified in Table 1. Acquisition sequences were defined with Wenner–Schlumberger set up, which provides the best trade-off between shallow resolution and depth of investigation [Dahlin and Zhou, 2004]. Each measurement was stacked 3 to 6 times with imposed electrical potential of 800 mV. The measurement time window was 500 ms.

After field acquisition, ERT and roll-along data were merged and processed using Prosys III software (v1.05, <https://www.iris-instruments.com/fr/download.html#processing>). In order to delete gap-fillers measurements, as well as measurements with negative apparent resistivity, low intensity and/or low electrical potential data were filtered out. Finally, data with poor standard deviation (above 10%) and insufficient stacking were also removed.

Data were then inverted using 11 norm in Res2dinv software version 3.59.121. The result is presented after 3 iterations. Similar parameters were used to invert all ERT profiles. The width of model grid cells was set to half the electrode spacing. The thickness of the top layer corresponds to 120% of the cell width, and the layer thickness increases by a factor of 1%. The depth of the model is defined with a vertical extension of 5%. The damping factors used are: (i) 0.15 for the initial, (ii) 0.02 for the minimum, and (iii) 5 for the top layer. The flatness ratio is set at 1 to ensure the imaging of both vertical and horizontal structures. Finally, the depth of investigation is estimated using a sensitivity estimation of the inverted resistivity model [Marescot *et al.*, 2003, Oldenburg and Li, 1999]. This approach uses two inversions with different initial models and calculates the difference between the two inverted results. Here, we considered that if that difference exceeds 10%, the inverted resistivity model is not valid. Inversions yielding differences between 5 and 10% are regarded as poorly reliable but still informative.

2D-inverted resistivity profiles were ground-proofed near the artesian springs using available geological logs defined with the drilling cores. They allowed us to identify the geological formations corresponding to the imaged geoelectrical layers. The ERT profiles obtained during the second survey allow us extending these interpretations up the slope, combined with surface geological observations.

### 2.3.3. Hydrology and geochemistry

The study site is located in Java Island characterized by a tropical climate. The year is divided in two seasons: the wet season characterized by strong moonson rainfall (October–March); the dry season with less or no rainfall at all [April–September; Dumont *et al.*, 2022]. The annual rainfall varies from 2 to 5 m/year, while the daily average temperature ranges from 18 to 25 °C.



**Table 1.** Summary of ERT acquisition parameters, which provides for each of the 12 profiles the total length, number of electrodes and the roll-along. No roll-along was implemented along profiles #8, #11 and #12. However, profile #8 is composed of two profiles with an overlap of 310 m, whereas profiles #11 and #12 do not overlap. The two processing columns present the number of data obtained after (without gap-fillers) and before processing, with indication of the percentage of data retained for the inversion. The last column summarizes the absolute error of the inverted model for each profiles. All the profiles are located in Figure 1

ERT profile	Length (m)	Electrodes	Roll-along	After processing	Before processing	Absolute error (%)
#1	630	64	—	1100	966 (88%)	6.20
#2	630	64	—	1100	1023 (93%)	10.45
#3	630	64	—	1010	821 (81%)	8.99
#4	950	64	32	1944	1592 (82%)	6.68
#5	950	64	32	1920	1720 (90%)	8.25
#6	950	64	32	1944	1716 (88%)	9.72
#7	950	64	32	1944	1698 (87%)	15.09
#8	950	2 × 64	—	2200	1700 (77%)	14.36
#9	950	64	32	1920	1749 (91%)	16.75
#10	1190	64	32 + 24	2520	2434 (97%)	13.25
#11	630 × 2	2 × 64	—	2110	2015 (95%)	7.96 & 3.15
#12	630 × 2	2 × 64	—	2200	1944 (88%)	4.03 & 4.78

Alongside the geophysical surveys, springs were inventoried. Their discharges and physico-chemical properties were measured in the field and water samples were taken for further analyses, in particular laboratory measurements of major elements and isotopic ratios. Major elements analysis has been proceeded by UnPad ITB laboratory using an ion chromatography instrument. The water isotopic ratios have been estimated by the BATAN laboratory (National Nuclear Energy Agency of Indonesia) using a liquid-water stable isotope analyzer Los Gatos Research DLT-100. Springs were classified according to discharge, major element composition, and groundwater electrical conductivity [Bénard *et al.*, 2019, Charlier *et al.*, 2011, Join *et al.*, 1997, 2016]. Water stable isotopic ratios were used to define the average elevation of the recharge area using the local meteoric line, which we defined from analysis of weighted rainfall collected within the watershed as described in Toulhier *et al.* [2019]. For the watershed, rainfall was collected at four different elevations during the hydrological cycle 2006–2007, both the amount of rainfall and its mean monthly isotopic composition were measured. To characterize the artesian aquifer

recharge, three campaigns of water sampling for water stable isotope analysis were carried out from 2006 to 2007. The variations in water stable isotope ratios over the period allowed us to assess the seasonal variability of the recharge processes and to obtain an estimate of the average elevation of the aquifer recharge. Thus the spread of isotopic results informs on changes in the dynamics of the recharge from dry to wet seasons.

Once the hydrogeological structure of the subsurface was defined, it was necessary to understand the distribution of water flow and its dynamics. This step was hampered by the absence of long-term monitoring of river discharges within the surface watershed. Besides, artesian aquifers extension is generally large and disconnected from the surface drainage network. On the one hand, part of the subsurface flow could not be quantified with hydrological monitoring. On the other hand, the model must represent the functioning of the artesian springs and not the whole system in order to be calibrated on the available spring discharges. The river flow measurements undertaken during the 2019 dry season and the 2020 wet season were used as a constraint for sur-

face discharge evaluation in the model. During these surveys, the discharge was repeatedly measured at downstream locations whereas the upstream rivers were measured only during the wet season. The specific discharge rates (as a function of surface area) derived from these site-specific measurements were used as proxies of the specific runoff of all upstream sectors (central and proximal) in order to compute the total runoff. The difference in flow between the dry and wet seasons provides some preliminary insights on annual variations in the dynamics of the rivers. Nevertheless, this information is not comprehensive enough to be used as input in the models and was therefore only used to constrain the simulation results.

Finally, to constrain the water cycle in the system, we use the discharge of the artesian springs tapped for the drinking water supply of the population (Figure 2). These springs have been gradually instrumented since 2015. It is thus possible to have a partial measurement of the artesian aquifer outlet. Indeed, these measurements do not integrate all the artesian springs identified on the field, nor the underground flows that may not be drained by the springs. Nevertheless, we summed the discharge of all the inventoried springs issuing from the artesian aquifer to obtain one discharge time series of more than 5 years, from 2015 to mid-2021, that recorded the seasonal fluctuations of the system. The climate forcing during this period evolved from one year to another with notably an ENSO (El-Niño Southern Oscillation) event between 2014 and 2016. These data allow us to understand how the system reacts to variable forcings.

## 2.4. Water cycle simulation

### 2.4.1. Lumped model structure

Since the final structure of the hydrological model depends directly on the characterization of the hydrosystems, this section will focus on the description of the equations used. The model was chosen to run at the daily timestep in order to account for the strong rainfall events that occur in such a tropical context.

The production function of the model consists of a soil production storage ( $SW$ ) defined to estimate

the balance between rainfall, evaporation, and vegetation transpiration. The Thornthwaite equation is used to compute the effective rainfall ( $ER$ ) as:

$$ER = R - AET + \frac{\Delta SW}{\Delta t} \quad (1)$$

where  $R$  is rainfall,  $AET$  is actual evapotranspiration, and  $SW$  is soil water content ( $SW_0 \leq W \leq SW_{\max}$ ), where  $SW_{\max}$  represents the soil water holding capacity and  $SW_0$  is the lower limit of evaporable soil water. The  $AET$  is calculated as follows:

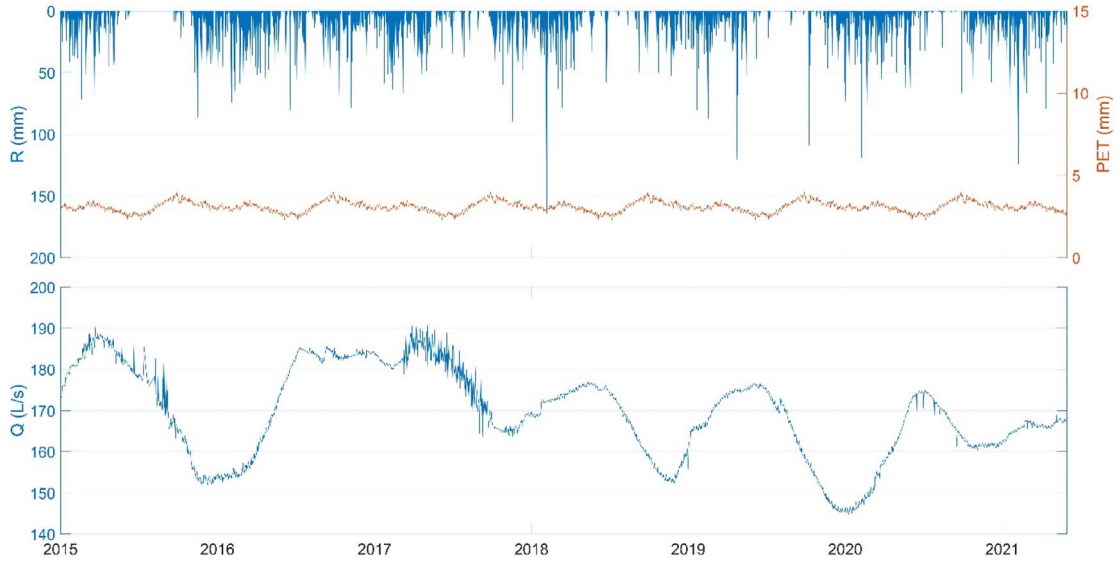
$$AET_n = \begin{cases} PET_n & \text{for } P_n + SW_{n-1} \geq PET_n \\ P_n + SW_{n-1} & \text{for } PET_n - P_n \geq SW_{n-1} \\ 0 & \text{else.} \end{cases} \quad (2)$$

Where  $n$  is the daily timestep and  $PET$  is the potential evapotranspiration estimated using the Penman–Monteith equations [Monteith, 1965], which explicitly accounts for climatic and vegetation characteristics:

$$PET = \frac{0.408 \times \Delta(R_n - G) + \frac{\rho_a * C_p}{r_a} * (e_s - e_a)}{\Delta + \gamma * \left(1 + \frac{r_s}{r_a}\right)} \quad (3)$$

with  $\Delta$  the slope of the saturation vapor pressure curve as a function of the air temperature ( $\text{kPa} \cdot \text{°C}^{-1}$ );  $R_n$  the net radiation ( $\text{MJ} \cdot \text{m}^{-2} \cdot \text{day}^{-1}$ );  $G$  the soil heat flux density ( $\text{MJ} \cdot \text{m}^{-2} \cdot \text{day}^{-1}$ );  $\rho_a$  the air density;  $r_a$  the aerodynamic resistance ( $\text{s} \cdot \text{m}^{-1}$ );  $r_s$  the stomatal resistance ( $\text{s} \cdot \text{m}^{-1}$ );  $e_s$  the saturation vapor pressure at the mean air temperature ( $\text{kPa}$ );  $e_a$  the actual vapor pressure of the air ( $\text{kPa}$ );  $\gamma$  the psychrometric constant equal to  $0.066 \text{ kPa} \cdot \text{°C}^{-1}$ ;  $C_p$  the specific heat of the air ( $\text{MJ} \cdot \text{kg}^{-1} \cdot \text{°C}^{-1}$ ).

Weather data used for the water cycle simulation come from the Stamet Citeko station of BMKG network (Badan Meteorologi, Klimatologi, dan Geofisika—<http://www.bmkg.go.id>) located at 14 km north of the watershed. While the station is not within the watershed, the rainfall analysis of Dumont *et al.* [2022] demonstrated the representativity and accuracy of this weather station monitored from 2000 to 2021 for the studied area (Figure 2). For the estimation of potential evapotranspiration, the Stamet Citeko station is the only station providing long-term climate data (temperature, radiation, wind, sun exposure time per day and relative air humidity) without long gaps. However, due to punctual gaps in the climate data, they have been averaged over the time series to obtain the climatological potential evapotranspiration for each day of the year



**Figure 2.** Daily rainfall ( $R$ ), potential evapotranspiration ( $PET$ ) and artesian springs discharge ( $Q$ ) time series from 2015 to mid-2021. Note that the higher noise of the data in 2015 and 2017 is a consequence of the pressure device; the noise is not representative of the real discharge of the spring.

(Figure 2). This long-term climatic time series allows us to assign a warm-up period to the model from the beginning of 2000 to the end of 2014, the latter corresponding to the beginning of spring discharge measurements. This warm-up period ensures that the model during the studied period is not impacted by the choices of model initialization.

Below the  $SR$  reservoir, the  $ER$  feeds runoff ( $RO$ ) and infiltration. Flux allocation is defined by the Baseflow Index ( $BFI$ ) parameter, used to calculate the infiltration/effective rainfall ratio. Deeper in the ground, the model structure is not defined at this stage, as it depends on spatially varying hydrogeological characteristics. In the model, each reservoir is characterized by a recharge and by the infiltration or the discharge from a shallower reservoir. Its discharge is calculated as a function of the reservoir state using the GR4J model equations [Perrin *et al.*, 2003]. For a reservoir level  $R$  (mm), the outflow  $Q_R$  thus depends on its current filling level  $R$  and its maximum capacity  $R_{max}$  (mm):

$$Q_R = R \left\{ 1 - \left[ 1 + \left( \frac{R}{R_{max}} \right)^4 \right]^{-\frac{1}{4}} \right\}. \quad (4)$$

#### 2.4.2. Probabilistic parameter exploration

The objective of our methodology is to define the range of possible values for input model parameters, based on the interpretation of field data. This will allow a better linkage between the numerical simulations and field reality while improving the future characterization of the studied hydrosystems. For this reason, the parameterization of the artesian springs flow simulation is based on a probabilistic approach. To do so, we used the multiple-chains differential evolution adaptive Metropolis algorithm [MT-DREAM<sub>ZS</sub>—Laloy *et al.*, 2015] to infer the posterior probability density functions (PDF) of model parameters. At the start, an a priori probability distribution is defined for each parameter. At each iteration, the MT-DREAM<sub>ZS</sub> algorithm uses a latin hypercube sampling to sample parameter values following the statistical methodology presented in Laloy *et al.* [2015]. Once the parameter values are accepted, the water cycle simulation is run in order to estimate a log-likelihood function used to compare simulation results and measured discharges. The use of multiple Markov chains allows us improving parameter space exploration, limiting the impact of local maximum likelihood on the search for best-fit parameters. In this study, five Markov chains are used with 50,000

iterations per chain. At each iteration, a candidate parameter set is tested via the Metropolis ratio. If it is accepted, the parameter set is validated and integrated into the calibration process.

### 3. Results and discussion

This section is divided into three parts. First, the conceptual model of the hydrological system is defined using field measurements and observations. Second, this conceptual model is used as the basis to build and set up the numerical lumped model which is used for the probabilistic parameters exploration. Third, the optimal parameter set is used to simulate the spring discharge.

#### 3.1. From observations to a conceptual model

To understand the geological structure of the watershed, our methodology relies on a combination of geological, geomorphological and geophysical information (Figure 3). In addition, hydrogeological knowledge such as the location and properties of the springs are of prime importance to decipher the structure of the conceptual model.

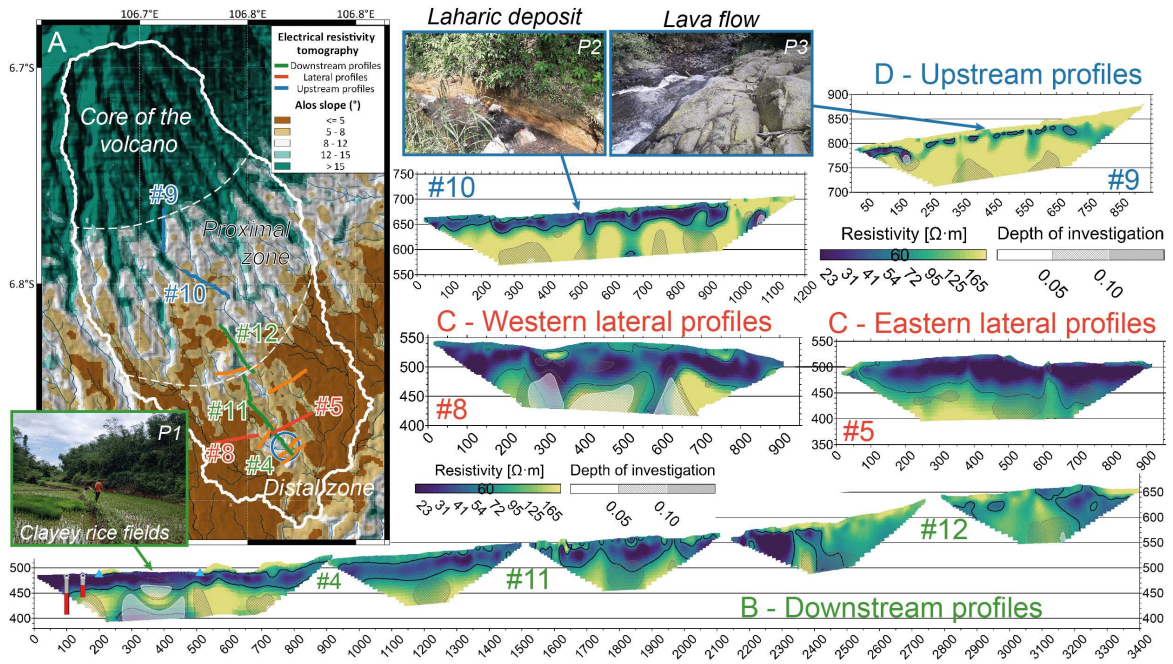
##### 3.1.1. Geomorphological and geophysical results

The slopes map allows dividing the watershed into three zones: the core of the volcano is characterized by steep slopes (average: 23.9°) and, on the opposite, the distal zone by lower slopes (average: 5.5°); the intermediate proximal zone shows variable slopes between 5° and 20° (Figures 3A and 4). In the latter, field geological observations showed that several morphological forms correspond to lava flow fronts, indicating that it is the meeting zone between lava flows tracking from the core of the volcano and the detrital deposits of the distal zone. Downstream of these massive flows, the valley is particularly eroded with slopes of up to 12°, which contrasts with the low erosion of the distal area. The outcrop of lava flow may have disconnected this area from the upper part of the watershed, limiting alluvium deposit or mark the upper limit of regressive erosion coming from downstream. This particular feature allows the underlying formations to outcrop.

The ERT profiles provide an imagery of the structures at depth (Figure 3B–D). The lowest profile

stretches from the artesian springs zone to the lower limit of the proximal zone. These profiles were geologically and hydrogeologically calibrated from the geological data of springs and boreholes located near the springs area (Figure 3B). Thus, these ERT profiles reveal a long, from 10 up to 50 m thick, superficial conductive layer ( $<60 \Omega\cdot\text{m}$ ), which extends to about 570 m a.s.l. and thickens over a distance of 2400 m. This conductive layer corresponds to a clayey laharic formation identified in the wells. There it constitutes an aquiclude that seals the underlying resistive ( $>60 \Omega\cdot\text{m}$ ) artesian aquifer. The artesian aquifer corresponds in the wells to mildly reworked volcanic breccias, pyroclastic rocks, and sandstones. This main deep reservoir feeds several high discharge ( $>10 \text{ L/s}$ ) springs observed in the springs area. Groundwater flows out from the aquifer through the locally thinner (notably in incised valleys) and/or locally more permeable laharic formations (Figure 4). Additional ERT profiles performed immediately downstream the spring area show that the aquifer is bounded there to the south-east by clayey conductive formations; this is an important feature as it surely prevents any groundwater to flow downstream the artesian springs. The western and eastern lateral ERT profiles confirm that both the aquifer and its confining layer extend laterally West and East. The deep resistive aquifer layer is less resistive in the western part, while it deepens towards the eastern limit. Unfortunately, these lateral limits of the aquifer were not reached by the current #5 and 8 ERT profiles (Figure 3C). However, given the absence of other known artesian springs laterally (see Figure 7 below), the aquifer is not expected to extend much farther than the chosen surface watershed boundaries.

At the transition between the distal and the proximal zones, a few tens of meters in elevation above the upper limit of the aquifer confining layer, the ERT profiles reveal a thick resistive layer (80–150  $\Omega\cdot\text{m}$ ) that reaches the surface on a steep slope (8–15°). Both the morphology and outcrops confirm a lava flow front issuing from the North. From this elevation, a conductive layer consistent with laharic outcrops is present in the areas with the shallow slopes ( $<8^\circ$ ). At the limit with the core of the volcano, the conductive layer deepens and becomes thinner (as seen in #9). This conductive layer is thus comprised between two resistive bodies (Figure 3D). The proximal zone appears as a transition area between a



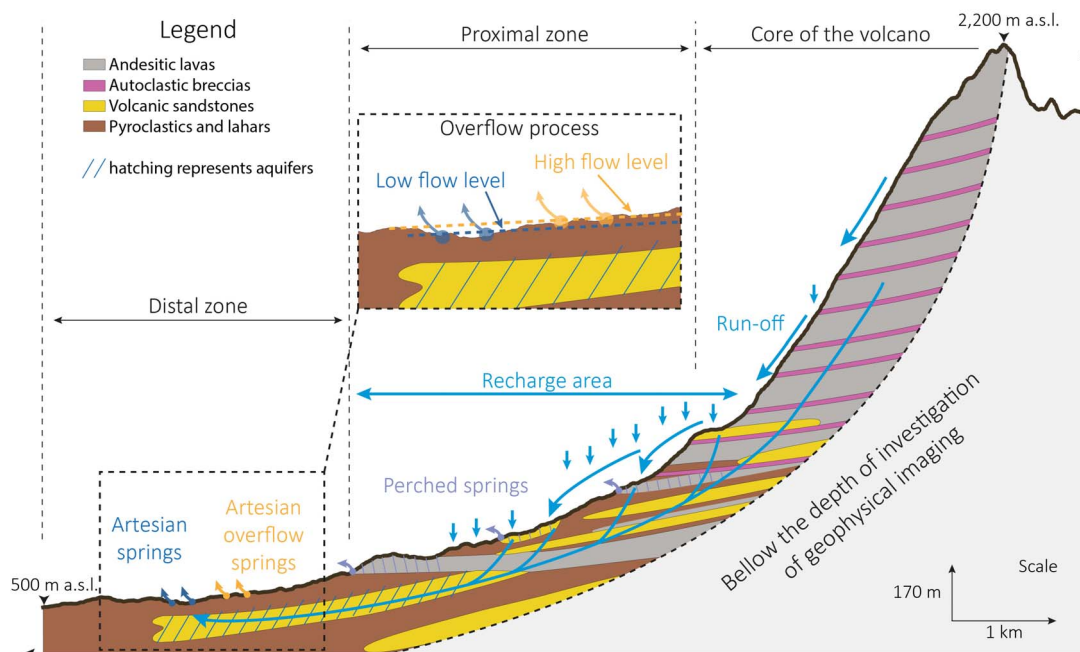
**Figure 3.** (A) Slopes map of the watershed extracted from the ALOS DEM with the location of the ERT profiles, (B) the downstream profiles starting from the artesian springs area (#4, #11 and #12—green), (C) lateral profiles oriented west–east (#5 and #8—red), and (D) upstream profiles (#9 and #10—blue). The distance along the profiles is presented in m (X axis), as well as the elevation (Y axis). Profiles not presented here are in orange in the map (A). In the slopes map, three geomorphological zones are delineated as: core of the volcano (slope > 15°), proximal (slope from 8° to 15°) and distal (slope < 8°) zones. The three field pictures illustrate the different types of vegetation cover and geological outcrops encountered in the catchment, with clayey rice fields (P1), laharc deposits of the proximal zone (P2), and lava formations of the core of the volcano (P3). In the lowest profile, two artesian wells are located with laharc formations in grey and pyroclastic ones in red. The two nearest springs are located with blue triangles.

reworked-pyroclastics dominant area (downstream) and a lava-flow dominant area (upstream). There, groundwater may flow from lava-flow to sandstone aquifers. There might also be surface–groundwater interactions, with both springs issuing from aquifers and, locally, streams infiltrating to aquifers. In the core of the volcano, only lava flows and breccias deposits crop out, which is consistent with alternating thinner conductive (lahars) and thicker resistive (lava flows) layers (Figure 3D).

### 3.1.2. Geochemical and isotopic results

Location and characterization of springs (discharge, geochemical and isotopic analyses) combined with hydrological measurements on

streams complete this geological/geomorphological/geophysical analysis. In the distal zone, there are no springs outside the spring zone (blue points in Figure 1). There, artesian springs are characterized by high discharges (between 10 and 40 L/s) and the highest measured electrical conductivities (130 to 150  $\mu\text{S}/\text{cm}$ —Figure 5A). During the wet season, additional artesian springs were identified a little further upstream in the field. This highlights that the artesian aquifer discharge varies according to the hydraulic head in the aquifer which is controlled by seasonal recharge. When this latter increases, temporary springs appear at a higher elevation. Several of these springs only flow during the wet season, accommodating fluctuations in discharge of the



**Figure 4.** 2D conceptual hydrogeological model of the volcano slope. The water flows are represented by blue arrows. A zoom of the artesian springs area provides information on the artesian aquifer water level and the overflow process.

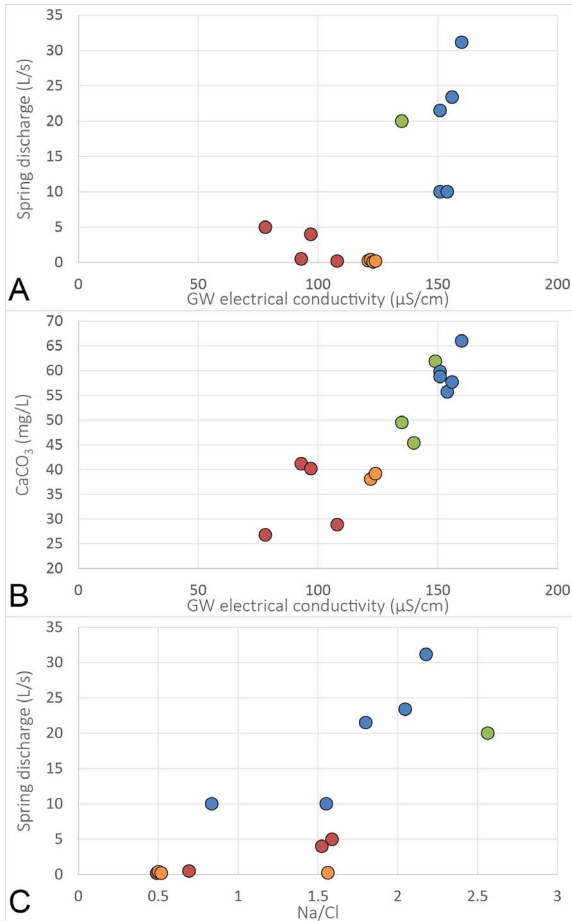
aquifer throughout the year (Figure 4). Upstream, in the proximal zone, all springs have a lower discharge rate ( $\leq 5$  L/s) and lower electrical conductivities ( $< 120$   $\mu\text{S}/\text{cm}$ —Figure 5A) which suggests the existence of only local perched aquifers of small spatial extent (Figure 4).

Globally, the groundwater electrical conductivity is well correlated to the discharge of the springs as well as to the carbonate mass concentration, which may be related to the contact time between groundwater and the aquifer rocks (Figure 5A and B). The sodium/chloride (Na/Cl) ratio, defined by Join *et al.* [1997], tends to be well correlated with flow discharge (Figure 5C). Higher ratios (longer water/rock interactions) are usually reached in the deep artesian aquifer while the shallow aquifers exhibit lower Na/Cl ratios (Na/Cl  $< 1.6$ ). The Na/Cl ratios support previous insights indicating that the proximal springs are fed by small extension aquifers while the artesian springs are supported by an extensive deep one (Figure 4).

Then, oxygen and hydrogen stable isotope ratios allowed to estimate the elevation of the recharge area of the artesian aquifer (Figure 6). The mea-

surement of monthly weighted rainfall isotopic ratios allow to estimate the water local meteoric line of Salak volcano (SMWL;  $\delta^2\text{H}\text{‰} = 7.83 \times \delta^{18}\text{O} + 12.35$ ). This meteoric line is consistent with Toulhier *et al.* [2019] Bromo–Tenger meteoric line (BMWL;  $\delta^2\text{H}\text{‰} = 7.71 \times \delta^{18}\text{O} + 12.37$ ). As for Toulhier *et al.* [2019], the SMWL appears above the global and local meteoric line, GMWL and LMWL respectively, demonstrating the local topographic effect on rainfall composition. The oxygen and deuterium ratios measured in the artesian aquifer have a low range of variation throughout the sampling periods compared to the variation observed in other springs and rivers sampled in the same time in the watershed (Figure 6A). This stability confirms the important inertia of the artesian aquifer.

In addition, the monitoring of stable isotopes of rainwater sampled at four different elevations (475, 628, 760 and 923 m) during the 2006–2007 hydrological year was used to calibrate the local elevation meteoric line:  $\delta^{18}\text{O} = -0.0026 \times \text{elevation} - 4.9672$ . We can thus estimate that the aquifer recharge occurs within the elevation range between 710 and 920 m a.s.l. according to Figure 6B, indicating a median



**Figure 5.** (A) Spring discharge as a function of the groundwater (GW) electrical conductivity; (B) carbonate mass concentration as a function of the groundwater (GW) electrical conductivity; and (C) spring discharge as a function of sodium/chloride ratio (concentrations are transformed in meq/L to estimate this ratio). The dot colors correspond to groundwater electrical conductivity clusters defined in Figure 1. Blue and green dots correspond to artesian springs.

recharge area of about 800 m a.s.l. Considering that the artesian aquifer recharge area can hardly extend below about 600 m a.s.l. (see section above on the geophysical interpretations; 200 m below the 800 m a.s.l. median recharge area), we can conclude that rainfall precipitated up to about 1000 m a.s.l. may recharge the artesian aquifer (200 m above the 800 m a.s.l. median recharge area), provided effective rain-

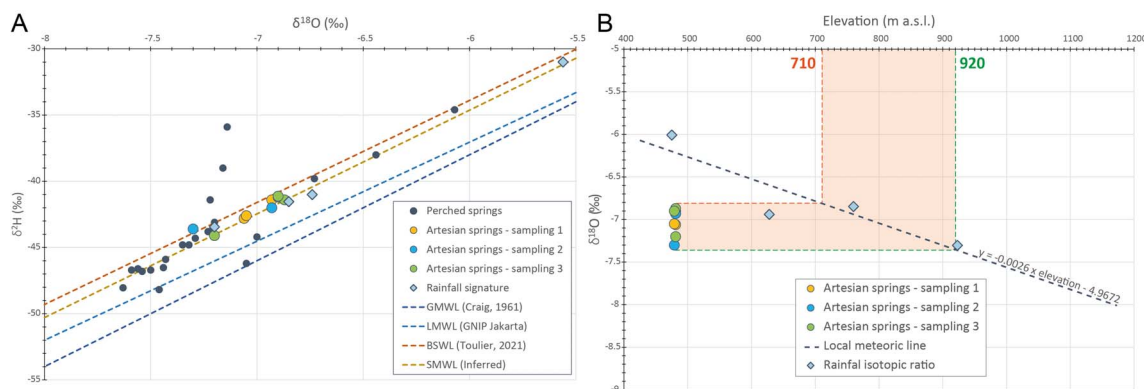
fall does not vary too much within this elevation range [Toulier *et al.*, 2019]. However, as recharge may also result from runoff after rainfall events, the upper limit of the recharge area is probably lower than 1000 m a.s.l. (Figure 4) as runoff may come from higher elevations than the recharge upper limit. More detailed research in that area should be performed to precise this upper limit.

### 3.1.3. Hydrological results

The final step in the characterization of the hydrosystem is the surface water dynamics. Concerning surface water, river discharge measurements showed limited flows in the upstream sectors with specific flows of 17 to 20 L/s/km<sup>2</sup> (river discharges between 120 and 200 L/s for a subcatchment surface of 6.5 to 11 km<sup>2</sup>). Downstream, the river flow increases up to 4140 L/s for a contributive surface of 40 km<sup>2</sup> (specific discharge of 104 L/s/km<sup>2</sup>). The overall watershed discharge evolves during the year with a decrease of the discharge by a factor of 8 during the dry season (a discharge of 510 L/s and a specific one of 12.75 L/s/km<sup>2</sup>). In fact, during the dry season, rainfall decreases drastically resulting in the drying up of perched springs and upstream rivers. The map of specific discharges is presented in Supplementary Material A.

The analysis of the artesian springs discharge time series allows a better understanding of the seasonal variations of hydraulic heads in the aquifer. These variations are buffered in the deep aquifer, with low water periods between October and March depending on the year. The annual variations range between 150 and 190 L/s. The suspected overflow effect described previously from the temporary discharge of the springs located above the springs area zone appears to be supported by a flow threshold around 190 L/s.

In this section, the combination of geological observations, geophysical imagery, slope analysis, hydrological and hydrogeological observations as well as geochemical and isotopic analyses allowed to characterize the hydrosystem's structure and limits, and provided a first robust hydrological conceptual model (Figure 4). The numerical model will then be based on this conceptual model. Nevertheless, the numerical model will also enable challenge and, if possible, confirm the coherence of this conceptual model.



**Figure 6.** (A) Oxygen and deuterium isotopic ratios of weighted rainfall (in light blue), artesian springs (in yellow, blue and green according to the sampling periods) and other springs and rivers of the watershed (in black). Meteoric lines are plotted with dashed lines: dark blue for the global meteoric water line [Craig, 1961]; light blue for the local meteoric line (GNIP station of Jakarta); orange for Bromo line [Toulier *et al.*, 2019]; and the Salak line (inferred in this study). (B) Analysis of water stable isotopes to estimate the mean elevation of the recharge of the artesian aquifer.

### 3.1.4. Water cycle characterization summary

In order to assist in understanding the transition from the conceptual to the numerical model, the characteristics of the hydrosystem are summarized below. In the studied area, a thick clayey layer extends over the distal area. It thins out at the center of the distal zone of the watershed, at places where the slope increases locally. At depth, the underlying aquifer is confined below this impervious layer over most of the area. Locally, thanks to the thinning of this clayey layer, it outflows through several artesian springs (Figure 7). Outside the artesian spring area (blue circle in Figure 7), no significant springs were identified in this distal zone. The layering imaged at low elevation in the distal zone appears to change in the proximal zone farther upslope. One of the reasons is the presence of lava flow fronts in the steep areas of the proximal zone, which could act as local, perched aquifers. These perched aquifers feed springs with different geochemical signatures (Figure 5), characterized by a lower mineralization, and shorter residence time than the deep aquifer, typical of local, perched aquifers [Join *et al.*, 2016]. Detritic and laharc accumulations probably produce the conductive clayey layers imaged near the surface. Farther upstream, slopes are directly underlain by high permeability pyroclastic flows with no interstratified low-permeability detrital formations.

There, the increase in slope, and progressive disappearance of clayey deposits blanketing the land surface, agree with the average recharge elevation estimated from stable isotopes (Figure 7). However, this result does not prevent the existence of a recharge from the core volcano slopes and lower proximal zone.

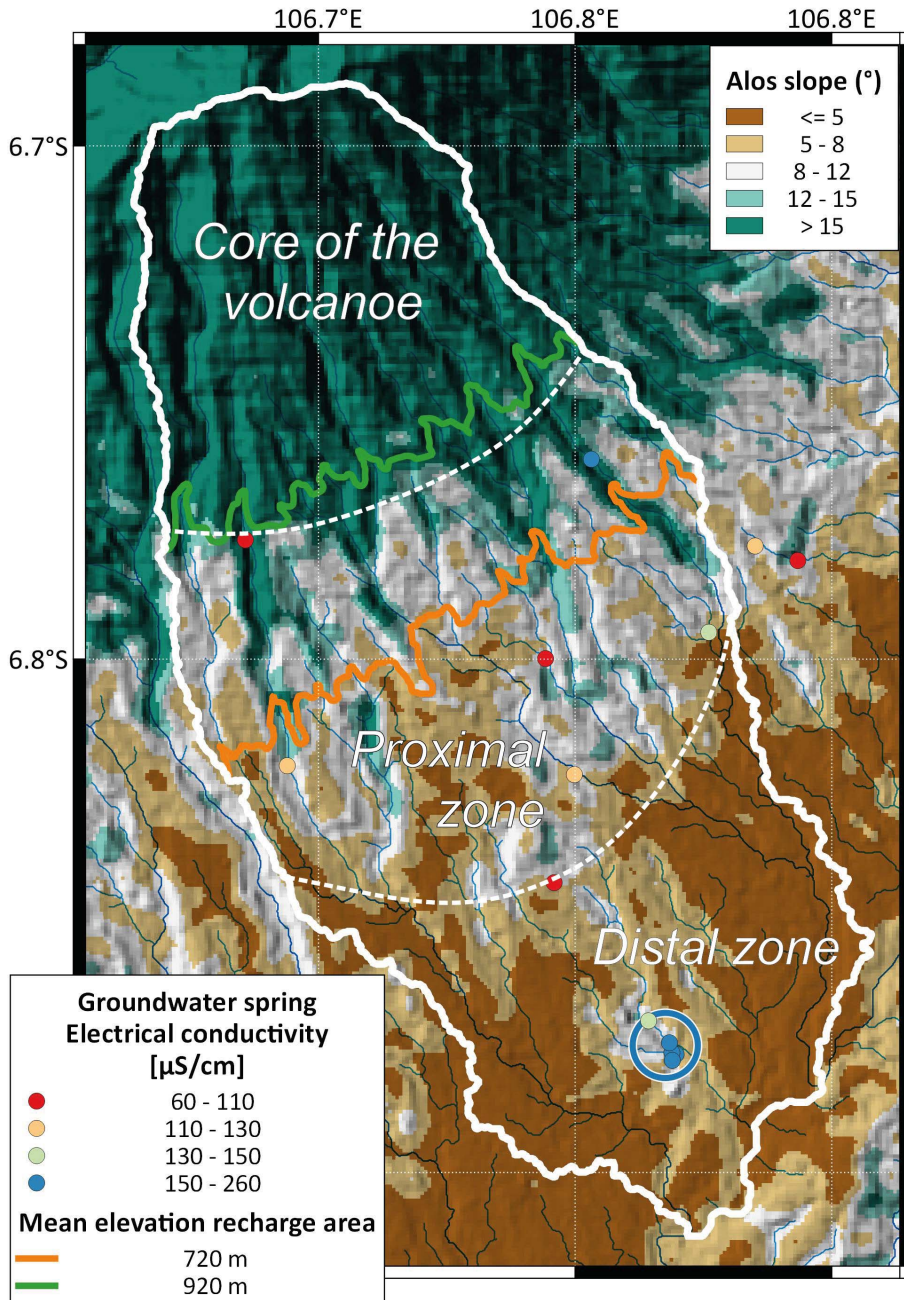
## 3.2. Probabilistic lumped model

### 3.2.1. Lumped model structure

The structure of the hydrological model is then set up based on hydrosystem architecture described in Figure 8. The lumped model relies on the equations previously presented in Section 2.4.

At the top of the model, seasonal fluctuations in rainfall ( $R$ ) and actual evapotranspiration ( $AET$ ) forcings control the water level in the soil water ( $SW$ ) reservoir (Equations (1)–(3)—Figure 8). When the  $SW$  reservoir is full ( $>SW_{max}$ ), excess water is transformed into effective rainfall ( $ER$ ). The  $ER$  is then divided between runoff ( $RO$ ), which supplies stream water, and infiltration to the shallow perched aquifer reservoir ( $P$ ). The parting of  $ER$  fluxes is defined by the  $BFI$  parameter. The discharge of  $P$  ( $Q_P$ ) is calculated using Equation (4) and divided into perched springs discharge ( $PSD$ ) and the recharge of the artesian aquifer ( $A$ ), according to the perched flow index

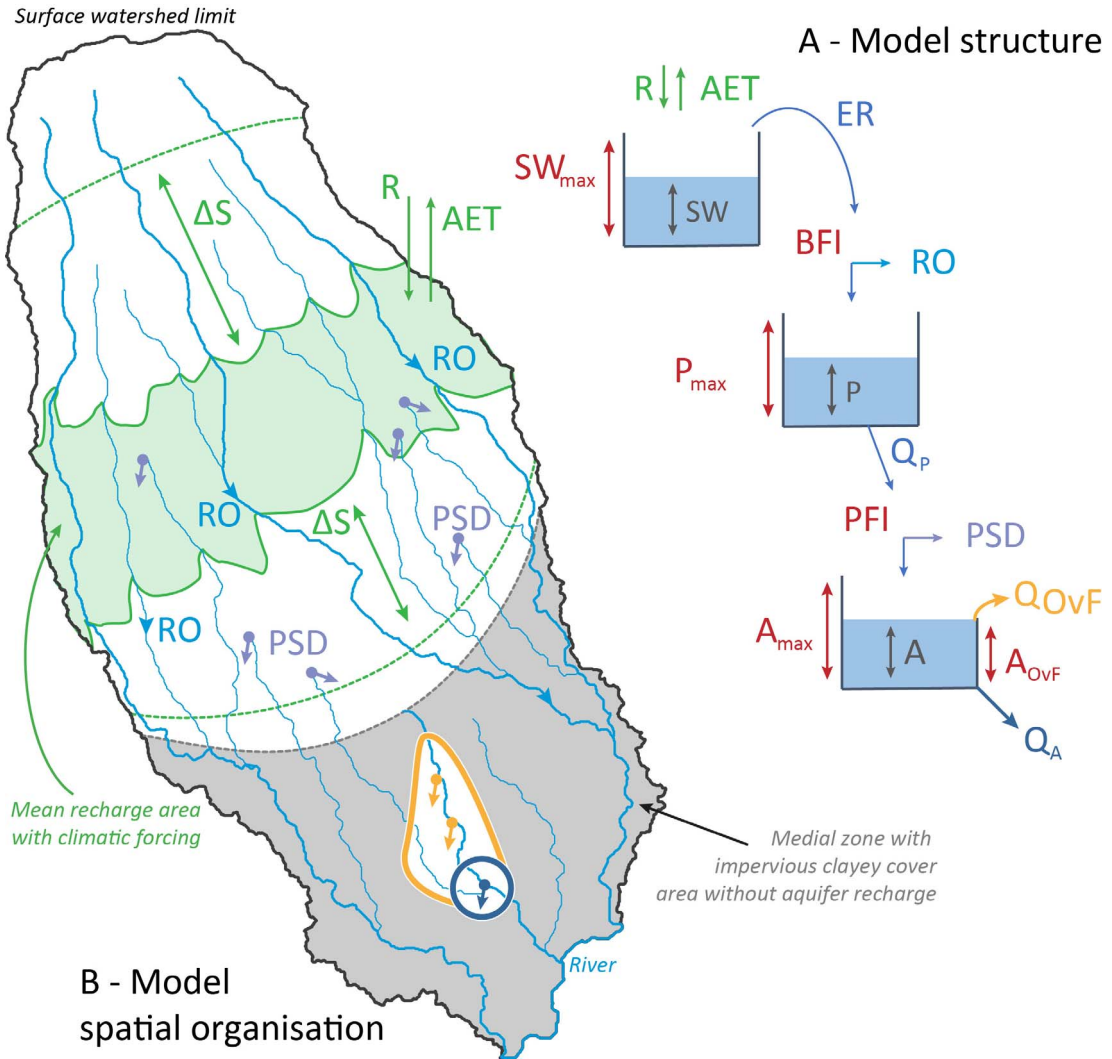




**Figure 7.** Slopes of the watershed extracted from ALOS DEM with the delimitation of the core of the volcano, proximal and medial zones. The artesian springs located in the artesian spring area (blue circle) are characterized by high electrical conductivity (130–250  $\mu\text{S}/\text{cm}$ ) and perched springs by a lower electrical conductivity (60–120  $\mu\text{S}/\text{cm}$ ). The mean recharge area is delimited by the lower (710 m, orange line) and higher (920 m, green line) limits estimated from water stable isotopes.

ratio (*PFI*). This latter corresponds to the ratio between water feeding perched springs discharge, and

water reaching the artesian aquifer. The discharge of reservoir *A* ( $Q_A$ ) represents the artesian spring's out-



**Figure 8.** Structure (A) and spatial organization (B) of the lumped model. In the structure, the parameters are represented in red, climatic forcing in green, model fluxes in light blue and model states in dark blue. In the map, the surface of the recharge area (S) of the model is represented in green. This surface may evolve according to the results of the isotopic ratios ( $\pm\Delta S$ ). The artesian (blue circle) and the overflow springs areas (yellow extension) are delineated according to the eroded area in the slopes map (Figure 7).

flow, which is simulated and compared to spring discharge measurements during the calibration process (blue circle—Figure 8B). However, considering that some springs only discharge during the wet season, we considered instead that, above a specific water level in reservoir A (threshold  $A_{OvF}$ ), the discharge of the springs no longer increases. This parameter consists of a filling ratio of reservoir A. This can be explained by the drainage of groundwater towards

non-perennial springs (not instrumented) located at slightly higher elevations (yellow extension—Figure 8B). To represent that shortcut in the model, an overflow function (*OvF*) was implemented into the artesian reservoir A. The excess water is thus transformed into a non-perennial springs discharge ( $Q_{OvF}$ ).

Consequently, the numerical model resulting from the proposed conceptual model is thus com-

posed of three reservoirs:  $SW$ ,  $P$  and  $A$  (Figure 8A). Each of them possesses a maximum filling level that constitutes 3 calibrated parameters:  $SW_{\max}$ ,  $P_{\max}$ ,  $A_{\max}$ . The other four parameters of the model are indexes  $BFI$  and  $PFI$ , overflow threshold level  $A_{OvF}$ , and the extension of the recharge area  $S$  that supplies the artesian springs (Figure 8A). As the discharge series we want to simulate correspond to only a part of the water flows in the watershed, the model extension should include only a part of the watershed. We delineate this feeding catchment  $S$  as the area enclosed within the elevation range of the recharge area, inferred from the analysis of the isotopic ratios of the artesian springs (Figure 6B). This surface can evolve towards a higher or lower altitude according to the hydrological cycles and its extension can also change.

### 3.2.2. Probabilistic exploration of lumped model parameters

The calibration of the model is conducted using the probabilistic approach MT-DREAM<sub>ZS</sub>. For this purpose, uniform a priori probability distributions are defined for each parameter. Field data are used to define the range of parameter values to improve the convergence of the simulations, prevent convergence toward local solutions, and improve the representativeness of spring discharge simulations.

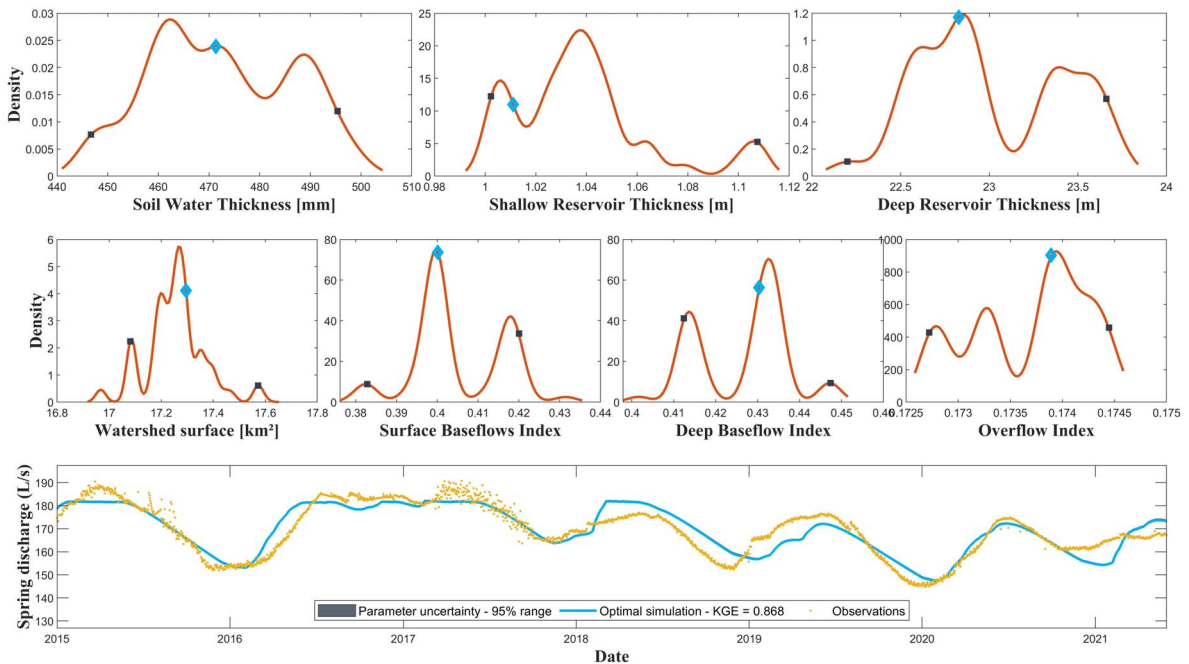
The first parameter  $SW_{\max}$  represents the thickness of the soil reservoir that serves as an interface between the atmosphere, the vegetation and the deep reservoirs. In the watershed, soils are thin on lava flows but thicken substantially on detrital formations. The a priori distribution of soil thickness  $SW$  was allowed to vary between 50 and 500 mm (of water). The ranges of values of reservoir filling parameters of the perched and deep reservoirs,  $P_{\max}$  and  $A_{\max}$ , were set based on geophysical imaging. In the central and proximal zones, resistant horizons, interpreted as perched aquifers, have thicknesses ranging from 1 to 20 m. At depth, the artesian aquifer displays thicknesses ranging from 10 m up to 60 m. The size of the feeding catchment  $S$  is about 7.2 km<sup>2</sup> from the area enclosed within the elevation of the recharge area, inferred from the analysis of the isotopic ratios. However, the recharge may take place over a larger area as demonstrated by Toulhier *et al.* [2019]. Therefore, the contributing catchment can

expand up to 30 km<sup>2</sup> according to the surface watershed (Figure 8B). An in-situ observable range of values for the remaining three parameters can not be constrained with field data. We therefore allowed these parameters to vary over a very broad range. Indexes  $BFI$  and  $PFI$  thus vary from 0 to 1, and overflow  $A_{OvF}$  varies between 10 and 50% of the  $A$  reservoir thickness.

Once the size of the parameters space was set, 50,000 MT-DREAM<sub>ZS</sub> iterations were run on 5 Markov chains (see Supplementary Material B). Over all these iterations, 1401 sets of parameters were validated by the Metropolis ratio. Each parameter quickly converged to a best-fit value. Parameters appear uncorrelated with each other, highlighting their independence (see the correlation matrix in Supplementary Materials C). Out of the 1401 simulations, the last 25% are considered as a posteriori simulation and are used for the analysis of the results. The posterior PDFs are thus reported on Figure 9.

The parameter exploration converges to a restricted range of posterior distributions. The best-fit values of  $SW_{\max}$ ,  $P_{\max}$  and  $A_{\max}$  are 471 mm, 1.01 and 22.83 m respectively, each one is close to the probability peak (Figure 9—top panel). Their posterior distribution at the 95% confidence level is restricted between 447 and 495 mm, 1 and 1.11 m, and 22.2 and 23.66 m, respectively. These results demonstrate fairly good confidence in the parameters estimates despite the low constraint on the deep reservoir. The area is also fairly well constrained, between 17.1 and 17.6 km<sup>2</sup>, with an optimum at 17.3 km<sup>2</sup>. The two indexes,  $BFI$  and  $PFI$ , are also focused around their optimal values 0.4 and 0.43 respectively (range from 0.38 to 0.42 and 0.41 to 0.44). Finally, the optimum of  $A_{OvF}$  is located at 0.174 (range from 0.173 to 0.175). The good convergence of the Bayesian approach is remarkable on the simulated flow curve since the 95% prediction interval is narrow (Figure 9—bottom panel).

With a KGE score of 0.87, the simulation represents fairly well the dynamics of the artesian springs, especially at the beginning (2015–2017) and for the hydrological cycle (2019–2020). In between, the quality of the simulation is slightly altered. For example, in 2018, the simulation overestimates the discharge during the wet season and early 2019, the simulated discharge keeps decreasing while the observations increased. Also, the simulation underestimates the



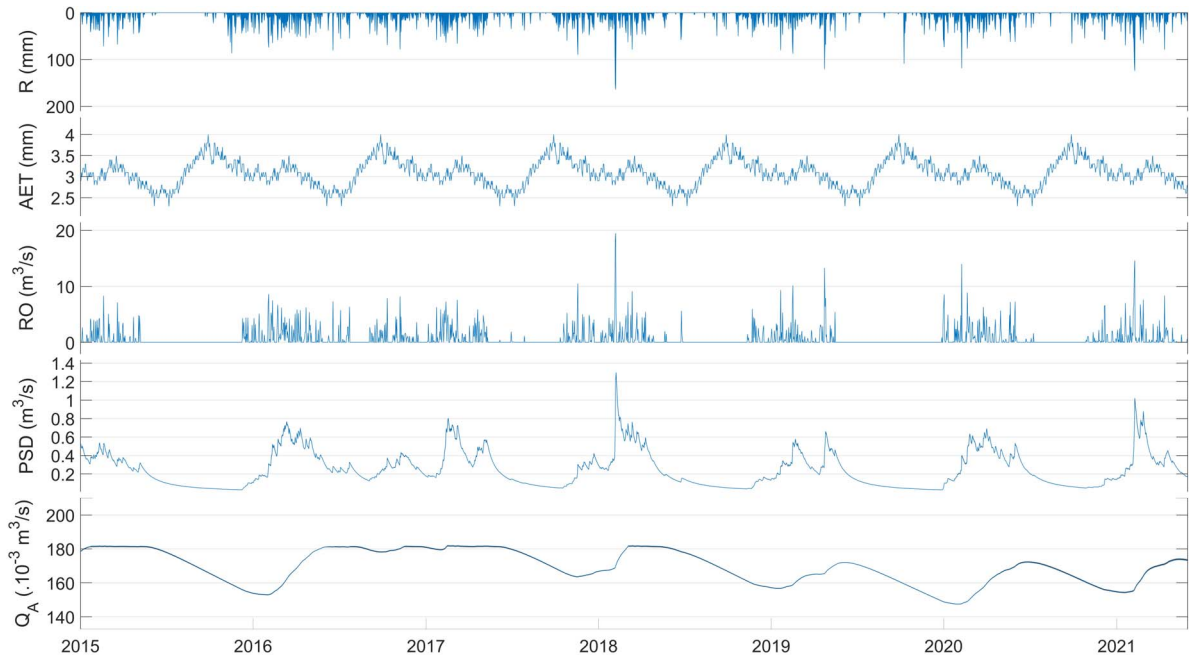
**Figure 9.** Probabilistic exploration of lumped model parameters. The red lines represent the seven probability density functions (PDFs) of the parameters. In each line, the posterior distribution of each parameter is defined by dark square. The optimum parameter value is pointed with a blue diamond. The bottom part represents the artesian springs observed discharge with yellow dots, the optimum simulated discharge in blue with the 95% range parameters uncertainty. The 95% range parameter uncertainty represents the variability in simulated spring discharge as a function of the posterior distribution of parameters. It does not represent a comprehensive assessment of the model uncertainties. The parameter uncertainty is small since the posterior distribution of the parameters is narrow.

recharge of the aquifer at the end of 2020 and the start of 2021. Although the simulation is less efficient during these periods, the model error does not exceed 6% of the observed discharge.

The last model validation is based on the use of river specific flows measured in the field. To do this, our model allows computing the discharge of the different water flows over the watershed (Figure 10). The simulated watershed runoff ( $RO$ ) varies between 0 and 7  $m^3/s$  on average with larger floods up to 16  $m^3/s$ . The perched springs discharge ( $PSD$ ) varies between 0.2 and 1.5  $m^3/s$ . The artesian springs discharge ( $Q_A$ ) is completely buffered between 150 and 180 L/s.

From these simulations, we extracted the average values corresponding to the months of field measurements. In December 2020, the runoff mea-

sured in the field was 18 to 20  $L/s/km^2$ . In comparison, the median specific flow of the model during December is 13  $L/s/km^2$ . However, the simulated floods occurring during this period have a strong impact on the average specific flow (50  $L/s/km^2$ ). Considering the dry/wet period comparison, while in the field the river flow is 8 times higher during the wet season, the model simulates an increase of a factor of 5.5. Nevertheless, the high-frequency variability of the river flow simulated during the wet season is directly related to rainfall dynamics. This high variability complicates the analysis of punctual river measurements. It seems therefore impossible to constrain the model without continuous monitoring of river flows. Herein lies the most important margin of progress to improve the simulations. The implementation of continuous monitoring of river flow, even limited in time,



**Figure 10.** Simulation inputs and outputs simulated at daily basis: rainfall ( $R$ ), actual evapotranspiration ( $AET$ ), runoff ( $RO$ ), perched spring discharge ( $PSD$ ), and the artesian spring discharge ( $Q_A$ ). The climatic forcings are in mm/day while the surface and groundwater flows are in  $m^3/s$ .

will improve the robustness of the simulations. This monitoring should be made at the outlet of the surface watershed as well as upstream in the proximal zone to calibrate perched and the artesian aquifer water balance. To do so, the upstream river should be instrumented to calibrate  $BFI$  parameter, as well as the downstream one for  $PFI$ .

### 3.2.3. Interest of probabilist approach for andesitic hydrosystem simulation

In this study, we developed a hydrological model aiming to represent the water cycle on the flanks of an andesitic volcano. To do so, we followed a two-step framework including (i) the proposal of a model structure as simple as possible to represent the complex and heterogenous surface- and ground-water behaviors; (ii) a probabilistic approach aiming at determining model parameters using a wider diversity of punctual measurements data.

Regarding the first issue, we focused on lumped reservoir models to balance the number of parameters with available data. The flexible structure of such models can be easily adapted to the complex

hydrogeological characteristics of the environment [Dubois *et al.*, 2020, Mazzilli *et al.*, 2019]. Such a modelling approach was already implemented in a French West Indies watershed by Charlier *et al.* [2008] who set up a conceptual hydrogeological model considering both surface and groundwater flows. Our study goes further by integrating multidisciplinary analyses combining geophysical, geochemical, and isotopic measurements to further constrain the model structure. Both studies show the ability of parsimonious reservoir models to represent complex subsurface flows of minimally instrumented watersheds. In these conditions, we assume that this parsimonious approach of lumped parameters models is the most consistent approach to represent the complexity of the volcanic environment. Process-based models would provide a finer representation of physical processes but would require extensive additional data to set up and validate. Given the amount of available data and the large heterogeneities of the studied areas, the feasibility of such a modelling approach is still to be proven.

Regarding the second aspect, we developed a

probabilistic approach aiming at integrating punctual measurements data from a wide variety of disciplines. Prior distributions of parameter values were derived using geophysical imagery, geochemical and isotopic analyses as well as expert knowledge, which allows strengthening the confidence in model simulations. In the future, other datasets could be integrated to improve the calibration/validation process as piezometric data [Charlier *et al.*, 2008], geophysical monitoring [Lesparre *et al.*, 2020], or groundwater dating [Kolbe *et al.*, 2016]. The integration of a wider diversity of dataset in water simulation increase the confidence in the processes captured by the model while limiting equifinality issues during the calibration of the model parameters.

Finally, the probabilistic approach provides posterior distribution of parameter values and an estimation of parameter uncertainty instead of a unique set of parameters. This allows representing fairly the parametric uncertainties of the model that are inherently large in poorly gauged environments. However, these parametric results have to be carefully interpreted as (i) DREAM algorithm try to disentangle error source but Bayes' law suffers from the interaction between individual sources of error (input, output, parameter, and model structure error); and (ii) the DREAM algorithm tends to provide a tighter posterior distribution of parameters than other informal Bayesian approach [Vrugt *et al.*, 2009]. Besides, the analysis of posterior distribution is a first step towards the identification of ways to improve the model structure and parametrization by further field measurement.

#### 4. Conclusion

The objective of this work was to integrate the multidisciplinary characterization of a previously unknown watershed to build a representative hydrological model of the water cycle in an andesitic volcanic aquifer. This characterization was needed for a better management of the wide variety of water uses for human needs within the watershed, while supporting biodiversity. Unfortunately, this characterization was hampered by a significant lack of measurements. To overcome this difficulty, we defined a parsimonious and realistic conceptual model

in order to understand the hydrological behavior in function of the climate forcing.

The multidisciplinary field characterization provides a characterization of hydrosystem structure as well as the identification of the aquifers distributed over the watershed. Our study revealed an extensive artesian aquifer at the foot of the andesite volcano confined by detritic clayey deposits. When the thickness of this latter decreases, due to erosion and/or thinner deposits, the aquifer supports significant artesian springs (150–190 L/s). On the slopes of the volcano, the alternation of lava flows and detritic formations allows the development of perched aquifers with small extension. While the recharge of both types of aquifers occurs on the volcano slopes, extension and outlet location of the aquifers induce spatially different behaviors. This characterization based on data analysis and field observations allowed to set up a conceptual lumped model characterized by various reservoirs representing closely the different functions observed. This model is used to simulate the artesian springs discharge.

Due to data scarcity, we improved the hydrological simulation by implementing a probabilistic estimation of the model parameters. The field characterization allowed to define a uniform a priori distribution for some of the parameters. From these distributions, the use of several Markov chains allowed to test a wide range of parameter sets while limiting the influence of local optima. The Bayesian simulation allows providing a PDF to analyze the uncertainty of parameter values. Finally, the uncertainty could be represented on the simulated water flows to estimate its impact.

This combination of a field-based approach and the probabilistic integration of field measurements allows reinforcing the simulation of ungauged watersheds. Despite a limited constraint on surface river flow, this simulation allows today to better manage the groundwater resource in this strategic sector. In the future, new field acquisitions could be guided by the simulation results. To go further, it would be possible to integrate into the model other reservoirs to represent the main anthropogenic activities that develop rapidly in the watershed. First, an exchange function with urban areas could be added. This function would withdraw groundwater from the deep aquifer and discharge water to the surface. A second reservoir representing the rice terraces would

allow the model to be refined by integrating the delayed effect of surface flows, the increase in evaporation, and the influence on the recharge of agricultural practices.

### Conflicts of interest

Authors have no conflict of interest to declare.

### Acknowledgements

This project has been carried out within the collaboration between Sorbonne University (METIS laboratory), Water Institute by Evian, Danone Aqua group and the Universitas Padjajaran (UNPAD). This research was funded by Sorbonne University and Danone Aqua, in the framework of Danone Aqua Waterstewardship acceleration plan on watershed preservation. The authors are grateful to Damien Jougnot (Sorbonne University—METIS laboratory) for his help in the implementation of the MT-DREAM<sub>ZS</sub> algorithm.

### Supplementary data

Supporting information for this article is available on the journal's website under <https://doi.org/10.5802/crgeos.192> or from the author.

### References

- Aldrian, E. and Djamil, Y. S. (2008). Spatio-temporal climatic change of rainfall in East Java Indonesia. *Int. J. Climatol.*, 28(4), 435–448.
- Bénard, B., Famin, V., Agrinier, P., Aunay, B., Lebeau, G., Sanjuan, B., Vimeux, F., Bardoux, G., and Dezares, C. (2019). Origin and fate of hydrothermal fluids at Piton des Neiges volcano (Réunion Island): A geochemical and isotopic (O, H, C, Sr, Li, Cl) study of thermal springs. *J. Volcanol. Geotherm. Res.*, 392, article no. 106682.
- Buwono, N. R., Risjani, Y., and Soegianto, A. (2021). Distribution of microplastic in relation to water quality parameters in the Brantas River, East Java, Indonesia. *Environ. Technol. Innov.*, 24, article no. 101915.
- Cerbelaud, A., Lefèvre, J., Genthon, P., and Menkes, C. (2022). Assessment of the WRF-hydro uncoupled hydro-meteorological model on flashy watersheds of the Grande Terre tropical island of New Caledonia (South–West Pacific). *J. Hydrol. Reg. Stud.*, 40, article no. 101003.
- Chang, S. W., Chung, I.-M., Kim, M.-G., and Yifru, B. A. (2020). Vulnerability assessment considering impact of future groundwater exploitation on coastal groundwater resources in northeastern Jeju Island, South Korea. *Environ. Earth Sci.*, 79(22), article no. 498.
- Charlier, J.-B., Cattan, P., Moussa, R., and Voltz, M. (2008). Hydrological behaviour and modelling of a volcanic tropical cultivated catchment. *Hydrol. Process.*, 22(22), 4355–4370.
- Charlier, J.-B., Lachassagne, P., Ladouche, B., Cattan, P., Moussa, R., and Voltz, M. (2011). Structure and hydrogeological functioning of an insular tropical humid andesitic volcanic watershed: A multi-disciplinary experimental approach. *J. Hydrol.*, 398(3–4), 155–170.
- Craig, H. (1961). Isotopic variations in meteoric waters. *Science*, 133(3465), 1702–1703.
- Dahlin, T. and Zhou, B. (2004). A numerical comparison of 2D resistivity imaging with 10 electrode arrays. *Geophys. Prospect.*, 52(5), 379–398.
- de Marsily, G. (2020). Will we soon run out of water? *Ann. Nutrition Metabolism*, 76(1), 10–16.
- de Marsily, G., Delhomme, J.-P., Delay, F., and Buoro, A. (1999). Regards sur 40 ans de problèmes inverses en hydrogéologie. *C. R. Acad. Sci. Ser. IIA - Earth Planet. Sci.*, 329(2), 73–87.
- Descloitres, M., Ritz, M., Robineau, B., and Courteaud, M. (1997). Electrical structure beneath the eastern collapsed flank of Piton de la Fournaise volcano, Reunion Island: Implications for the quest for groundwater. *Water Resour. Res.*, 33(1), 13–19.
- Dubois, E., Doummar, J., Pistre, S., and Larocque, M. (2020). Calibration of a semi-distributed lumped karst system model and analysis of its sensitivity to climate conditions: The example of the Qachqouch karst spring (Lebanon). *Hydrol. Earth Syst. Sci. Discussions*, 24, 1–25.
- Dumont, M., Peltier, A., Roblin, E., Reninger, P.-A., Barde-Cabusson, S., Finizola, A., and Ferrazzini, V. (2019). Imagery of internal structure and destabilization features of active volcano by 3D high res-

- olution airborne electromagnetism. *Sci. Rep.*, 9(1), 1–11.
- Dumont, M., Reninger, P. A., Aunay, B., Pryet, A., Jougnot, D., Join, J. L., Michon, L., and Martelet, G. (2021). Hydrogeophysical characterization in a volcanic context from local to regional scales combining airborne electromagnetism and magnetism. *Geophys. Res. Lett.*, 48(12), article no. e2020GL092000.
- Dumont, M., Saadi, M., Oudin, L., Lachassagne, P., Nugraha, B., Fadillah, A., Bonjour, J.-L., Muhammad, A., Hendarmawan, Dörfliger, N., and Plagnes, V. (2022). Assessing rainfall global products reliability for water resource management in a tropical volcanic mountainous catchment. *J. Hydrol. Reg. Stud.*, 40, article no. 101037.
- Gharari, S., Hrachowitz, M., Fenicia, F., Gao, H., and Savenije, H. H. G. (2014). Using expert knowledge to increase realism in environmental system models can dramatically reduce the need for calibration. *Hydrol. Earth Syst. Sci.*, 18(12), 4839–4859.
- Gómez-Delgado, F., Roupsard, O., le Maire, G., Tau-gourdeau, S., Pérez, A., van Oijen, M., Vaast, P., Rapidel, B., Harmand, J. M., Voltz, M., Bonnefond, J. M., Imbach, P., and Moussa, R. (2011). Modelling the hydrological behaviour of a coffee agroforestry basin in Costa Rica. *Hydrol. Earth Syst. Sci.*, 15(1), 369–392.
- Gresse, M., Uyeshima, M., Koyama, T., Hase, H., Aizawa, K., Yamaya, Y., Morita, Y., Weller, D., Rung-Arunwan, T., Kaneko, T., Sasai, Y., Zlotnicki, J., Ishido, T., Ueda, H., and Hata, M. (2021). Hydrothermal and magmatic system of a volcanic island inferred from magnetotellurics, seismicity, self-potential, and thermal image: An example of Miyakejima (Japan). *J. Geophys. Res. Solid Earth*, 126(6), article no. e2021JB022034.
- He, B., Takase, K., and Wang, Y. (2008). Numerical simulation of groundwater flow for a coastal plain in Japan: Data collection and model calibration. *Environ. Geol.*, 55(8), 1745–1753.
- Hrachowitz, M. and Clark, M. P. (2017). HESS Opinions: The complementary merits of competing modelling philosophies in hydrology. *Hydrol. Earth Syst. Sci.*, 21(8), 3953–3973.
- Join, J. L., Coudray, J., and Longworth, K. (1997). Using principal components analysis and Na/Cl ratios to trace groundwater circulation in a volcanic island: The example of Reunion. *J. Hydrol.*, 190(1–2), 1–18.
- Join, J.-L., Folio, J.-L., Bourhane, A., and Comte, J.-C. (2016). Groundwater resources on active Basaltic Volcanoes: Conceptual models from La Réunion Island and Grande Comore. In Bachelery, P., Lenat, J.-F., Muro, A. D., and Michon, L., editors, *Active Volcanoes of the Southwest Indian Ocean*, pages 61–70. Springer, Berlin, Heidelberg.
- Khasanah, N., Tanika, L., Pratama, L. D. Y., Leimona, B., Prasetyo, E., Marulani, F., Hendriatna, A., Zulkarnain, M. T., Toulrier, A., and van Noordwijk, M. (2021). Groundwater-extracting rice production in the Rejoso Water-Shed (Indonesia) reducing urban water availability: Characterisation and intervention priorities. *Land*, 10(6), article no. 586.
- Kolbe, T., Marçais, J., Thomas, Z., Abbott, B. W., de Dreuz, J.-R., Rousseau-Gueutin, P., Aquilina, L., Labasque, T., and Pinay, G. (2016). Coupling 3D groundwater modeling with CFC-based age dating to classify local groundwater circulation in an unconfined crystalline aquifer. *J. Hydrol.*, 543, 31–46.
- Lachassagne, P., Aunay, B., Frissant, N., Guilbert, M., and Malard, A. (2014). High-resolution conceptual hydrogeological model of complex basaltic volcanic islands: A Mayotte, Comoros, case study. *Terra Nova*, 26(4), 307–321.
- Laloy, E., Linde, N., Jacques, D., and Vrugt, J. A. (2015). Probabilistic inference of multi-Gaussian fields from indirect hydrological data using circulant embedding and dimensionality reduction. *Water Resour. Res.*, 51, 4224–4243.
- Lesparre, N., Girard, J.-F., Jeannot, B., Weill, S., Dumont, M., Boucher, M., Viville, D., Pierret, M.-C., Legchenko, A., and Delay, F. (2020). Magnetic resonance sounding measurements as posterior information to condition hydrological model parameters: Application to a hard-rock headwater catchment. *J. Hydrol.*, 587, article no. 124941.
- Lo, W., Purnomo, S. N., Sarah, D., Aghnia, S., and Hardini, P. (2021). Groundwater modelling in urban development to achieve sustainability of groundwater resources: A case study of Semarang City, Indonesia. *Water*, 13(10), article no. 1395.
- Marescot, L., Loke, M. H., Chapellier, D., Delaloye, R., Lambiel, C., and Reynard, E. (2003). Assessing reliability of 2D resistivity imaging in mountain permafrost studies using the depth of investigation in-



- dex method. *Near Surf. Geophys.*, 1(2), 57–67.
- Mazzilli, N., Guinot, V., Jourde, H., Lecoq, N., Labat, D., Arfib, B., Baudement, C., Danquigny, C., Dal Soglio, L., and Bertin, D. (2019). KarstMod: A modelling platform for rainfall—discharge analysis and modelling dedicated to karst systems. *Environ. Model. Softw.*, 122, article no. 103927.
- Monteith, J. L. (1965). Evaporation and environment. *Symp. Soc. Experiment. Biol.*, 19, 205–234.
- Mottes, C., Lesueur-Jannoyer, M., Charlier, J.-B., Carles, C., Guéné, M., Le Bail, M., and Malézieux, E. (2015). Hydrological and pesticide transfer modeling in a tropical volcanic watershed with the WATPASS model. *J. Hydrol.*, 529, 909–927.
- Oldenburg, D. W. and Li, Y. (1999). Estimating depth of investigation in dc resistivity and IP surveys. *Geophysics*, 64(2), 403–416.
- Perrin, C., Michel, C., and Andréassian, V. (2003). Improvement of a parsimonious model for streamflow simulation. *J. Hydrol.*, 279(1), 275–289.
- Satapona, A., Putra, D. P. E., and Hendrayana, H. (2018). Groundwater flow modeling in the Malioboro, Yogyakarta, Indonesia. *J. Appl. Geol.*, 3(1), 11–22.
- Selles, A., Deffontaines, B., Hendrayana, H., and Violette, S. (2015). The eastern flank of the Merapi volcano (Central Java, Indonesia): Architecture and implications of volcanoclastic deposits. *J. Asian Earth Sci.*, 108, 33–47.
- Suriadikusumah, A., Mulyani, O., Sudirja, R., Sofyan, E. T., Maulana, M. H. R., and Mulyono, A. (2021). Analysis of the water quality at Cipeusing river, Indonesia using the pollution index method. *Acta Ecol. Sin.*, 41(3), 177–182.
- Tadono, T., Ishida, H., Oda, F., Naito, S., Minakawa, K., and Iwamoto, H. (2014). Precise global DEM generation by ALOS PRISM. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inform. Sci.*, II4, 71–76.
- Thornthwaite, C. W. (1948). An approach toward a rational classification of climate. *Geogr. Rev.*, 38(1), 55–94. JSTOR.
- Toulier, A., Baud, B., de Montety, V., Lachassagne, P., Leonardi, V., Pistre, S., Dautria, J.-M., Hendrayana, H., Miftakhul Fajar, M. H., Satrya Muhammad, A., Beon, O., and Jourde, H. (2019). Multidisciplinary study with quantitative analysis of isotopic data for the assessment of recharge and functioning of volcanic aquifers: Case of Bromo-Tengger volcano, Indonesia. *J. Hydrol. Reg. Stud.*, 26, article no. 100634.
- Toulier, A., Lachassagne, P., Hendrayana, H., Fadillah, A., and Jourde, H. (2022). A cost-effective device and methodology to compute aquifer transmissivity and piezometry from free-flowing artesian wells. *Hydrogeol. J.*, 30, 1917–1931.
- Vittecoq, B., Deparis, J., Violette, S., Jaouën, T., and Lacquement, F. (2014). Influence of successive phases of volcanic construction and erosion on Mayotte Island's hydrogeological functioning as determined from a helicopter-borne resistivity survey correlated with borehole geological and permeability data. *J. Hydrol.*, 509, 519–538.
- Vittecoq, B., Reninger, P.-A., Lacquement, F., Martelet, G., and Violette, S. (2019). Hydrogeological conceptual model of andesitic watersheds revealed by high-resolution heliborne geophysics. *Hydrol. Earth Syst. Sci.*, 23(5), 2321–2338.
- Vrugt, J. A., ter Braak, C. J. F., Gupta, H. V., and Robinson, B. A. (2009). Equifinality of Formal (DREAM) and Informal (GLUE) Bayesian Approaches in Hydrologic Modeling? *Stoch. Environ. Res. Risk Assess.*, 23(7), 1011–1026.





Research article

Geo-hydrological Data & Models

# Contribution of the dynamic visualization of a 3D hydrogeological conceptual model to the participatory management of groundwater resources

Arnauld Malard<sup>Ⓢ,a</sup>, Nathalie Dörfliger<sup>Ⓢ,\*,b</sup>, Pierre-Yves Jeannin<sup>Ⓢ,a</sup> and Silvan Laube<sup>c</sup>

<sup>a</sup> ISSKA, Institut Suisse de Spéléologie et de Karstologie, Rue de la Serre 68, CH-2301 La Chaux-de-Fonds, Switzerland

<sup>b</sup> Water Institute by Evian, DANONE, 11 avenue du Général Dupas, BP 87, F-74503 Evian-les-Bains, France

<sup>c</sup> i4ds, Institute for Data Science, Fachhochschule Nordwestschweiz, CH-5210 Windisch, Switzerland

E-mails: [arnauld.malard@isska.ch](mailto:arnauld.malard@isska.ch) (A. Malard), [nathalie.dorfliger@danone.com](mailto:nathalie.dorfliger@danone.com) (N. Dörfliger), [pierre-yves.jeannin@isska.ch](mailto:pierre-yves.jeannin@isska.ch) (P.-Y. Jeannin), [silvan.laube@fhnw.ch](mailto:silvan.laube@fhnw.ch) (S. Laube)

**Abstract.** Groundwater is an essential water resource for domestic use, agriculture and other socioeconomic activities, as well as ecological services around the world. Participative approach to water resources management is essential, and requires all stakeholders to have sufficient knowledge and understanding of processes taking place, from rainfall to recharge, drainage and discharge [De Marsily and Besbes, 2017]. Understanding unseen processes require abstraction and imagination. In many cases, detailed 3D conceptual models may help non-specialists to understand underground environments and thereby address questions related to the management of groundwater resources [Jeannin et al., 2013, Malard, 2018]. In order to increase the understanding of dynamic hydrogeological processes, a flow animation module (FAM) has been developed within Visual KARSYS (<https://www.visualkarsys.com/>); a web tool dedicated to 3D conceptual modelling of complex aquifers. Concepts and principles of 3D conceptual models and of the new FAM are presented and discussed in this paper with an application to a volcanic aquifer in Indonesia.

**Keywords.** Groundwater, Conceptual model, 3D visualization, Flow animation, Management, Stakeholders.

*Manuscript received 1 April 2022, revised 10 August 2022, accepted 6 September 2022.*

## 1. Introduction

Groundwater resources are essential for humanity as well as for nature and are the most important reserve of fresh water on Earth. Groundwater represents an important source of water supply for peo-

ple and for the development of livelihoods, economic and living conditions around the world. It is thus critically important for sustainable development as it provides water for domestic use, agriculture, industries and other socioeconomic sectors and as well for many ecological services. It is estimated that at global scale, groundwater withdrawals have increased sixfold in the past 60 years [Van der

\* Corresponding author.

Gun, 2012], linked to population growth, industrialization, urbanization and change of diet [De Marsily and Besbes, 2017]. Depletion of aquifers<sup>1</sup> with important drawdowns of water table occurred in several regions around the world, leading in some cases to irreversible degradation of groundwater chemical quality and to land subsidence [Bierkens and Wada, 2019]. Groundwater resources are thus threatened in terms of both quantity and quality, due to anthropogenic pressure and climate changes. Managing them is a great challenge, as groundwater is a distributed and a hidden resource. The resource being limited, aquifers should be utilized collectively and responsibly, requiring a sustainable management approach involving all stakeholders. However, in practice, water tends to be pumped out independently, leading to a chaotic and non-sustainable exploitation. To improve the situation, all stakeholders need to have sufficient knowledge and understanding of processes taking place, from rainfall to recharge, drainage and discharge [De Marsily and Besbes, 2017]. Without an inclusive and integrated water resources management of groundwater, socio-political issues may arise and worsen (conflicts of use), specifically in areas where water scarcity is already a burden.

Key groundwater knowledge refers to understanding the basic principles of hydrogeology, namely definition of aquifer and aquitard among the underground layers, spatial heterogeneity of underground reservoir as well as the abstract concept of a groundwater table. As mentioned by [De Marsily, 1986], a groundwater table is an abstract concept of the reality of the underground water world. One depicts the spatial representation of a water body with its upper and lower boundaries, and possibly also its side limits. This can be compared to a fog bank.

The perception and understanding of groundwater by stakeholders around the world are diverse, largely because of the regional and cultural dimensions related to the underground. One vision is related directly to springs, natural water holes and wetlands, and their observation across seasons. For

example, in Australia, the Rainbow Serpent is a spiritual entity recognized by Aboriginal people and credited with the creation of the Australian landscape. According to a regional study in the Northern Territory with a comparative analysis of hydrogeology map and Aboriginal knowledge, the outline of the major regional aquifer in central Arnhem Land has a rainbow serpent shape. Local people are aware that groundwater inflow supports flow of major rivers during dry season, and that they are living on top of water stored underground [Zaar *et al.*, 2002]. Similarly, in other parts of the world, outflowing water at springs or at artesian free-flowing water wells as in volcanic sedimentary plains in Java Island (Indonesia); water is considered as a gift from the goddess of nature. Putting a valve on these free-flowing water wells can lead to fears among farmers about the possible scarcity of water. Finally, in European countries for example, rather than imaging a water table as a fog bank, as defined by de Marsily, the common image of groundwater is an underground river or even a lake!... The popular imagination about groundwater is definitively not the unique perception shared by all stakeholders, as strong effort has been made by scientists to transfer knowledge (e.g. The Groundwater project<sup>2</sup> or several videos on YouTube of sand-boxes).

The explanation of complex hydrogeological processes (from recharge to discharge; from mineralization to contamination) calls for simplification and didactic explanations for non-specialists. Driving principles should be explained with precise key-elements; for instance, the role of vegetation growth and of air temperature on the groundwater recharge or the range of groundwater travel times as it varies over several orders of magnitude.

All knowledge related to hydrogeology principles are usually carried out via static schematization of the reality, most of the time using one cross section to address the vertical dimension of an aquifer. Both spatial and temporal dimensions are difficult to be grasped by non-specialists: stakeholders, water users, farmers or people from industrial and domestic water sectors. With this aim, a conceptual model, which is a schematization of the complex

<sup>1</sup>In this paper, the term "aquifer" refers to both unsaturated and saturated zones of the permeable rocks (European definition).

<sup>2</sup>On Line platform for Groundwater Knowledge (<https://gw-project.org>).

reality of the invisible groundwater world, is used at the scale of a hydrogeological system or catchment unit. However, it often remains static and simplified in 2D. 3D groundwater flow animations are not always available for all studies. Comprehensive full data set to set up such 3D groundwater flow numerical model is necessary. In the perspective of groundwater resource management, 3D groundwater flow, modeling is not realized, and not used as a decision support tool. Occasionally, very simplified sketches or even animation clips are used without being able to transfer accurately hydrogeological processes, at least not anchored in the real local setting. It may even carry an incorrect understanding of the complexity of groundwater systems. The velocity of hydrogeological process is hardly transcribed. Demonstration of sand-tank groundwater models, i.e. physical models, concern one type of hydrogeological setting, porous media generally, and are not applicable for all other aquifer types.

In integrated water resources management and especially in water stewardship actions concerning groundwater, the engagement of actors in a collaborative management is acknowledged to be important. A collective action plan and a strong involvement of a variety of stakeholders are recognized to lead to a more efficient and sustainable management of aquifers. Co-management is a key element within water policies, involving power sharing between water users and government entities, as reported in [Shalsi *et al.*, 2019]. Such a multilateral approach implies mutual learning between water users, policy makers, technical experts on groundwater resources and management, including culture heritage, community values and technical expertise. Stakeholders' ignorance concerning groundwater resources and the impacts of their various activities, dissociates them from any long-term responsibility [De Marsily and Besbes, 2017]. Sharing knowledge and information, even participation to monitoring activities provides a better understanding of the groundwater resources and their hydrodynamics under climatic and anthropic conditions (withdrawals, land use changes, etc.).

Stakeholders may be from various backgrounds; those directly concerned about water uses by their activities (e.g. agriculture under irrigation, industry, water supply) or within organizations (author-

ities, companies, NGOs); those who will take financial decisions. These stakeholders play a role and have an influence, based on their understanding of water resources issues and solutions. Their interactions and contribution will lead to a more efficient and/or rapid implementation of water resources management. Their involvement in a participatory approach can rely on several methodologies [e.g. Baccar *et al.*, 2021, Etienne *et al.*, 2011], including stakeholder mapping and analysis regarding the issues to be addressed via different solutions—technical ones or policies [e.g. Reed *et al.*, 2009]. Everyone has his own knowledge based on observation, self-monitoring and perception about groundwater. Engaging stakeholders implies exchanges among water users, technical experts, policy makers, and also general managers. For communication of knowledge about groundwater, technical experts use sketches, graphs, maps, and possible outcomes of mathematical modelling, but also translate that knowledge into narratives, in the frame of participatory scenario analysis methods [Rouillard *et al.*, 2022].

From this perspective, Danone Waters looked for a way to combine 3D representations of subsurface of earth and a visualization of groundwater water flow from rainfall to springs. KARSYS is a pragmatic approach for developing 3D conceptual models of groundwater systems in karst regions. It was developed at the Swiss Institute for Speleology and Karst Studies [Ballesteros *et al.*, 2015, Jeannin *et al.*, 2013, Malard, 2018] and has been made available online via the Visual KARSYS website. In the frame of the collaboration, new developments have led to addition of a Flow Animation Module (FAM) which could transmit the principles of hydrogeology to stakeholders associated with the groundwater resources management. For a few years now, KARSYS is being applied to hydrogeological systems in volcanic island environments too, highlighting also the ability of this approach to be applied to non-karstic environments. The aim of this development is to represent explicitly the underground processes, which cannot be seen directly and thus require abstract thinking for a non-specialist to imagine these over different time and spatial scales (from mm to km and from a few hours to multiple decades). Sharing such concepts to non-initiated persons is a real challenge but essential to manage water resources.

In this paper, after a brief review about relevant

general hydrogeological concepts, we present the main principles of the KARSYS approach. Then, the principles of 3D dynamic visualization are sketched, as well as the way FAM was included within Visual KARSYS. Finally, the implementation to a case study in East Java, in a volcanic hydrogeological setting, is presented.

## 2. Material and method

### 2.1. *Hydrogeological conceptual model—short review*

A conceptual model in hydrogeology is a qualitative and schematic (i.e., simplified) description of an existing hydrosystem based on characteristics and processes (hydraulics, gravity, permeability, etc.) as defined by [Anderson and Woessner, 1992] or [Ghasemizadeh *et al.*, 2012]. Conceptual models may be generic (focusing on processes and related to well-defined contexts: coastal aquifers, alluvial, hypogenic, etc.) or site-specific (focusing on site characteristics, with possible high-frequency monitoring on a demo site [Cooper *et al.*, 2018]). They can be static (steady state only) or dynamic if they address different water stages (i.e., low-flow, high-flow, drought, etc.).

Theoretically, conceptual models are seen as a prerequisite for the construction and analysis of numerical models [Betancur *et al.*, 2012], as they allow to point out the main characteristics and processes which must be computed in order to reproduce the expected dynamics (i.e., recharge, effects of pumping, intrusion of salted water, etc.). In parallel, numerical models are expected to support and to confirm the conceptual model with a more quantitative approach (Figure 1).

In practice, the building of the conceptual model prior to the design of a numerical model often remains implicit. The modeler usually keeps the qualitative and schematic description of the studied hydrosystem in mind. The results of the numerical model are then presented as the explicit representation of the hydrosystem, making it difficult to understand the main characteristics and processes (recharge, evapotranspiration (ET), transfer in the unsaturated zone, etc.), to validate them and use them critically. By presenting numerical results in which some of the characteristics and processes remain implicit, the modeler cannot be guaranteed

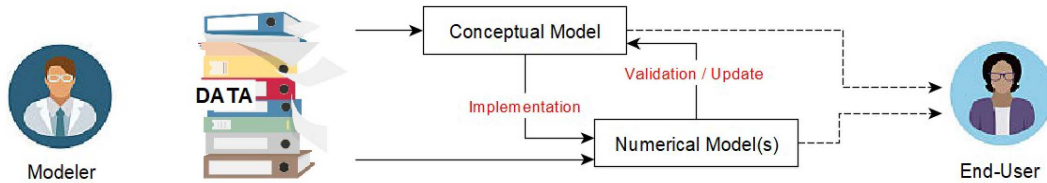
that the end-user has properly perceived the message (the interpretations), and that he will be able to use them for decision making.

In order to better communicate the concepts, processes and results, we can build an explicit conceptual model in which principles are spatially presented, according to data of the studied site. This can be presented prior to any numerical models or results of simulations to address the basic questions: “where are the main recharge zones”, “where does the groundwater flow through?”, “where could I drill to intersect the water table?”, etc. It also provides an detailed image which is the same for all stakeholders, at least concerning (i) the geometry of the aquifer, (ii) the way groundwater flows and (iii) the way it is recharged. Such an unified representation enables easier understanding and further management strategies. Concepts and processes can be validated at any time by simulations.

### 2.2. *“KARSYS” and “Visual KARSYS”*

KARSYS [Jeannin *et al.*, 2013] is a pragmatic approach developed for geologists and hydrogeologists (i.e., “modelers”) to build an explicit conceptual representation of groundwater flow systems in heterogeneous rocks via a 3D geological modeling and a series of principles (i.e., algorithms). An interesting aspect of this approach is to provide guidelines for a step-by-step construction of conceptual models (i.e., the representation of the modeler’s understanding of the groundwater flow system), following a well-defined approach (defined data and principles), and in a detailed and easy to understand way (a 3D model which can be viewed by anyone). Although KARSYS was first developed for karst aquifers, most of the principles can be applied to other heterogeneous aquifers (volcanic, highly fractured, glacial, etc.).

Visual KARSYS is a webtool developed for users to apply the steps of KARSYS by themselves to their site of interest. It offers a database to capture and to format all required data, a 3D viewer and different tools that allow the construction of complex geological models (including faults, erosion phases, etc.) and the extrapolation of the supposed saturated zones in the permeable layers. Another advantage of the webtool is that outputs from Visual KARSYS (data, information, models and maps) can be directly shared between users for being consulted or edited.



**Figure 1.** The modeling workflow assumes the explicit conceptual model to be built before implementing numerical models. These two models can be presented to the end-user for a better understanding of the characteristics and processes taking place in the hydrosystem.

A dedicated output page (Figure 2) has been developed for the end-users to access the KARSYS conceptual model and all associated data. From this page, end-users can extract information from anywhere in the modelled area by generating geological or hydrogeological cross-sections or maps or expected logs of virtual drillholes. They can also export data and results to use them in other applications.

Visual KARSYS not only supports the development of conceptual models for stakeholders (end-users) but also for scientists from geosciences (i) to frame the relevant issues, define outcomes and manage complexity, as a common starting point [Jakeman et al., 2016] and (ii) to improve conceptual models, to test the role of different structures within the groundwater flow path before carrying out a numerical modeling.

### 2.2.1. 3D groundwater (GW) conceptual model development

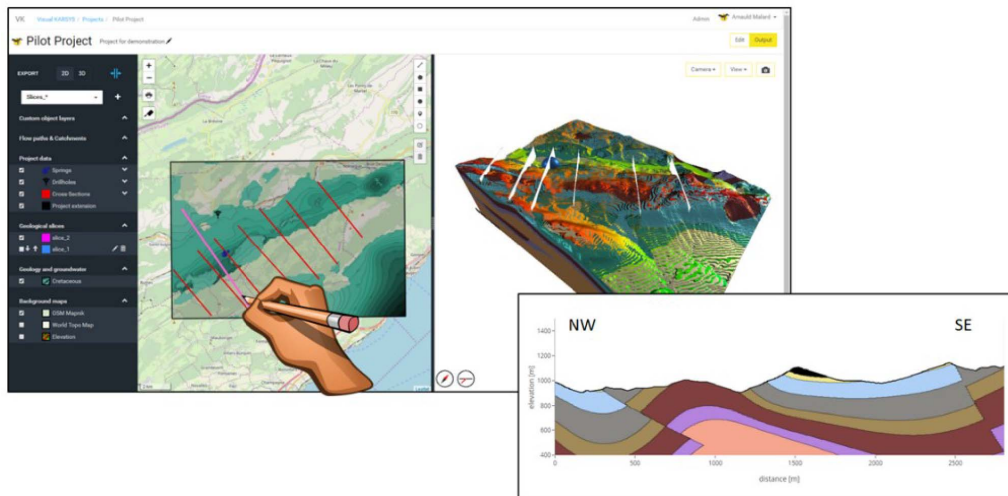
Four basic steps are required for building a KARSYS model. Prior to these steps, users are expected to define the boundaries of the project area, the coordinates system and to enter a series of required data and information in a specific order (geological cross-sections, maps, drillholes, etc.). In step (1), users are required to define the geological units which form the geological framework and to assign hydrogeological roles (pervious vs. impervious, type of permeability, etc.) as well as hydraulic properties (basically porosity and hydraulic conductivity of pervious units, with the exception of the karstified units). Step (2) is dedicated to the building of the 3D geological model by using GmLib [Lopez et al., 2018], a specific library which allows the implicit modelling of geological data in 3D. Step (3) allows the user to infer the location and the geometry of

the main groundwater zones based on hydrological evidences (location of springs, water-head measurements in drillholes, etc.) and a series of principles based on the extrapolation of the hydraulic gradients. In step (4), users can sketch the main expected drainage axes in the aquifer and delineate the boundaries of the catchment area. This last step is completed by FAM which allows a more dynamic visualization, from rain to the aquifer outlets (basically the springs).

### 2.2.2. 3D dynamic conceptual model development

Till the end of 2020, Visual KARSYS merely allowed the user to build “static” conceptual models, basically for low-flow stages. By early 2021, in the frame of a joint project between Danone Waters, SSKA and i4ds, new developments were made in Visual KARSYS in order to allow the users to build a more “dynamic” visualization of the groundwater and surface flow in 3D (FAM). This module is designed for karst, volcanic and other highly heterogeneous aquifers. The aim of the FAM is to depict visually, the way surface water and groundwater flows through the hydrosystem via the following processes: the runoff, the allogenic vs. autogenic recharge, the infiltration through the unsaturated zone, the flow paths in the saturated zone, the overflow from the springs, the effects of drillholes pumping, etc.

In the FAM, users can create different recharge scenarios for different timescales: starting from 30 days, 10 years, to more than 200 years and run the corresponding animation in order to visualize processes taking place at the selected timescale and for the different range of permeabilities. FAM is coded in JavaScript and runs in real time for 120 s. When running, the FAM allows the visualization of the following processes for the selected timescale (Table 1).



**Figure 2.** Overview of the output page. End users can display geological layers on the map or create geological cross-sections or logs for any virtual drillholes.

Generated particles of rainfall are emitted following the provided rainfall time-series:  $X$  mm of rain over one timestep corresponds to  $60X$  particles emitted from the clouds. The evapotranspiration (ET) is generated once particles of rainfall hit the ground surface (Figure 3). When reaching the top of the ground surface, the probability that a particle turns into ET is based on the ET rate for the simulated period (month), modulo, a random variable (*#Check 1*). The remaining particles are partitioned to runoff particles or to recharge particles depending on the hydraulic conductivity ( $K$ ) of the geological unit (*#Check 2*). Runoff particles flow downwards following the steepest slope of the relief. When moving to another place (3D voxel), particles might have the probability to again infiltrate or flow further down as runoff depending on  $K$ .

Infiltrated particles (I) flow vertically through permeable units until reaching: (i) the aquifer basement or (ii) the saturated zone. The velocity of particles is a function of the aquifer's zonation (unsaturated, saturated) and of the hydraulic conductivity of the geological unit ( $K$ ).

In the unsaturated zone, the vertical velocity of the particles ( $v$ ) is in the form:  $v = K \cdot B$  where  $B$  is a constant parameter depending on the timescale.

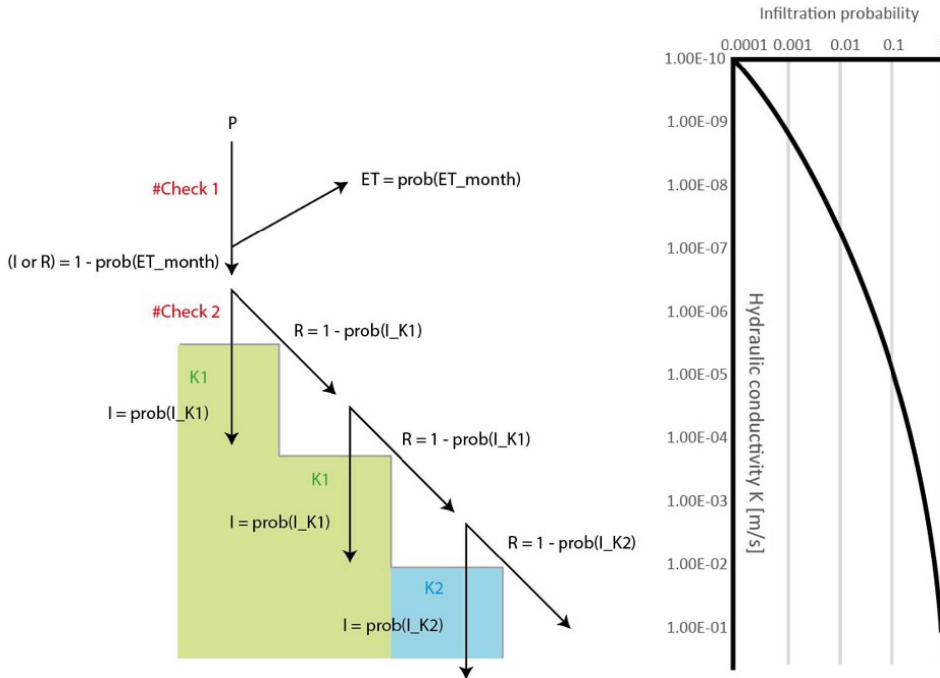
In the phreatic zone, a predefined flow field is generated based on the geometry of the saturated zone and on the location of the springs. This will govern

the trajectory of the particles towards the spring(s). The velocity of the particles is controlled by  $K$  and by an empirical factor based on the relative significance of the different outlets draining the saturated zone ( $F$ ). The equation is expressed as:  $v = C \cdot K^x \cdot F^y$  where  $C$ ,  $x$  and  $y$  are constant parameters.

An example of animation for a few-days event with a located rainfall is given in Figure 4.

Location of the clouds and variations of the rain intensity ( $P$ ) in time can be parametrized. The ET is computed and represented as small clouds (or fumaroles) rising up from the ground surface. The amount of ET is adjustable at monthly time-step.  $P$  minus ET generates runoff ( $R$ ) on impervious rocks and recharge or infiltration ( $I$ ) on permeable rocks depending on  $K$ .  $R$  is represented as moving dark blue particles leaving light blue lines behind. The lines clearly illustrate the merging of flow paths forming streams and rivers at ground surface. Recharge particles are white and flow vertically downwards through the unsaturated zone leaving dark blue vertical lines behind. Particles follow the basement of the aquifer until reaching the saturated zone where they flow towards the spring and return to surface runoff. A glowing effect indicates the groundwater body inflating and deflating according to the input/output balance of the particles entering/leaving the saturated zone. This particle tracking produces a clear idea of the various flow conditions and velocities tak-





**Figure 3.** Left: The partition of the falling rain (P) between evapotranspiration (ET), runoff (R) and infiltration (I) is controlled by two independent checks. #Check 1 applies at the first collision with the ground surface. Particles are converted to ET depending on a randomness probability value based on the ET monthly rate. #Check 2 applies on each collision with the ground surface. Right: Depending on a randomness probability value based on the hydraulic conductivity (K), the particle will infiltrate or will turn to runoff.

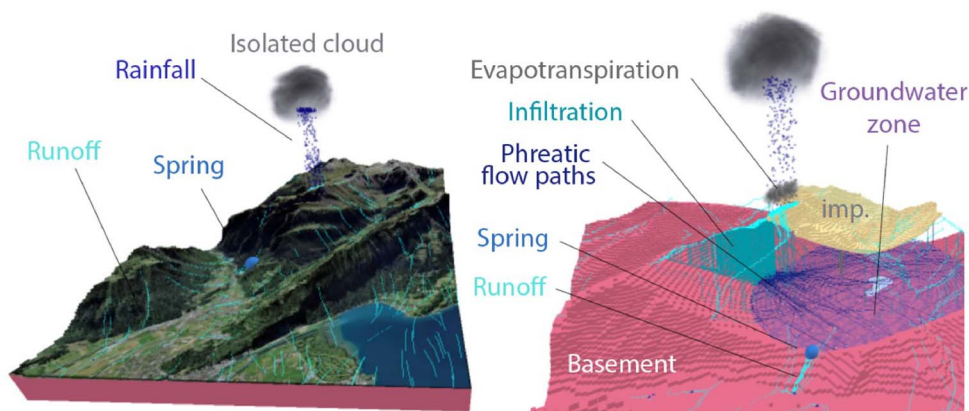
**Table 1.** List of the timescale scenarios available in the FAM and their associated processes

Groundwater processes	Timescale
<ul style="list-style-type: none"> <li>• Partition between ET/runoff/infiltration</li> <li>• Fast recharge/discharge</li> <li>• GW fluctuations visible via glowing effect, and hydrograph</li> <li>• Fast flow (high permeable units, <math>K &gt; 10^{-3}</math> m/s)</li> </ul>	<ul style="list-style-type: none"> <li>Few days event</li> <li>Duration: 30 days</li> <li>Timestep: 6 hours</li> </ul>
<ul style="list-style-type: none"> <li>• Seasonal GW recharge</li> <li>• Seasonal variations of ET</li> <li>• Moderate flow (mid-permeable units, <math>10^{-3} &gt; K &gt; 10^{-6}</math> m/s)</li> </ul>	<ul style="list-style-type: none"> <li>Few seasons</li> <li>Duration: 10 years</li> <li>Timestep: 1 month</li> </ul>
<ul style="list-style-type: none"> <li>• GW residence time</li> <li>• Mineralization and mixing processes</li> <li>• Slow flow (low permeable units, <math>K &lt; 10^{-6}</math> m/s)</li> </ul>	<ul style="list-style-type: none"> <li>Few decades</li> <li>Duration: 240 years</li> <li>Timestep: 2 years</li> </ul>

ing place at ground-surface and in the underground.

Additional tools allow users to take full advantage of the results of the animation. They can hide the 3D

model along selected directions and orientations in order to create sliced views and to focus on particular places without the view being disturbed by the



**Figure 4.** Overview of a short-term animation depicting the expected surface and groundwater flow paths for a rain isolated event at the top of the volcano. Left: surface processes (rainfall, runoff). Right: underground processes (the aquifer layers have been disabled for making the flow lines visible, “imp.” refers to the impervious cover). It is possible to track the trajectory of the infiltrated particles from the surface to the spring and back to the surface downstream the spring.

complexity of the model. A live graphic of the expected hydrological regime for a saturated zone or a spring (i.e., recharge minus discharge) can also be displayed to provide a pseudo-quantitative estimate of the groundwater flow. At the end of the animation, the cumulated rainfall and the respective partition between ET, R and I are given (in mm per unit of time and the total amount over the simulation period).

### 3. 3D Application of dynamic conceptual model of Pasuaran

#### 3.1. Test site of Pasuaran

The 3D GW dynamic conceptual model under KARSYS is applied to the case study of Pasuaran, East Java, Indonesia. The study area is a large one of about 560 km<sup>2</sup> (20 km along EW and 28 km along NS) and covers the northern flank of the Bromo-Tengger volcano, the Monte Bromo being the highest point (2470 m). It encompasses the highest caldera, Tengger caldera, and the coastal plain. The high discharges of groundwater flow, both springs and from free-flowing wells in the plain, are located on this side of the volcano, and therefore of high hydrogeological interest as well for water resources management. The 3D geological model has been built

according to information and data found in the following documents [Abidin *et al.*, 2004, Toulhier *et al.*, 2019, Tugas Akhir, 2019].

From a geological point of view, the Bromo-Tengger mountain is a complex andesitic stratovolcano. It began to grow about 1.4 million years ago. Over time, numerous eruptions produced sequences of lava flows accumulation and subordinate pyroclastic deposits, all of them constituting the framework of this composite volcano [Thouret, 1999 in Toulhier *et al.*, 2019]. The volcano is incised by steep slopes canyons on the flank due to erosion. The Pasuaran north plain, as a peripheral zone of the volcano, is constituted by quaternary volcano-sedimentary deposits. Their thickness is more than 300 m as reported by [Mazzini *et al.*, 2018]. The volcano-sedimentary series comprise a large variety of typical facies, from proximal (lava flows, pyroclastic falls and flows, lahars) to distal sequences (volcano-sedimentary deposits with higher content of clay). A fine clayey sediments layer of about 10 m thickness forms the top of these series in the northern plain [Toulhier *et al.*, 2019].

From a hydrogeological point of view, according to [Toulhier *et al.*, 2019], the volcano system is composed of 6 hydrogeological units, from the caldera to the volcano-sedimentary plain: (i) the central zone forms a sand sea caldera, with dark sand of high

porosity, probably hosting an unconfined aquifer (ii) the proximal zone of the edifice with several hundred meters of pyroclastics, overlying lava flow complex, with high porosity and permeability, (iii) the median zone with non-weathered lava flow complex, a huge unit that can be considered as homogeneous as a whole, with few low discharges at the inter lava flows locations (perched and local aquifers), (iv) distal zone with slope deposits characterized by a low permeability due to their clayey content, and (v) the volcano-sedimentary plain, with a confined multi-layers aquifer, as the main hydrogeological feature providing groundwater at the artesian springs (Umbulan spring namely, the largest with a discharge of  $3250 \text{ L}\cdot\text{s}^{-1}$  and Banyu Biru,  $200 \text{ L}\cdot\text{s}^{-1}$ ) and free flowing wells. [Toulier *et al.*, 2019] inventoried 460 artesian wells in the plain, with an average discharge of  $4.1 \text{ L}\cdot\text{s}^{-1}$  each. Considering the density of wells ( $\sim 6 \text{ wells km}^{-2}$ ) and the area of the plain, it was estimated that 20% of wells were out of the survey. A maximum of 600 wells is then estimated, with a total of  $2400 \text{ L}\cdot\text{s}^{-1}$ . Hydroclimatological data from a monitoring network considering the altitudinal gradient are available in [Toulier, 2019].

There is a groundwater resource management issue in the Pasuruan Plain. The Umbulan spring discharge is decreasing. The monitoring of the spring since 2007 confirms the decreasing trend in discharge. If oral testimonies report a discharge of about  $5000 \text{ L}\cdot\text{s}^{-1}$  in the 1980s, the discharge in 2007 was around  $4000 \text{ L}\cdot\text{s}^{-1}$ . Nowadays it is about  $3250 \text{ L}\cdot\text{s}^{-1}$ . In 2021, the Umbulan spring was utilized for the water supply of Surabaya city located nearby, the second largest town in Java Island. An uncontrolled groundwater use through free-flowing wells in the plains for paddy fields—rice cultivation—reduces the groundwater level, and consequently impacts the lowland pressure-driven artesian springs, e.g., Umbulan spring [Khasanah *et al.*, 2021]. The pressure on the groundwater resource is getting higher, with potential tensions between water users. The decrease of the groundwater level already led to the increased need of pumps in the wells and social conflicts may arise due to the possible unaffordable cost of such equipment for some local farmers.

### 3.2. *Building of the KARSYS conceptual model*

Geological units have been simplified and grouped to improve the readability of the model (Table 2). The building principles follow the chronology of the main lava (and pyroclastic deposits) flow events and are based on the followings: (i) Each new lava flow is channeled into existing valleys and overflows from the topography into downstream areas (fans-like morphology), (ii) Pyroclastic flows are deposited in stratum around the places of emission, (iii) Between two distinct episodes of flows, deep valleys could potentially have been incised and (iv) Deep valleys do not incise below the sea level. The resulting geological and hydrogeological 3D models are presented in Figure 5.

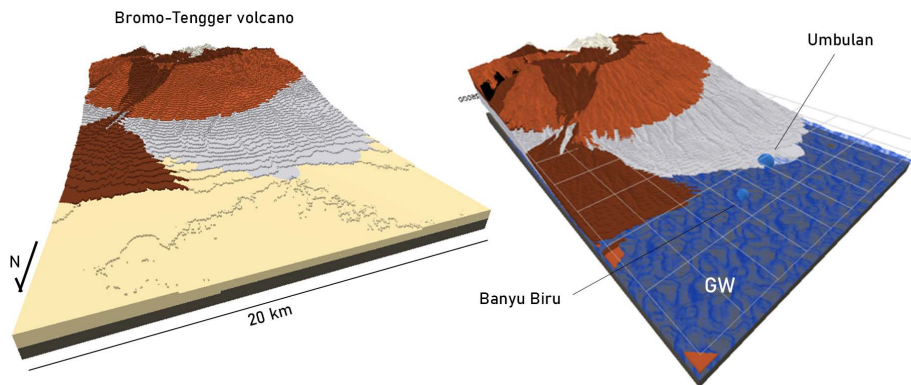
The KARSYS model shows that a large groundwater zone develops in the volcano-sedimentary plain and extends upstream in the different lava flow units. In the volcano-sedimentary plain, the groundwater zone is confined below successive layers of low pervious ashes. In the upstream part of the groundwater zone, most of the groundwater is supposed to be drained by the springs while in the downstream part it directly discharges into the sea.

### 3.3. *Scenarios visualization*

Two animation scenarios are set: (i) a short-event of rain and (ii) a long period of recharge over several decades in order to point out specific processes that take place for these contrasted timescales. The short-term animation figures out the dynamics of the surface processes (runoff and evapotranspiration) and of the groundwater in the mid to high permeable units (i.e.,  $K > 10^{-3} \text{ m/s}$ ) while the long-term animation makes it possible to show the dynamic of the groundwater in the low permeable units (i.e.,  $K < 10^{-6} \text{ m/s}$ ) and the progressive mineralization of the particles according to the time they spend in the different aquifers. Results are displayed as pictures but it should be kept in mind that the results are displayed in a 3D animated scene where water particles are computed live.

#### 3.3.1. *Short rain event*

A 30-day rainfall series with a 6-hour timestep (i.e., 120 values) has been created to reproduce a tropical rain event in wet season (Figure 6). Clouds (i.e., the



**Figure 5.** Left: Geological 3D model of the Pasuruan test site. Dark-brown formations refer to ashes and pyroclastic fall deposits, light-brown formations refer to pyroclastic deposits, grey formations refer to basalt/andesite lava flows, light-yellow formations refer to plain sediments. Right: Hydrogeological 3D model with visualization of springs (blue spheres) and groundwater bodies (“GW”) under the impervious overlying layer in the plain (the light-yellow layer is removed).

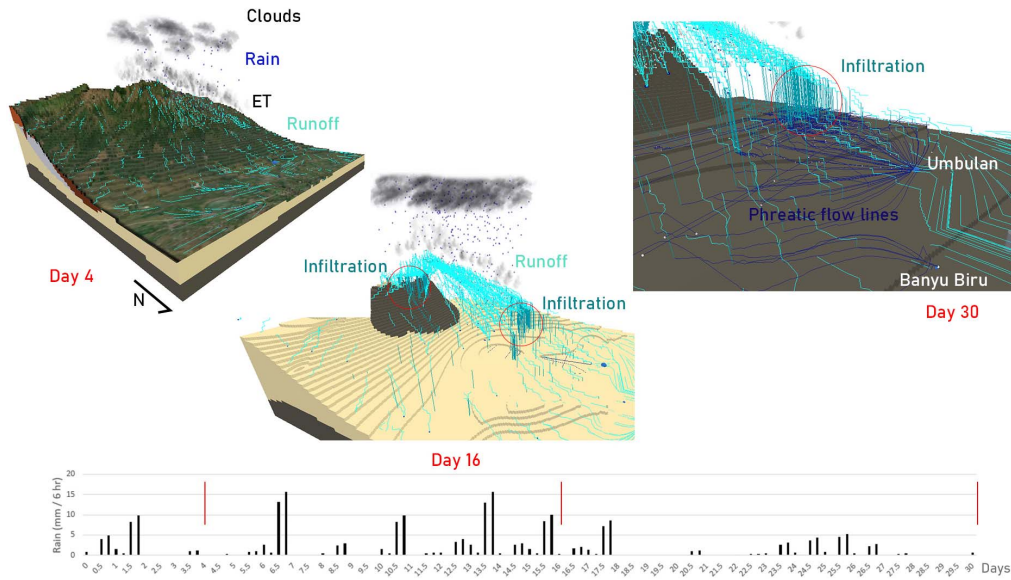
**Table 2.** Hydraulic characteristics of the 6 hydrogeological units of the Pasuruan model

Modeled geological units (Visual KARSYS)	Hydraulic conductivity (K)
Sand sea caldera	Highly permeable unit, $1 \times 10^{-2}$ m/s
Ashes and pyroclastic fan deposits	Low permeable unit, $1 \times 10^{-6}$ m/s
Pyroclastic deposits	Low permeable unit, $1 \times 10^{-5}$ m/s
Basalts, andesites	High permeable unit, $1 \times 10^{-3}$ m/s
Volcano sedimentary plain	High permeable unit, $1 \times 10^{-3}$ m/s
Old weathered volcanic deposits (Basement)	Non-permeable unit

place where the water is emitted as rain) are located in the upstream part of the area at an elevation of 4000 m asl above the upper flanks of the Bromo-Tengger volcano. In January, the cumulated rainfall is considered as 219 mm and ET as about 20% of the raw precipitation.

At day 4, after a first rain event, most of the particles flow as runoff over the top of the pyroclastic units with moderate permeability. The infiltration is limited. At day 16, we can see that most of the infiltration takes place within the ravine beds at the point of contact with the more permeable basalts ( $K = 10^{-3}$  m/s) or in the caldera which is composed of permeable

sands ( $K > 10^{-2}$  m/s). At day 30, we can see that the infiltration still occurs in the basalts and no significant amount of water flow through the pyroclastic units. The water reaches the saturated zone and forms phreatic flow-lines that converge in the direction of the springs. As expected because of the location of the cloud, most of the phreatic flow lines go to Umbulan which displays a higher regime than Banyu Biru (about a factor 10 which is reliable with the proportion of displayed lines). At the end of the animation, the computed ET is about 67 mm, the infiltrated water 36 mm and the runoff 116 mm. Because of the relative low permeability of the pyroclastics, 50% of



**Figure 6.** Three selected stages (day 4, 16 and 30) of the 30-day event animation applied to the Pasuruan test site.

the water is drained as runoff, the infiltration being about 16% in this particular case.

### 3.3.2. Long term

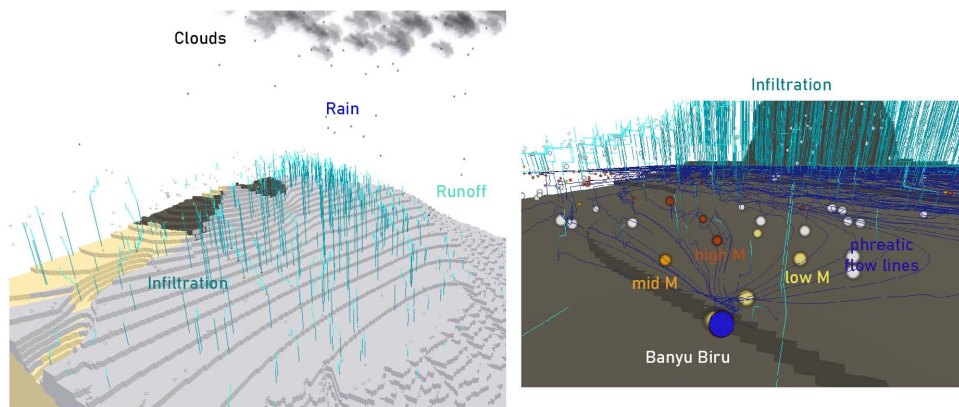
A 240-year series with a 2-year timestep has been created to reproduce the groundwater flow over the long term, especially in the low permeable units (pyroclastics and fan deposits). The cumulated rainfall over 2 years is about 2580 mm (i.e., 1290 mm/year) and the clouds extend over the whole area. Overviews of the animation results are given in Figure 7.

After a few years, we can see the infiltration through the low permeable pyroclastic units which was not visible for the short-term animation. In the groundwater zone, the color of the particles allows to distinguish the low-mineralized water (i.e., the “freshwater”), the mid-mineralized one and the highly mineralized one as well as the mixing processes taking place upstream of the springs, resulting in mixed mineralized water at the springs. High-mineralized water refers to particles which have a longer transit time in the aquifer (at least >2/3 of the total duration of the animation). This applied, for instance, to particles which first infiltrate through the low-permeable ashes and pyroclastic fan deposits ( $K = 10^{-6}$  m/s) and finally reaches the basalts

where they mix with freshly infiltrated particles. At the end of the animation, the cumulated rainfall is about, 310,000 mm, the ET 129,000 mm (41%), the runoff and infiltration 181,000 mm (59%).

## 4. Discussion

Different questions may arise by reading the previous sections. First of all, we can think about the usefulness of such a tool, including FAM. This could be seen as a “nice to have” or a gadget in the sense that this type of dynamic 3D conceptual model remains a fairly qualitative model (even if it can provide pseudo-quantitative results) which is usually based on a few data and a series of assumptions and principles. This is exactly the expected goal of Visual KARSYS and it does not intend to replace numerical models which provide more quantitative and physically based results. The spirit of such a tool is to really offer the possibility for a modeler (geologist, hydrogeologist) to explicitly describe the hydrosystem as all parameters and information are modeled (i) in 3D and (ii) over a period of time and to present preliminary interpretations regarding main characteristics and processes (groundwater resources, flow, recharge, etc.) even to non-specialists. The development of the FAM has raised many questions about



**Figure 7.** Overview of a long-term event animation applied to the Pasuruan test site. “low M”, “mid M” and “high M” respectively refer to low-mineralized, mid-mineralized and high-mineralized particles.

how to represent explicitly, the hydrogeological processes, especially with timescales (short recharge events, seasonal variations, long-term trends) and permeability issues, which make it virtually impossible to represent the dynamics of these processes in a single model. This implies that if all these multi-scale processes cannot be visualized altogether, they are difficult to communicate—and even impossible without dynamic conceptual models to address these issues. The FAM offers the possibility to get an idea of the main characteristics and process dynamics of the studied hydrosystem within a few minutes and requires a limited effort to be built compared to numerical models which may require weeks for setup and calibration. When working on site with limited data and information, such an approach may be more efficient than building a sophisticated numerical model which cannot be properly parameterized.

The second question may be related to the reliability of this type of explicit conceptual model. Indeed, the implementation and the calibration of these models are usually based on (i) a few “hard” data (observed geological outcrops, drillholes, caves and mines survey, etc.), (ii) a few “already interpreted data” as geological maps, cross-sections, interpretation of geophysical data, etc., (iii) a series of geological and hydraulic principles (lateral continuity, thickness conservation, gradients, etc.) and (iv) the knowledge or the experience of the modeler to describe the “reality” of the invisible underground environment. In this context, the tool is usually not the limiting factor for interpretation, but rather the available data and the quality of the information on the

site of interest. Introducing and combining as much data and information as available in an explicit conceptual model also allows for direct discussion of the density and reliability of the data and thus the uncertainties of the model, which is less often possible with numerical models.

## 5. Conclusion

This paper intends to introduce the usefulness of a dynamic modeling and visualization tool (Visual KARSYS and the FAM) which allow modelers (geoscientists) to build specific conceptual models of the studied hydrosystem (i) in 3D to show the main characteristics and (ii) in time to visualize the hydrogeological processes taking place in the aquifer. This tool is developed for the modelers to specifically describe the hydrosystem on which they work prior to design and implement numerical models in which most of the site characteristics and processes might remain implied. It is also useful for non-specialists (stakeholders, water authorities, etc.) to communicate the main characteristics of the aquifer and hydrogeological processes or to answer pragmatic questions which remain qualitative (via an interactive 3D viewer, video-clips, etc.). Such models do not intend to replace numerical ones which provide more quantitative results. However, they can and should be used as a basis for interpretation and discussion before undertaking numerical modeling work.

We expect increased general use of detailed and specific conceptual models to support and improve the management of the groundwater resources in a

more integrated approach (tapping strategies, protection policies, conflicts of use, etc.). Past and current applications to different sites in different countries (Switzerland, France, Spain, Slovenia, etc.) demonstrate the interest of the end users in this type of representation and support for their groundwater management activities.

Visual KARSYS and the FAM are still work in progress and evolving. New developments are possible according to needs and funding. Any inputs for improvement are welcome too.

### Conflicts of interest

Authors have no conflict of interest to declare.

### Acknowledgements

The Visual KARSYS project is supported by the Swiss Federal Office for Environment via two grants for the promotion of environmental technology: UTF 537.13.16 (2016–2019) and UTF 647.26.20 (2020–2021). The flow animation module has been supported by Danone Waters, in the frame of a contract with ISSKA in 2021.

The authors thank the three reviewers for their relevant comments and suggestions that allow to improve the general quality of the paper.

More information about Visual KARSYS on Twitter: @Visual\_KARSYS

### References

- Abidin, H. Z., Andreas, H., Gamal, M., Hendrasto, M., Suganda, O., Purbawinata, M., Meilano, I., and Kimata, F. (2004). The deformation of bromo volcano (Indonesia) as detected by GPS surveys method. *J. Glob. Position. Syst.*, 3, 16–24.
- Anderson, M. P. and Woessner, W. W. (1992). *Applied Groundwater Modeling: Simulation of Flow and Advective Transport*. Academic press, New York.
- Baccar, M., Bergez, J.-E., Couture, S., Sekhar, M., Ruiz, L., and Leenhardt, D. (2021). Building climate change adaptation scenarios with stakeholders for water management: a hybrid approach adapted to the South Indian water crisis. *Sustainability*, 13, article no. 8459.
- Ballesteros, D., Malard, A., Jeannin, P. Y., Jiménez-Sánchez, M., García-Sanseguendo, J., Meléndez-Asensio, M., and Sendra, G. (2015). KARSYS hydrogeological 3D modeling of alpine karst aquifers developed into geological complex areas: Picos de Europa National Park (Spain). *Environ. Earth Sci.*, 74, 7699–7714.
- Betancur, V. T., Palacio, T. C. A., and Escobar, M. J. F. (2012). Conceptual models in hydrogeology, methodology and results. In Kazemi Dr, G. A., editor, *Hydrogeology - A Global Perspective*, pages 203–222. IntechOpen, London.
- Bierkens, M. F. P. and Wada, Y. (2019). Non-renewable groundwater use and groundwater depletion: a review. *Environ. Res. Lett.*, 14, article no. 063002.
- Cooper, R. J., Hiscock, K. M., Lovett, A. A., Dugdale, S. J., Sunnenberg, G., Garrard, N. L., Outram, F. N., Hama-Aziz, Z. Q., Noble, L., and Lewis, M. A. (2018). Application of high-resolution telemetered sensor technology to develop conceptual models of catchment hydrogeological processes. *J. Hydrology*, 1, article no. 100007.
- De Marsily, G. (1986). *Quantitative Hydrogeology*. Academic Press, London.
- De Marsily, G. and Besbes, M. (2017). Les Eaux souterraines. *Annales Des Mines - Responsabilité et Environnement*, 86(2), 25–30.
- Etienne, M., Du Toit, D., and Pollard, S. (2011). ARDI: a co-construction method for participatory modeling in natural resources management. *Ecol. Soc.*, 16, article no. 44.
- Ghasemizadeh, R., Hellweger, F., Butscher, C., Padilla, I., Vesper, D., Field, M., and Alshawabkeh, A. (2012). Review: Groundwater flow and transport modeling of karst aquifers, with particular reference to the North Coast Limestone aquifer system of Puerto Rico. *Hydrogeol. J.*, 20, 1441–1461.
- Jakeman, A. J., Barreteau, O., Hunt, R. J., Rinaudo, J.-D., and Ross, A., editors (2016). *Integrated Groundwater Management: Concepts, Approaches and Challenges*. Springer International Publishing, Cham.
- Jeannin, P. Y., Eichenberger, U., Sinreich, M., Vouillamoz, J., Malard, A., and Weber, E. (2013). KARSYS: a pragmatic approach to karst hydrogeological system conceptualisation. Assessment of groundwater reserves and resources in Switzerland. *Environ. Earth Sci.*, 69(3), 999–1013.
- Khasanah, N., Tanika, L., Pratama, L. D. Y., Lei-

- mona, B., Prasetyo, E., Marulani, F., Hendriatna, A., Zulkarnain, M. T., Toulrier, A., and van Noordwijk, M. (2021). Groundwater-extracting rice production in the Rejoso Watershed (Indonesia) reducing urban water availability: characterisation and intervention priorities. *Land*, 10, article no. 586.
- Lopez, S., Courrioux, G., Calcagno, P., Bourguine, B., and Allanic, C. (2018). Yet another geological modeling library.... In *Delivering Subsurface Models for Societal Challenges - 4th Meeting of the European 3D Geomodelling Community, 21st to 23rd February 2018, Orléans, France*, page 71. BRGM, Paris.
- Malard, A. (2018). *Hydrogeological Characterization of Karst Aquifers in Switzerland using a Pragmatic Approach*. University of Neuchâtel, Neuchâtel.
- Mazzini, A., Scholz, F., Svensen, H. H., Hensen, C., and Hadi, S. (2018). The geochemistry and origin of the hydrothermal water erupted at Lusi, Indonesia. *Mar. Pet. Geol.*, 90, 52–66.
- Reed, M. S., Graves, A., Dandy, N., Posthumus, H., Hubacek, K., Morris, J., Prell, C., Quinn, C. H., and Stringer, L. C. (2009). Who's in and why? A typology of stakeholder analysis methods for natural resource management. *J. Environ. Manag.*, 90, 1933–1949.
- Rouillard, J., Neverre, N., and Rinaudo, J. D. (2022). Initiating collective action for the management of deep confined aquifer systems: application of a participatory scenario approach in France. *Hydrogeol. J.*, 30, 21–36.
- Shalsi, S., Ordens, C. M., Curtis, A., and Simmons, C. T. (2019). Can collective action address the “tragedy of the commons” in groundwater management? Insights from an Australian case study. *Hydrogeol. J.*, 27, 2471–2483.
- Thouret, J.-C. (1999). Volcanic geomorphology—an overview. *Earth Sci. Rev.*, 47, 95–131.
- Toulrier, A. (2019). *Multidisciplinary Study for the Characterization of Volcanic Aquifers Hydrogeological Functioning. Case of Bromo-Tengger Volcano (East Java, Indonesia)*. Montpellier University, France.
- Toulrier, A., Baud, B., de Montety, V., Lachassagne, P., Leonardi, V., Pistre, S., Dautria, J.-M., Hendrayana, H., Miftakhul Fajar, M. H., Satrya Muhammad, A., Beon, O., and Jourde, H. (2019). Multidisciplinary study with quantitative analysis of isotopic data for the assessment of recharge and functioning of volcanic aquifers: Case of Bromo-Tengger volcano, Indonesia. *J. Hydrol. Reg. Stud.*, 26, article no. 100634.
- Tugas Akhir, A. (2019). *Vulkanostratigrafi dan petrogenesis gunung Bromo dan Sekitarnya, kabupaten probolinggo, provinsi Jawa Timur*. Institut Teknologi Bandung, Bandung.
- Van der Gun, J. (2012). Groundwater and global change: trends, opportunities and challenges. In *UNESCO World Water Assessment Programme*. UNESCO, Paris. (No. ISBN: 978-92-3-001049-2).
- Zaar, U., George, D. C., Matsuyama, A., and Haig, T. (2002). Aboriginal views on “Balancing the Budget”. In *CD Proceedings: International Groundwater Conference Darwin 2002 – “Balancing the Budget”*. Northern Territory Administration, Darwin.





Research article

Geo-hydrological Data & Models

# Hydrodynamic relationships between gravel pit lakes and aquifers: brief review and insights from numerical investigations

Anne Jost<sup>ⓧ,\*</sup>, Shuaitao Wang<sup>ⓧ,a</sup>, Thomas Verbeke<sup>b</sup>, François Colleoni<sup>c</sup> and Nicolas Flipo<sup>ⓧ,d</sup>

<sup>a</sup> Sorbonne Université, CNRS, EPHE, Metis, F-75005 Paris, France  
URL: <https://www.metis.upmc.fr/~jost/index.html>

<sup>b</sup> Ecole Normale Supérieure Paris-Saclay, Laboratoire de Mécanique Paris-Saclay, Gif-sur-Yvette, France

<sup>c</sup> INRAE-AMU, Aix-en-Provence, France

<sup>d</sup> Geosciences Department, Mines Paris, PSL University, Paris, France  
E-mails: [anne.jost@sorbonne-universite.fr](mailto:anne.jost@sorbonne-universite.fr) (A. Jost), [shuaitao.wang@upmc.fr](mailto:shuaitao.wang@upmc.fr) (S. Wang), [thomas.verbeke@ens-paris-saclay.fr](mailto:thomas.verbeke@ens-paris-saclay.fr) (T. Verbeke), [francois.colleoni@inrae.fr](mailto:francois.colleoni@inrae.fr) (F. Colleoni), [nicolas.flipo@minesparis.psl.eu](mailto:nicolas.flipo@minesparis.psl.eu) (N. Flipo)

**Abstract.** Supply of aggregate materials for every construction requires mining of sand and gravel, which leads to the formation of a myriad of freshwater lakes, a now common feature of the landscape in the valleys of large rivers. Typically small in size and shallow, they are filled with waters from the adjacent aquifers and directly exposed to the atmosphere. The creation of gravel pit lakes has various and contrasting effects on their immediate environment. This article first provides a review of these impacts from the hydrodynamic point of view, and illustrates them on simple numerical test cases. It also introduces the gravel pit lake module developed for the occasion within the integrated modelling platform CAWAQS, which formulation was tested on the same test cases against the LAK package, its MODFLOW counterpart. By accurately simulating gravel pit lake interactions with groundwater in different configurations, this modelling exercise also aims to identify the preponderant factors leading water level fluctuations of those artificial lakes, whose temporal monitoring will soon be accessible to satellite observation.

**Keywords.** Gravel pit lake, Groundwater–lake interaction, Numerical modelling, In silico experiments, Natural resources.

## 1. Introduction Manuscript received 23 May 2022, accepted 15 November 2022

The low-carbon transition requires an increasing amount of raw materials, from the most traditional, such as aggregates, to the most emerging and rare metals, raising the question of their availability [de Marsily and Tardieu, 2018]. In particular, sand

and gravel will be needed for large-scale infrastructure development such as for wind turbine foundations and structures or for insulation materials. Although sand is one of the most abundant materials on Earth [Sverdrup et al., 2017], its exploitation faces major environmental issues, associated economical difficulties and even regional scarcity concerns [Ioannidou et al., 2020], in particular when it comes to its quality requirements. Population growth, rapid

\* Corresponding author.

urbanisation and infrastructure development have led to increasing demand since the 1950s [OECD, 2019]. Mostly used in the construction industry as a key ingredient in the production of concrete, for road bases and land reclamation, sand and gravel are the world's most consumed primary materials after water, up to 50 billion tonnes produced each year [Bendixen *et al.*, 2021], at unsustainable levels far greater than their natural renewal [Bendixen *et al.*, 2019, Peduzzi, 2014].

Sand and gravel resources are derived from various geomorphological settings such as beach deposits, streambeds, river floodplains and terraces, alluvial fans, and glacial deposits. Extraction of material from terraces and upland areas is generally perceived as having less impact than removing sand and gravel from active floodplains and stream channels [Kondolf, 1997, Sandercock and Ladson, 2014]. However, when aggregates are mined below the water table, artificial water bodies fed by groundwater appear as a new landscape feature. These thousands of gravel pit lakes are now a common freshwater lake type significantly influencing the morphology of the watershed, the natural hydrologic system and regional biogeochemical cycles [e.g., Mollema and Antonellini, 2016]. In these environments, surface water and groundwater will mix and interact with the atmosphere. By offering open water surfaces where direct evaporation can occur, sustained by groundwater inflow, gravel pit lakes are generally recognised as a sink for adjacent aquifers in temperate and Mediterranean climates [Mollema and Antonellini, 2016], particularly in dry years [Schanen, 1998], although they may also act as a temporary buffer reservoir [Sinoquet, 1987], especially during low-amplitude floods [Czernichowski-Lauriol, 1998]. This freshwater loss is of concern, in areas where the mining of sand and gravel from those productive reservoirs is already in competing use with drinking water supply.

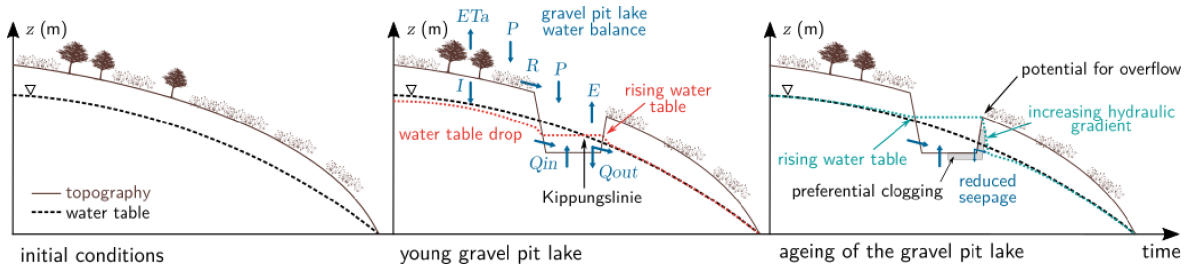
Furthermore, gravel pit lakes, characterised by infinite transmissivity and a unit storage coefficient, also alter the hydraulic gradient in the adjacent aquifer, causing the water table to rise or fall and thus disturbing the groundwater drainage pattern [Peaudecerf, 1975]. By establishing a surface of equipotential head, gravel pit lake levels may nonetheless be representative of the average groundwater level, like giant piezometers. With the develop-

ment of ever more efficient remote sensing systems, satellite observation will soon provide regular monitoring of temporal fluctuations of open continental water surfaces with unprecedented precision. Despite their small size, gravel pit lakes are a good candidate for future monitoring by the SWOT (Surface Water and Ocean Topography) satellite [Ottlé *et al.*, 2020]. For landscapes where few *in situ* groundwater level measurements are available, gravel pit lakes could be used as proxy indicators of local water resource trends. This will require the use of a modelling tool for the coupled gravel pit-aquifer system.

A lot of effort has recently been put into analysing lake-aquifer interactions but future work still needs to assess the potential of using the lakes, and in particular the increasingly common gravel pit lakes, as monitoring wells of shallow groundwater for better water resources planning and management [Shrestha *et al.*, 2021]. Understanding how gravel pit lake level will fluctuate in response to groundwater exchange, overland flow and atmospheric conditions (precipitation and evaporation) is therefore fundamental for this purpose, especially as artificial lakes interact differently with groundwater compared to natural lakes [El-Zehairy *et al.*, 2018]. Special emphasis in this paper is on the dynamics of these interactions. By first reviewing the results of field and numerical studies, we recall which consequences gravel mining may have on groundwater systems from a quantitative point of view. We then present the numerical code we have developed in order to simulate the gravel pit lake/aquifer interaction and introduce the test case used to validate the lake module. On the basis of the same test case, we finally illustrate numerically the hydrodynamic effects associated with gravel pit lakes and also discuss the main factors influencing lake level changes.

## 2. A brief review of the hydrodynamical impacts of gravel pit lakes

The interactions between gravel pit lakes and their environment have received attention for many years and have recently been summarised by Mollema and Antonellini [2016]. This additional state of knowledge report is meant to be an overview of relevant information on the hydrodynamic aspects associated



**Figure 1.** Illustration of gravel pit lake–aquifer interactions over time. The terms of the water balance are: precipitation  $P$ , land actual evapotranspiration  $ET_a$ , open-water evaporation  $E$ , diffuse runoff  $R$ , infiltration  $I$ , groundwater inflow  $Q_{in}$ , groundwater outflow  $Q_{out}$ .

with the presence of gravel pit lakes. The main elements discussed in this section are summarised in Figure 1.

### 2.1. Characteristics of gravel pit lakes

By definition, a gravel pit lake develops where sand and gravel extraction extends below the water table. Gravel excavations are therefore typically found in areas with a shallow water table, in floodplains and adjoining terraces of large rivers, in glacial valleys or coastal areas, where most coarse-grained sediments were deposited. Scattered throughout these landscapes, pit lakes are permanent water bodies of recent formation, mostly dredged since the second half of the last century. Usually located near urban areas, where aggregates are needed, water-filled pits are used for a variety of purposes, as reservoirs for water supply and irrigation, for recreational activities, and wildlife habitat. As such, they play an important ecological role [Seelen *et al.*, 2021].

The gravel pit pools are usually small, with a surface area varying from a few hundred square metres to several hectares. Many of them have steep sides, uneven but relatively flat bottoms and irregular shorelines [e.g., Hindák and Hindáková, 2003, Kondolf, 1997]. Their maximum depth depends on the thickness of the gravel layers, which is usually limited, resulting in generally shallow lakes but they tend to be deeper than natural lakes [Vucic *et al.*, 2019] and some can reach a depth of several tens of metres [Mollema and Antonellini, 2016].

Gravel pit lakes are often in close proximity to each other, with large areas of the plains turning into open pits, up to nearly 25% [Peckenham *et al.*, 2009], but

they are usually disconnected from permanent watercourses, except during exceptional floods [Kondolf, 1997]. They therefore generally have no natural surface inlet and outlet. When they are located in relatively flat floodplains, diffuse runoff is also limited. In contrast, they are in close hydrologic continuity with the surrounding groundwater body [Søndergaard *et al.*, 2018], as highly permeable sand and gravel deposits allow significant groundwater seepage. Groundwater inflow is thus considered a key component in the water balance of quarry lakes. This inflow returns either downstream to the aquifer or to the atmosphere, making most of these artificial lakes so called flow-through or seepage lakes [Wilson, 1984].

### 2.2. Hydrodynamic effects associated with gravel pit lakes

Gravel pits are young objects in the landscape whose environmental effects have been observed *in situ* for about 50 years. A number of studies on their hydrodynamical impacts have been published since the pioneering work of Peaudecerf [1975] and Vandenberg [1975] in France, Wrobel [1980] in Germany or Wilson [1984] and Morgan-Jones *et al.* [1984] in Great Britain. Most of them have been carried out in temperate and western countries where sand and gravel mining was historically practised [Koehnken and Rintoul, 2018] and are summarised in Table 1.

A major concern in the development of aggregate mines is to avoid or minimise negative effects on local water resources. Peaudecerf [1975] was among the first to consider their impacts on groundwater and to question the compatibility of simultaneous

**Table 1.** Summary of the scientific literature on the hydrodynamic impacts of gravel pit lakes, by site and in chronological order

Reference	Site	Deposit type
Vandenbeusch [1975]		
Saplaïroles et al. [2007]	Garonne River, France	Alluvial
Bessière et al. [2013]		
Morgan-Jones et al. [1984]	River Colne, England	Alluvial
Wilson [1984]	Rivers Thames and Avon, England	Gravel terraces
Durbec [1986]		
Sinoquet [1987]	Rhine River, France	Alluvial
Marsland and Hall [1989]	Southern coastline, England	Beach
Mazenc et al. [1990]	Oise River, France	Alluvial
Panel [1991]	Marne River, France	Alluvial
Blanchard et al. [1991]		
Mimoun [2004]	Loire River, France	Alluvial
Hatva [1994]	Finland	Glaciofluvial
Mead [1995]	Thurston county, Washington, USA	Unconsolidated deposits of glacial and nonglacial origin
Schanen [1998]	Seine River, France	Alluvial
Kattner et al. [2000]	Danube River, Germany	Alluvial
Michalek [2001]	Columbia River, Oregon, USA	Floodplain
Green et al. [2005]	Minnesota, USA	Unconsolidated deposits of glacial and nonglacial origin
Kuchovský et al. [2008]	Morava River, Czech Republic	Alluvial
Peckenham et al. [2009]	Hancock county, Maine, USA	Unconsolidated deposits of glacial and nonglacial origin
Smerdon et al. [2012]	Boreal Plains, Canada	Glacial outwash plains
Apaydın [2012]	Kazan Plain, Turkey	Fluvial
Mollema and Antonellini [2016]	River Meuse, the Netherlands	Alluvial
Mollema and Antonellini [2016]	Adriatic coast, Italy	Coastal areas

water abstraction and gravel mining. First, gravel extraction produces an area of high permeability within the aquifer, which alters the direction of groundwater, either by providing preferred groundwater flow paths towards the lake or conversely, as an obstacle to groundwater flow due to clogging of the pit bed and banks. Second, what was previously the water table becomes an horizontal lake surface. As a result, groundwater levels immediately adjacent to the open pit must fall upgradient and increase at its downgradient end (Figure 1). Wrobel [1980] postulates that the initial level of the gravel pit lake should theo-

retically stabilise at the height of the pre-extraction water table that existed approximately halfway between the upstream and downstream ends of the open pit. He calls the line where the pre-extraction water table intersects the surface of the lake the “Kippungslinie”. It divides the lake into an upstream section where groundwater inflow is prevalent, and a downstream area where outflow takes place. Any further shift in the position of this line would thus determine the presence and extent of lake sealing [Wilson, 1984]. The effects of existing mines on groundwater have been confirmed by field observations of local

changes in groundwater levels [Morgan-Jones *et al.*, 1984], of up to a few tens of centimetres in the immediate perimeter of the gravel pits and measurable over a maximum of several hundred metres around the gravel pit [e.g., Bessi re *et al.*, 2013, Gravost, 1988, Sinoquet, 1987]. Assessing these changes quantitatively, however, is often difficult due to the lack of site-specific data on water levels prior to gravel mining.

Changes in groundwater levels can affect the water supply of nearby wetlands [Green *et al.*, 2005], streams [Kuchovsk y *et al.*, 2008] or other surface water features [Smerdon *et al.*, 2012], either adversely upstream of the lake or favourably where downgradient increases in water levels create opportunities for wetland enhancement [Maliva *et al.*, 2010]. Of major concern and source of land-use conflicts is the potential impact on groundwater supply, as the deposits that contain significant aggregate resources also host valuable unconfined aquifers [e.g., Apaydın, 2012, Nadeau *et al.*, 2015]. Marsland and Hall [1989] recall in this respect the dispute between the “water” and “gravel” interests during the 1970s in Kent County, England. The former argued that the water resource available for abstraction would decrease due to evaporative losses from open water surfaces. The latter replied that the additional storage provided by lakes mitigated the drawdowns of the neighbouring production wells and that it was water withdrawal that was the cause of the decline in water levels. Marsland and Hall [1989] eventually conclude that the average groundwater level had fallen due to both gravel extraction and groundwater abstraction, making this coastal aquifer more vulnerable to saline intrusion.

It is still necessary to clarify in which direction aggregates extraction affects the water balance components (Figure 1). When examined on a monthly basis, precipitation generally exceeds open-water evaporation in winter and the gravel pit lakes recharge the aquifer; however evaporation exceeds precipitation in other months of the year, resulting in a loss to the aquifer [Wilson, 1984].

On one hand, the removal of vegetated, low-permeability soil layers in the excavation area can promote enhanced recharge and a higher rate of water cycling [Smerdon *et al.*, 2012], provided that the lakes capture available excess precipitation and spring snowmelt [Hatva, 1994], river water infiltra-

tion [Kattner *et al.*, 2000] or local surface water runoff [Apaydın, 2012]. However, the topography is often flat in the vicinity of gravel pits and they are also designed to be protected from surface water intrusion which could cause pollution in the aquifer [Saplaioles *et al.*, 2007].

Furthermore, removing the gravel itself increases the storage capacity, which can act to maintain locally higher aquifer water levels. Particularly during floods or after a rainfall event, gravel pit lakes may temporarily act as a buffer reservoir to dampen groundwater level fluctuations, depending mainly on the hydraulic conductivity of the aquifer [Sinoquet, 1987]. However, the storage volume available at the time of overflow may be negligible so that this effect is only observed for small overflows in spring and summer, but is cancelled out for large floods [Mazenc *et al.*, 1990]. Gravel mines can instead facilitate the transfer of water [Mazenc *et al.*, 1990] or even be captured by the active channel in case of flooding when they are located close to a watercourse [Kondolf, 1997].

On the other hand, the gravel pits create windows through the unsaturated soil layer into the aquifer, which is directly exposed to the atmosphere and thus, to increased water losses through evaporation, which can lead to a water balance deficit, especially during dry years [Saplaioles *et al.*, 2007, Schanen, 1998]. Two processes operate that are difficult to quantify, direct evaporation from the open water surface and transpiration by emergent plants, depending on meteorological factors and local conditions such as the depth of the water body, the presence of riparian trees along the shoreline [Hayashi and van der Kamp, 2021] or the connection with the surrounding groundwater body that replenishes the evaporated water. In most cases, evaporation from the artificial pond should be larger than actual land evapotranspiration notably in temperate and Mediterranean climates [Mollema and Antonellini, 2016], even though it can be substantially smaller than potential evapotranspiration [Hayashi and van der Kamp, 2021]. Evaporation losses from gravel pit lakes have been estimated at an average of 6 to 11 m<sup>3</sup> per day and per hectare of pits [Panel, 1991, Schanen, 1998] but varies according to the hydrological year. In addition to the amount of water that would have infiltrated in the absence of the quarries, it may represent a significant share of the renew-

able resource, especially as compared to groundwater withdrawals, and given the relatively small area occupied by the water bodies [Bessière *et al.*, 2013]. If the lakes are concentrated in close proximity to each other, the cumulative effect of their large number may be sufficient to cause a measurable drop in the water table [Marsland and Hall, 1989, Wilson, 1984].

In conclusion, the increase in open water resulting from gravel extraction surely affects the water balance of their catchment. The processes involved are, however, often difficult to quantify, vary from one year to another and depend on the regional context, including previous land use prior to the land-water conversion. For example, one should also consider former groundwater-consuming activities such as irrigation in areas replaced by gravel pits [Maliva *et al.*, 2010]. One of the major hydrological impacts of the flooded gravel pits is their significant open water evaporation, which is expected to increase further under climate change [Mollema and Antonellini, 2016], as does the global evaporative water loss, especially from artificial lakes [Zhan *et al.*, 2019, Zhao *et al.*, 2022].

### 2.3. Factors that matter

The aforementioned impacts of gravel pits on groundwater flow patterns depend on several key factors: (i) the extent of clogging of the sides and the bottom of the lakes, (ii) the geometry of the excavations, (iii) their position with respect to the general direction of groundwater flow, and (iv) the characteristics of the aquifer itself, which influence the hydraulic gradient.

Clogging occurs as the result of a series of phenomena leading to a decrease in the permeability of the solid matrix at the interface between surface water and groundwater. Among the mechanisms involved in clogging, the most important is the sedimentation of fine particles in suspension, the origin of which lies partly in the extraction and processing phase itself, when silt and clay tend to be washed out of the gravel as it is excavated [Wilson, 1984]. The partial filling of the excavations by the low-permeability overburden in the restoration phase is a further source of fine sediment and another clogging factor [Vandenbeusch, 1975]. To a lesser extent,

chemical and biological processes are also responsible for progressive clogging over time. Indeed various geochemical reactions of oxidation–reduction, precipitation/dissolution or dissolved complex formation take place in the lake water when it mixes with groundwater and can lead to clogging of the lake boundary [Wilson, 1984]. The development of algae, bacterial flora or rooted aquatic vegetation in summer and the deposition of organic matter in winter will also cause biological clogging [Blanchard *et al.*, 1991].

The amount of clogging in gravel pit lake varies significantly depending on a number of factors such as the morphology of the pit and in particular the slope of the banks, the mining method and subsequent reclamation, the current use of the lake, the presence of vegetation on the banks, or the turbidity of the lake water. According to field observations, clogging is not evenly distributed along the banks (Figure 1): it is usually predominant on the downstream banks of the gravel pits, in the direction of flow [Vandenbeusch, 1975], and preferentially occurs on their lower fringe [Blanchard *et al.*, 1991], whereas there is little or no clogging when the slope is greater than 20% [Eberentz and Rinck, 1987, Gravost, 1988]. The lake bottom is also subject to long-lasting clogging, particularly due to the collapse of the steep slopes [Zhang *et al.*, 2019], but otherwise does not vary significantly throughout the pit. Clogging starts from the first stages of mining and is generally established within a few years from the cessation of excavation [Muellegger *et al.*, 2013, Vandenbeusch, 1975]. It increases over time but with less intensity [Eberentz and Rinck, 1987], at a variable rate of evolution that depends mainly on the quality of the lake water [Wilson, 1984], and eventually becomes imperceptible to field measurements [Darmendrail, 1986]. Observations made *in situ* show clogged layers of varying thickness, from 0.5 to 1.5 m, and hydraulic conductivity between  $10^{-8}$  and  $10^{-3}$  m/s [Durbec, 1986, Schanen *et al.*, 1998].

The level of the lake, compared to the average pre-existing water table in the gravel pit area, is a good indicator of the presence and extent of lake sealing [Wilson, 1984]. Likewise, variations in the level of the gravel pit lake that are not synchronised with those of the water table and of smaller amplitude characterise the buffering role played by partially clogged reservoirs [Sinoquet, 1987]. Indeed,

sealing of the downstream boundary of gravel lake influences the long-term lake and groundwater levels, by raising the water level in the lake, as well as in the aquifer up-gradient from the lake, while downstream, the water table is lowered and the hydraulic gradient is locally increased [Peaudecerf, 1975, Vandenbeusch, 1975] (Figure 1). It is sufficient to maintain lake–aquifer exchanges [Sinoquet, 1987] but the low-permeability gravel lake sediments notably reduce the rate of groundwater seepage through the banks [Schanen, 1998, Wilson, 1984]. In the case of deep gravel pits and because clogging occurs primarily at depth, a low water table during a dry period will not favour groundwater seepage on the lower banks, whereas efficient exchanges between the lake and the aquifer are still possible when the water level can reach the upper unclogged fringes of the banks [Eberentz and Rinck, 1987, Mead, 1995].

The ageing of a water-filled gravel pit therefore results in a gradual slowing down of its exchanges with the adjacent aquifer. Several techniques are now available to estimate groundwater inflow and outflow in lakes and map their spatial distribution and temporal variability, using for example seepage meters, onshore and offshore geophysical measurements, or environmental tracers such as stable isotopes and temperature [Kidmose *et al.*, 2011, Masse-Dufresne *et al.*, 2021].

With regard to the influence of shape and size of the excavations on lake–groundwater interactions, Peaudecerf [1975] mentions that the disturbance of the original equipotential lines will be accentuated if gravel pit lakes are excavated in line parallel to the regional hydraulic gradient whereas elongated excavations, with their long axis perpendicular to the groundwater flow direction, will have relatively little effect on flow conditions. The creation of small water bodies rather than a large pond is also preferable to limit the risk of causing temporary overflows during high-flow periods [Mazenc *et al.*, 1990]. This is an important point to take into account when digging a gravel pit to avoid any potential overflow, should the raised lake level downstream exceed the topographic surface, especially when the latter is flat (Figure 1). Other lake parameters, such as the lake bed slope, have been identified as controlling the amount and spatial distribution of seepage [Genereux and Bandopadhyay, 2001], while excavation depth will have virtually no effect [Peaudecerf, 1975].

As for the position of the open water body within the regional flow system, it is obviously important for the gravel pit lake's water budget [Peaudecerf, 1975]: as numerically simulated by Cheng and Anderson [1994], groundwater inflow and outflow in lakes located lower in a watershed are likely to be higher and more important to the budget of lakes relative to precipitation than for uppermost lakes since lakes located in the discharge area intercept deeper groundwater than in the upper portion of the watershed where groundwater flows to and from the lakes originate from more variable shallow flow system.

Last but not least, it has long been known to what extent the hydrodynamic properties of the porous medium itself strongly influence the interaction of lakes and groundwater [Winter, 1976]. The magnitude of seepage is often governed by the regional groundwater conditions, *i.e.* hydraulic head gradient, anisotropy and heterogeneity of the aquifer which determine the background height of the water table relative to the lake level, especially on the downslope side of the lake.

#### 2.4. *Modelling gravel pit lakes*

The large volume of literature referred to above provides guidelines for sand and gravel mining operations, well known to professionals, which aim to ensure a balance between responsible economic development and mitigation strategies to protect local water resources and avoid overflows. Mining policies vary considerably between states and local jurisdictions but in most countries, sand mining is not only formally regulated by national mining legislation but must also comply with environmental legislation [Botta *et al.*, 2009]. Accordingly, an integrated environmental assessment, management and monitoring programme must be implemented in order to obtain permission to start operating a sand and gravel quarry. In this context, relatively simple groundwater modelling is now commonly used in planning aggregate excavation to make predictive and quantitative assessments.

The models make it possible to assist in the management of water resources [Fouché *et al.*, 2020] and examine the impact of various developments or restoration plans. The impact of gravel extraction on groundwater conditions is usually investigated by simulating three states [Kuchovský *et al.*, 2008, Panel,

1991]: pre-mining, with the presence of the existing open pits and integrating future gravel pits. In particular, the models enable the predictions of the cumulative effects of multiple extractions at the scale of the entire alluvial system [Bessière *et al.*, 2013] or to consider that all remaining alluvial resources are exploited [Mazenc *et al.*, 1990]. Field observations are used to adjust selected hydraulic parameters used in the hydrodynamic model such as the degree of clogging of the gravel pit banks [Durbec, 1986]. On theoretical case studies, sensitivity analyses are performed to assess the dominant parameters determining the response of the aquifer-gravel pit lake system.

Such numerical studies of lake-groundwater interactions are mainly conducted on natural lakes, from the early work of Winter [1976] who examined the general principles of these interactions to the more recent investigations of Jazayeri *et al.* [2021] who modelled the effects of lakes on groundwater wave propagation. They are also instructive with respect to artificial lakes when they aim to identify the main factors controlling lake-groundwater systems by varying lake and aquifer characteristics [e.g., Genereux and Bandopadhyay, 2001].

Different types of representation have been used to simulate the hydraulic effect of gravel pit lakes in groundwater flow models: (i) the simplest way is to specify the lake level as a constant head over the areal extent of the pit, assuming that it does not vary as a result of atmospheric exchanges or interactions with surface water and groundwater [e.g., Mimoun, 2004]; (ii) the famous and widely used “high K” technique, which proved to adequately simulate lakes [Winter, 1976], where they are considered to be a domain of very high hydraulic conductivity, of specific recharge and with a storage coefficient equal to 1 [e.g., Bessière *et al.*, 2013, Durbec, 1986, Michalek, 2001, Mimoun, 2004]; (iii) as the latter method is nevertheless subject to some numerical instabilities and faces difficulties in representing the seepage through the clogged boundaries of the lake, it may be necessary to consider more sophisticated lake modules [e.g., Smerdon *et al.*, 2012, Zhang *et al.*, 2019], such as the LAK package developed for MODFLOW [Merritt and Konikow, 2000] or the new SLM package [Lu *et al.*, 2021]. One of the particularities of lake packages is that they allow lake water levels to fluctuate in response to the lake-aquifer interaction, driven by the conductance of the interfaces.

### 3. Numerical modelling of groundwater-gravel pit lake exchanges

The next step was to provide the integrated modelling platform CAWAQS with such a tool capable of simulating the fluctuations in gravel pit lake water levels, in relation to their environment. We chose to develop a library dedicated to the hydrological simulation of gravel pits, given the modular architecture of CAWAQS and thus, on the basis of the one already available in MODFLOW [Harbaugh, 2005]. This module is based on the calculation of the water balance of the gravel pit, taking into account precipitation, evaporation, runoff and exchanges with adjacent aquifers. It has therefore been validated on a simplified alluvial plain case by comparing its performances with those of its counterpart and precursor, the LAK3 package [Merritt and Konikow, 2000].

#### 3.1. *Including a lake module in the CAWAQS hydrosystems modelling platform*

##### 3.1.1. *The CAWAQS platform*

CAWAQS (CAtchment WAter Quality Simulator) is a distributed and modular modelling platform for regional hydrosystems [Flipo, 2005, Labarthe, 2016]. This tool couples specific packages to simulate water transfers within and between the different reservoirs of the water cycle, from the surface to the underground compartment. The platform is conceptually divided into three components representing the surface, the unsaturated and saturated zones. Among its various libraries, it includes a package for the calculation of hydraulic heads in multi-layer aquifer systems, applying a semi-implicit finite volume scheme for the numerical resolution of the groundwater flow equation. We relied on the functionalities already existing in this library to develop a new module to simulate the interactions between the aquifer and a surface water body in order to estimate the hydrodynamic impacts of gravel pits.

##### 3.1.2. *Mathematical formulation of the gravel pit lake module*

The gravel pit lake module is designed to compute the lake level based on the volumetric exchanges of water into and out of the lake with the atmosphere, surface waters and adjacent aquifers, summarised



in the overall lake water balance. As described in Equation (1), the lake water level is controlled by the balance between the following terms of dimension ( $L^3 \cdot T^{-1}$ ): direct precipitation onto the lake  $P$ , diffuse runoff  $R$ , evaporation  $E$ , groundwater inflow  $Q_{in}$  and outflow  $Q_{out}$  from the aquifers, both laterally through the gravel pit banks and vertically across its bed. As gravel pit lakes usually have no permanent surface inflow or outflow streams, this term was not included in their water budget. During a given period  $\Delta t$  (T),

$$\frac{h_g^t - h_g^{t-1}}{\Delta t} = \frac{P - E + R + Q_{in} - Q_{out}}{A}, \quad (1)$$

where the first term ( $L \cdot T^{-1}$ ) expresses the rate of lake water level  $h_g$  (L) change between the current time step  $t$  and the previous time step  $t - 1$ , and  $A$  ( $L^2$ ) is the surface area of the gravel pit. The exchange rate of water per unit area  $q$  ( $L \cdot T^{-1}$ ) across a gravel pit-aquifer interface depends on the hydraulic head gradient between the two units and on a specific conductance  $C$  ( $T^{-1}$ ), i.e., a conductance per unit interface area, that is based on the material properties and grid cell dimensions. According to Darcy's law:

$$q = C(h_a - h_g), \quad (2)$$

where  $h_a$  (L) is the hydraulic head in the aquifer. As written,  $q$  is positive when the flow of water is from the aquifer to the lake.  $C$  ( $T^{-1}$ ) is equal to the harmonic mean of the specific conductances of the aquifer and of the lakebed:

$$\frac{1}{C} = \frac{\Delta l}{K_a} + \frac{b}{K_g}, \quad (3)$$

with  $\Delta l$  (L) the half size of the aquifer mesh in the direction of flow,  $K_a$  the hydraulic conductivity of the aquifer ( $L \cdot T^{-1}$ ),  $b$  (L) the thickness of the bed or banks of the gravel pit and  $K_g$  their hydraulic conductivity ( $L \cdot T^{-1}$ ). Groundwater seepage  $Q_{in} - Q_{out}$  is the sum of all individual flows exchanged through the  $N$  gravel pit-aquifer interfaces:  $Q_{in} - Q_{out} = \sum_n^N C_n (h_{an}^t - h_g^t)$ , where  $h_{an}$  is the hydraulic head in the adjacent element of the  $n$ th aquifer-gravel pit interface at time step  $t$ ,  $h_g$  is the lake level at time step  $t - 1$  and  $C_n$  is the conductance ( $L^2 \cdot T^{-1}$ ) of the  $n$ th aquifer-gravel pit interface. Equation (1) is an explicit scheme, used for steady state resolution. In transient state, an implicit or semi-implicit scheme can also be introduced, using  $h_g = (1 - \theta)h_g^{t-1} + \theta h_g^t$ , where  $0 \leq \theta \leq 1$ . In order to determine the equilibrium state of the system,

hydraulic heads in the aquifer and lake level are calculated iteratively, until convergence. The lake water level from the lake water balance is determined after solving the groundwater flow equation:

$$h_g = \frac{P - E + R + \sum_n^N C_n h_{an}}{\sum_n^N C_n}. \quad (4)$$

In transient state, the calculated lake water level in the  $(t - 1)$ th iteration is first used as a Dirichlet boundary condition when solving the groundwater equation then calculated at the current time step according to the following equation:

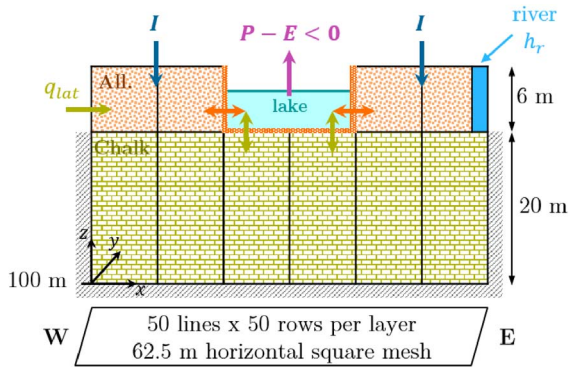
$$h_g^t = \frac{h_g^{t-1} + \Delta t \frac{P - E + R + (\sum_n^N C_n h_{an}^{t-1} - (1 - \theta)h_g^{t-1} \sum_n^N C_n)}{A}}{1 + \frac{\theta \Delta t}{A} \sum_n^N C_n}. \quad (5)$$

### 3.2. Validation on a test case

The numerical performance of the gravel pit package was evaluated by comparison with the state-of-the-art LAK package [Merritt and Konikow, 2000] associated with MODFLOW [Harbaugh, 2005], on a test case consisting of a gravel pit connected to two aquifers, as representative conditions in alluvial aggregate mining areas, such as those found in the alluvial plain of the Seine River, upstream of Paris, France [Schanen, 1998].

#### 3.2.1. Case description

A gravel pit lake, 250 m wide and 500 m long in the direction of flow thus covering about 14 ha, is dug into a 6 m thick first layer of alluvial deposits down to a 20 m thick second layer of chalk, in the centre of a domain of dimensions 3125 m  $\times$  3125 m, with a grid cell size of 62.5  $\times$  62.5 m (Figure 2). The top surface of the model is flat, at an altitude of 126 m. The alluvial aquifer is fed to the west by a lateral flow of  $1.85 \times 10^{-3} \text{ m}^3 \cdot \text{s}^{-1}$ , at the surface by a recharge of about 218 mm/year and is hydraulically connected to the chalk aquifer, itself bounded by no flow conditions on each side and at the bottom. Constant head boundaries were placed along the eastern edge of the first alluvial layer and no flow boundaries, parallel to the west to east groundwater flow. Both aquifer layers are confined, with a horizontal hydraulic conductivity of respectively  $6 \times 10^{-3} \text{ m} \cdot \text{s}^{-1}$  for the top layer and  $5 \times 10^{-4} \text{ m} \cdot \text{s}^{-1}$  for the deeper aquifer, and a vertical-to-horizontal anisotropy ratio of 0.1.



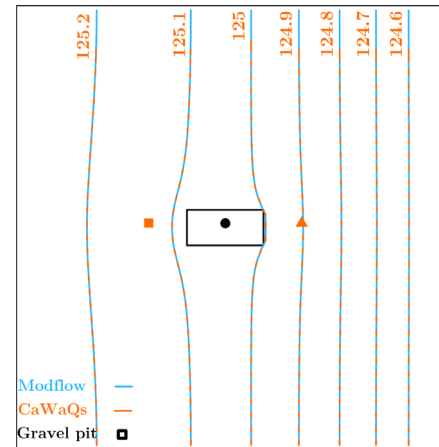
**Figure 2.** Description of the hypothetical two-layer case study.

The surface runoff is considered to be zero. The gravel pit is fed by precipitation of 675 mm/year but is subject to a higher evaporation of 710 mm/year. The same specific conductance is chosen for the bed and the banks of the gravel pit, which translates using Equation (3) into equivalent specific conductances of  $4.0 \times 10^{-5} \text{ s}^{-1}$  for the vertical interfaces separating the gravel pit from the alluvial deposits and  $4.5 \times 10^{-6} \text{ s}^{-1}$  for the horizontal interfaces between the lake bottom and the chalk. Multiplying by the surface area of the interface gives conductances  $C_n$  of  $1.5 \times 10^{-2} \text{ m}^2 \cdot \text{s}^{-1}$  and  $1.8 \times 10^{-2} \text{ m}^2 \cdot \text{s}^{-1}$  respectively.

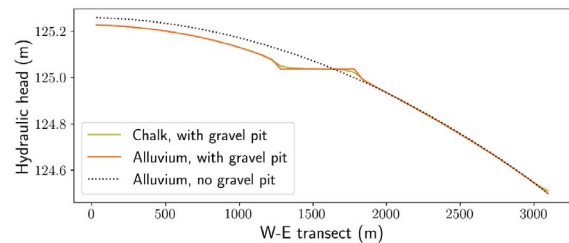
Two simulations were carried out, in steady and transient states, using a time step of one day and the explicit procedure for the lake scheme. The threshold for groundwater convergence was set to 0.0001 m. The forcings remain constant during the transient simulation, which is initialised under arbitrary conditions but sufficiently far from the steady state, with the initial lake level at its minimum and contrasting conditions in the two aquifers. For a sufficiently long simulation, the transient solution must converge towards the simulated steady-state. The proposed setup and model parameters are summarised in Table 2.

### 3.2.2. Simulation results

A first set of simulations made it possible to ensure the similarity of the results produced by the two codes CAWAQS and MODFLOW in the absence of a gravel pit. In steady-state, the simulated head differences over the study area are less than 1 mm and negligible in terms of calculation accuracy. The



(a) Groundwater level contours (m) simulated using MODFLOW or CAWAQS (equipotential lines overlap) in steady-state in the alluvial aquifer in the presence of a gravel pit in the centre of the study area. The modelled area is 3125 m on each side.



(b) Water table profiles with and without a gravel pit lake as simulated by CAWAQS along a west-east flow line and across the water body.

**Figure 3.** Simulation results for validation.

introduction of a gravel pit has an impact on the distribution of equipotential lines and therefore on the flow pattern in the vicinity of the water body, converging upstream and diverging downstream (Figure 3a). The water table is lowered upstream of the gravel pit lake up to more than 8 cm, while a high hydraulic gradient develops downstream (Figure 3b). Between the two models CAWAQS and MODFLOW, the simulated head differences in the alluvial aquifer do not exceed 0.1 mm, within the margin of error of the case without a gravel pit. The water budget of the gravel pit is detailed in Table 3. The atmospheric moisture deficit  $E - P$  of  $35 \text{ mm} \cdot \text{year}^{-1}$  results in a net inflow to the gravel pit of the same magnitude, of which nearly three quarters are from the alluvial aquifer. There is no difference between the flows calculated by each code.

When using the explicit procedure at a daily time step, the simulated lake level evolution was almost identical between the lake module of CAWAQS and

**Table 2.** Definition of the study area: model setup and parameters

	Alluvium (layer 1)	Chalk (layer 2)	Gravel pit (bed & banks)	River
Horizontal hydraulic conductivity $K_h$ ( $\text{m}\cdot\text{s}^{-1}$ )	$6 \times 10^{-3}$	$5 \times 10^{-4}$		
Vertical hydraulic conductivity $K_v$ ( $\text{m}\cdot\text{s}^{-1}$ )	$6 \times 10^{-4}$	$5 \times 10^{-5}$		
Storage $S$ (-)	0.06	0.001		
Thickness $e$ (m)	6	20		
Specific conductance $C_g$ ( $\text{s}^{-1}$ )			$5 \times 10^{-5}$	
<i>Initial conditions</i> $h_0$ (m)	123	122	120	124.5
<i>Boundary conditions</i>				
Dirichlet $h_r$ (m)				124.5
Neumann $q_{\text{lat}}$ ( $\text{m}^3\cdot\text{s}^{-1}$ )	$1.85 \times 10^{-3}$			
Recharge $I$ ( $\text{m}\cdot\text{s}^{-1}$ )	$6.9 \times 10^{-9}$			
Precipitations $P$ ( $\text{m}\cdot\text{s}^{-1}$ )			$2.14 \times 10^{-8}$	
Evaporation $E$ ( $\text{m}\cdot\text{s}^{-1}$ )			$2.25 \times 10^{-8}$	

**Table 3.** Gravel pit lake water balance in steady-state

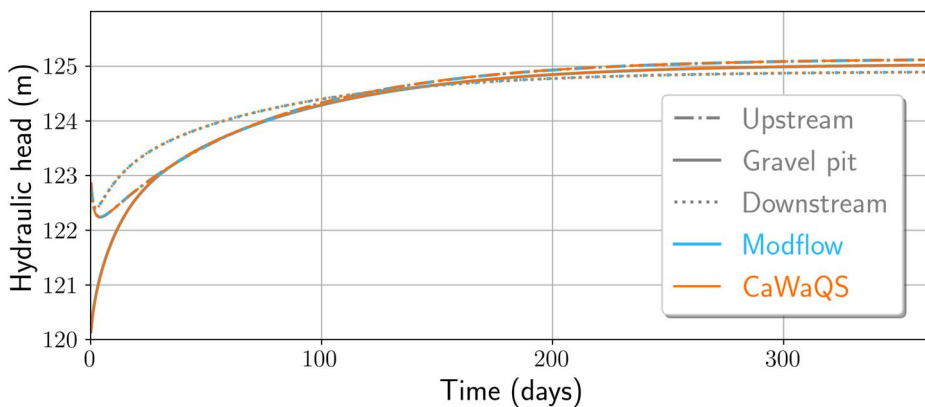
	Precipitation ( $\text{m}^3\cdot\text{day}^{-1}$ )	Evaporation ( $\text{m}^3\cdot\text{day}^{-1}$ )	Groundwater flows ( $\text{m}^3\cdot\text{day}^{-1}$ )			
			Inflow		Outflow	
			InV	InH	OutV	OutH
CAWAQS	260.0	273.4	436.1	164.3	425.1	161.9
MODFLOW						

$V$  and  $H$  indicate flow through vertical interfaces between the gravel pit and the first layer and through horizontal interfaces between the gravel pit and the deeper layer, respectively.

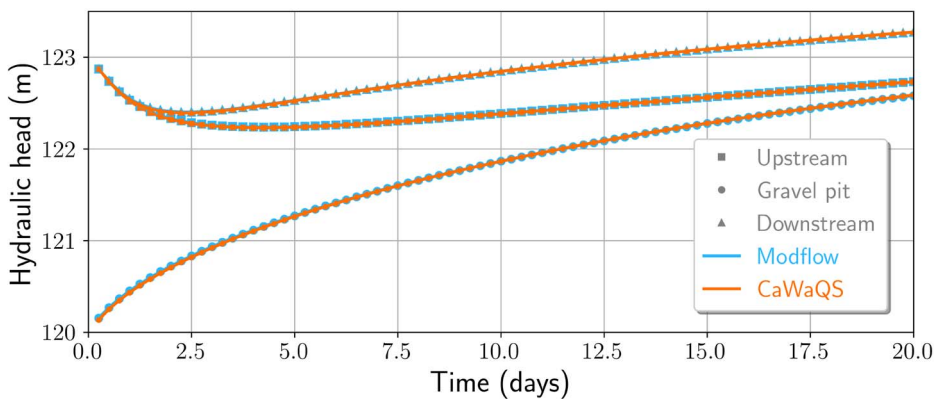
LAK, despite some differences at the beginning of the simulation, maximum on the first day (0.1 m). By decreasing the time step from one to half and then a quarter day, the transient CAWAQS model results in a consistent lake level, whatever the time step, with a maximum difference of less than 0.5 m at the beginning of the simulation, equivalent to that obtained with MODFLOW (0.48 m versus 0.42 m respectively) but which lasts longer. The temporal evolution of the heads simulated by the two numerical codes is compared at a time step of 0.25 day at two points in the alluvial aquifer, upstream and downstream of the gravel pit (see location in Figure 3a), and in the gravel pit itself. Figure 4a illustrates the convergence of the water levels in the aquifer and the gravel pit lake towards their equilibrium values. For the gravel pit, the difference in lake level is maximum on the first day of simulation (0.021 m) and decreases with

time (Figure 4). It is higher, although still acceptable, than that obtained at the same point in the alluvial aquifer in the absence of a gravel pit (0.002 m). Water balance calculations for the lake show a similar evolution of the inflow and outflow of the gravel pit from one code to another (Figure 5). After an initial filling phase, the lake acts as a flow-through system, with the rate of groundwater inflow nevertheless always exceeding the outflow (see steady-state water balance in Table 3).

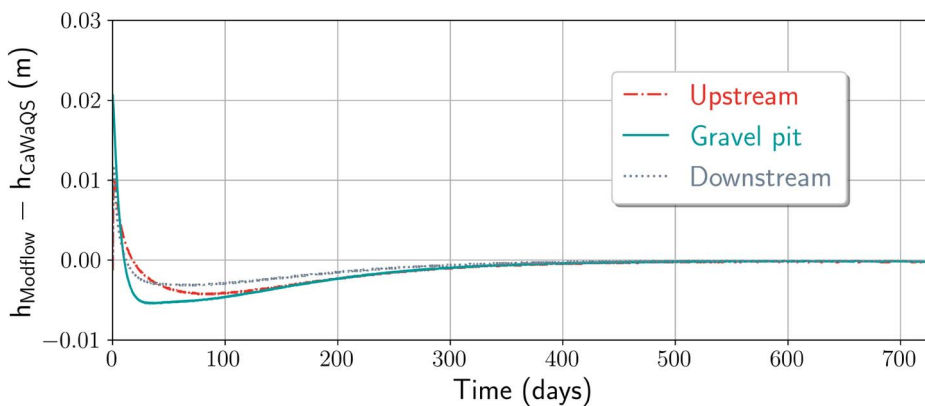
Thus validated and operational, the lake module has been integrated into the modelling platform as a specific library, which can be activated if needed. The test case configuration was taken as the baseline model for further sensitivity analysis to explore the impact of the introduction of a gravel pit lake on the behaviour of a hydrogeological system composed of two connected aquifers in interaction with a river.



(a) Temporal evolution of heads simulated by CAWAQS and MODFLOW in the gravel pit lake and in the alluvial aquifer, on both sides of the gravel pit along a flow path. The overlap between CAWAQS and MODFLOW is further illustrated in Figure 4b.

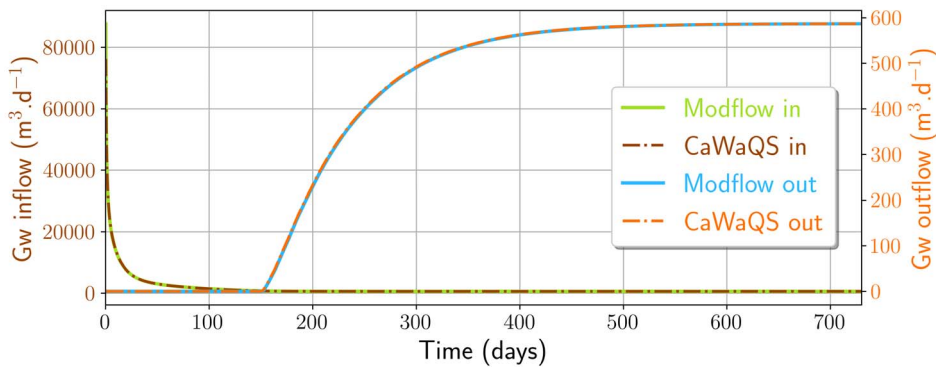


(b) Focus on the first 20 days of simulation.



(c) Simulated head differences ( $h_{Modflow} - h_{CaWaQS}$ ).

**Figure 4.** Detailed comparison CAWAQS–MODFLOW at three points (see location in Figure 3a) under transient conditions (explicit scheme, 0.25 day time step).



**Figure 5.** Flows into and out of the gravel pit lake during the transient run, as simulated by the two codes.

#### 4. Numerical experiments characterising gravel pit lake–aquifer interactions

In a third step, *in silico* experiments are run to illustrate the general principles governing the interactions between a gravel pit lake and the aquifer system in which it is embedded, principles which were recalled in the first part of the present work. The modelling approach offers the possibility of quantifying the exchange flows between the gravel pit and the aquifers and of simulating the changes in lake level and water table following gravel excavation. We first simulate the gravel pit lake–groundwater system under steady-state conditions. By varying the various parameters that could impact groundwater seepage to the lake and local water levels, we analyse their relative significance. We then examine the response of the gravel pit lake to cyclic transient boundary conditions, whose influence on the seasonal water balance terms is also investigated.

##### 4.1. Simulations design

Steady-state simulations were carried out from the base case (#1) using various configurations relative to (i) geometrical factors (#2 to #8): size, shape and depth of the gravel pit lake, number of lakes, lake orientation with respect to the direction of flow, distance of the gravel pit from the river, (ii) hydrodynamical factors (#9 to #15): hydraulic conductivity contrasts within the groundwater system, ratio of horizontal to vertical hydraulic conductivity and clogging of the lake bed and banks, and (iii) meteorological factors (#16 to #19): groundwater recharge

( $I$  and  $q_{lat}$ ) and atmospheric moisture deficit ( $E - P$ ). The values of the parameters chosen for the tests are shown in Table 4. Although hypothetical settings are simulated, it is expected that these values and the selected magnitude in their change from the baseline conditions are realistic and representative of gravel pit lake environments.

Two sets of transient simulations are used, one involving a step change in water inflow and the other a sinusoidal variation in its boundary conditions. The first set, based on all the simulations carried out in steady-state and defined as initial conditions, consists of cancelling all the recharge terms of the lake–groundwater system ( $q_{lat}$ ,  $I$  and  $P - E$ ) and following the rate of gravel pit lake level recession until a new equilibrium is approached. The response time  $\tau_c$  of the gravel pit lake is then deduced. This characteristic time is given by the exponential decay constant, adjusted between 10 and 90% of the total lake level change due to the cessation of recharge, in order to take into account the time necessary for the system to reach a purely exponential decay [Cuthbert, 2014]. A value of 0.06 is assigned to the storage coefficient of the upper aquifer and of 0.001 for the lower aquifer, resulting in hydraulic diffusivity  $T/S$  of 0.6 and  $10 \text{ m}^2 \cdot \text{s}^{-1}$ , respectively.

Additional transient simulations are developed to examine the response of the groundwater–lake system to periodically oscillating recharge ( $I$  and  $P - E$ ) and boundary conditions ( $q_{lat}$ ,  $h_r$ ), independently at first and then by applying all the forcings simultaneously. The sinusoidal input signal is expressed in the form of  $a \cos(\omega t)$ , where  $a$  is the driving force

**Table 4.** Design of the steady-state simulations: model parameters

Simulations	Geometrical factors					
	Size (m <sup>2</sup> )	Shape	Depth (m)	Number	Orientation	Position
#1 ref.	140,625	Rectangular	6	1	Parallel	Central
#2 size	257,812.5					
#3 square		Square				
#4 depth			20			
#5 × 3	46,875			3	Perpendicular	
#6 ppdcl					Perpendicular	
#7 up						Upstream
#8 down						Downstream
Hydrodynamical factors						
	Transmissivity $T$ (m <sup>2</sup> ·s <sup>-1</sup> )		Anisotropy		Clogging $C_g$ (s <sup>-1</sup> )	
	Alluvium	Chalk	$K_v/K_h$			
#1 ref.	$3.6 \times 10^{-2}$	$10^{-2}$	0.1			$5 \times 10^{-5}$
#9 homo		$3.6 \times 10^{-2}$	1		Bed	$5.09 \times 10^{-5}$
#10 T <sub>all</sub>	$3.6 \times 10^{-1}$					
#11 T <sub>ch</sub>		$10^{-4}$				
#12 $\alpha$			0.5			
#13 noC					No clogging	
#14 C <sub>4</sub>					All but upstream	$10^{-8}$
#15 C <sub>tot</sub>					All banks & bed	$10^{-8}$
Meteorological factors						
	$q_{lat}$ (m <sup>3</sup> ·s <sup>-1</sup> )		$I$ (m·s <sup>-1</sup> )/(mm·year <sup>-1</sup> )		$P - E$ (m·s <sup>-1</sup> )/(mm·year <sup>-1</sup> )	
#1 ref.	$1.85 \times 10^{-3}$		$6.9 \times 10^{-9} / 218$		$-1.1 \times 10^{-9} / -35$	
#16 E <sup>+</sup>					$-1.38 \times 10^{-8} / -44$	
#17 P					$1.1 \times 10^{-9} / 35$	
#18 P <sup>+</sup>					$1.38 \times 10^{-8} = 2 \times I_{\#1} / 44$	
#19 q <sup>+</sup>	$6.51 \times 10^{-2} = I_{\#1}$		$1.96 \times 10^{-10} = q_{lat\#1} / 6$			

amplitude (L, L·T<sup>-1</sup> or L<sup>3</sup>·T<sup>-1</sup>),  $\omega = 2\pi/T_o$  is the oscillation frequency (T<sup>-1</sup>),  $T_o$  is the oscillation period (T) and  $t$  is the time (T). At the river boundary, the head amplitude is about 0.3 m. In this scenario, mean driving forces were applied so that  $q_{lat} = I = 2.23 \times 10^{-2}$  m<sup>3</sup>·s<sup>-1</sup> and  $I = E - P = 2.36 \times 10^{-9}$  m·s<sup>-1</sup>. The period is 1 year, i.e., seasonal variations are investigated. The resulting cyclic stresses are illustrated in Figure 7. Steady-state groundwater heads and lake level are used as initial conditions and the transient model is run with daily time steps until a quasi-steady oscillatory state is reached. Models are used to evaluate the response of the gravel pit lake's level to

periodically changing hydraulic conditions, for different configurations involving increasing clogging of the lake (ref., C<sub>4</sub> and C<sub>tot</sub>).

## 4.2. Simulation results

### 4.2.1. Steady-state analysis

The hydrodynamic disturbance caused by the gravel pit lake, as simulated by the gravel pit lake-aquifer system models, is assessed using several criteria (see Table 5) that measure (i) the impacts in terms of groundwater and lake levels, as compared

**Table 5.** Main results of the steady-state and transient simulations: water levels and balance components

Simulations	Water levels					Water balance					
	$\Delta h_-$ (m)	$\Delta h_+$ (m)	$h_L$ (m)	$\kappa$ (%)	$\tau_c$ (days)	$Q_{in}$ ( $m^3 \cdot day^{-1}$ )	$E/(Q_{in} + P)$ (%)	InV (%)	InH (%)	$\tau_r$ (years)	
#1 ref.	<b>-0.084</b>	<b>0.038</b>	<b>125.037</b>	<b>70</b>	73	600	10	32	<b>73</b>	<b>27</b>	<b>3.2</b>
#2 size	-0.116	0.035	125.017	77	<b>84</b>	851	<b>15</b>	38	72	28	4.2
#3 square	-0.059	0.019	125.035	73	74	569	10	33	73	27	3.4
#4 depth	-0.088	0.033	125.033	73	71	738	<b>13</b>	<b>27</b>	84	16	2.6
#5 $\times 3$	-0.032	0.004	125.048	75	74	729	<b>13</b>	28	79	21	2.7
#6 ppdcl	-0.040	0.004	125.054	80	75	590	10	32	74	26	3.3
#7 up	-0.050	0	125.201	$\emptyset$	<b>85</b>	225	<b>4</b>	<b>56</b>	73	27	<b>8.9</b>
#8 down	-0.122	0.074	124.718	63	<b>60</b>	988	<b>17</b>	<b>22</b>	73	27	<b>1.8</b>
#9 homo	-0.055	0.023	124.841	71	78	661	11	30	<b>35</b>	<b>65</b>	2.8
#10 $T_{all}$	<b>-0.009</b>	<b>0.006</b>	124.568	62	<b>10</b>	180	<b>3</b>	62	<b>78</b>	<b>22</b>	<b>9.9</b>
#11 $T_{ch}$	-0.106	0.049	125.182	69	<b>94</b>	583	10	32	<b>100</b>	<b>0</b>	3.4
#12 $\alpha$	-0.084	0.037	125.034	70	71	627	11	31	<b>65</b>	<b>35</b>	3.1
#13 noC	-0.091	0.031	125.029	<b>75</b>	70	840	15	25	<b>90</b>	<b>10</b>	2.3
#14 $C_4$	-0.038	0.143	<b>125.143</b>	<b>103</b>	102	25	<1	96	<b>100</b>	<b>0</b>	<b>79.7</b>
#15 $C_{tot}$	-0.171	0.019	<b>124.950</b>	<b>-4</b>	1110	13	<1	<b>100</b>	<b>7</b>	<b>93</b>	<b>142.6</b>
#16 $E^+$	-0.110	0.011	125.011	<b>87</b>	73	667	12	35	72	28	2.9
#17 P	-0.079	0.042	125.042	<b>66</b>	73	589	10	30	73	27	3.3
#18 $P^+$	-0.053	0.069	125.068	<b>47</b>	73	524	9	28	73	27	3.7
#19 $q^+$	-0.134	0.095	125.205	<b>58</b>	73	1138	<b>20</b>	20	73	27	1.8

$\Delta h_-$  and  $\Delta h_+$  are respectively the maximal fall and rise in water levels along a flow path across the gravel pit lake, compared to the equivalent case without a lake,  $h_L$  is the simulated gravel pit lake level,  $\kappa$  is the position of the Kippungslinie within the lake as counted from upstream,  $\tau_c$  is the response time of the gravel pit lake,  $Q_{in}$  is the groundwater inflow to the gravel pit lake, also expressed as a percentage of total flow in the aquifer system,  $E/(Q_{in} + P)$  is the fraction of precipitated water and inflowing groundwater evaporated from the gravel pit lake, InV and InH are inflows through vertical interfaces between the gravel pit and the alluvial layer and through horizontal interfaces between the gravel pit and the chalk layer, respectively, here expressed as a percentage of total inflow  $Q_{in}$ , and  $\tau_r$  is the water residence time of the gravel pit lake. Numbers in bold are discussed more specifically in the text.

to equivalent baseline simulations with no gravel pit: to this end, we calculate the difference in water levels and the position of the Kippungslinie ( $\kappa$ ); (ii) changes in the water budget components, (iii) the water residence time of the gravel pit lake.

For the reference simulation #1, the steady-state model results in a gravel pit lake level of 125.037 m, a level that corresponds to the pre-extraction water table established at a distance of 70% of its size from its upstream side. For an unclogged gravel pit lake, it moves not at mid-point of the gravel pit lake

[Wilson, 1984] but downward at 75%. This is in line with the parabolic water table profile, here related to the specified recharge and which should apply further in the case of an unconfined shallow aquifer. The position of the Kippungslinie therefore reflects not only the presence and extent of the lake sealing as expected [Wilson, 1984, simulations #14 and #15] but also the recharge conditions of the system (see simulations #16, #18 and #19). A positive atmospheric lake balance (#18) equilibrates the lake at mid-point while as the  $E - P$  deficit increases, so

decreases the predicted level of the gravel pit lake (#16, Figure 6a), as well as the water table downstream of the lake. The main changes in lake water levels are nonetheless predicted as a consequence of lake sealing. While the two simulated water table profiles in the case of clogged lake (#14 and #15) are similar, characterised by slight rise and fall respectively upstream and downstream of the lake, e.g., a limited impact on alluvial groundwater levels, the predicted lake levels are opposite. A partially clogged lake gives rise to a high lake level (125.143 m), as often described in the literature [e.g., Peaudecerf, 1975] and acts as a reservoir fed by the alluvial aquifer upstream (100% of inflowing groundwater) and mainly drained by the chalk aquifer. In contrast, if totally sealed, the gravel pit lake ends as a terminal lake, with a low level (124.950 m), nearly 20 cm lower, mainly fed by the deep aquifer and where groundwater inflow exactly compensates for the atmospheric moisture deficit. A high hydraulic head gradient, of 3.3‰, i.e., more than ten times greater than the initial gradient, is found upstream of the excavation, as opposed to its downstream location in  $C_4$  simulation. The hydraulic gradients on either side of the lake are all the more pronounced as the clogging is significant. In the absence of clogging, the transition between the gravel pit lake and the water table is smoothed. In this case, the simulated water table is similar to that obtained from an additional run using the “high K” technique (referred to as #13b in Figure 6a).

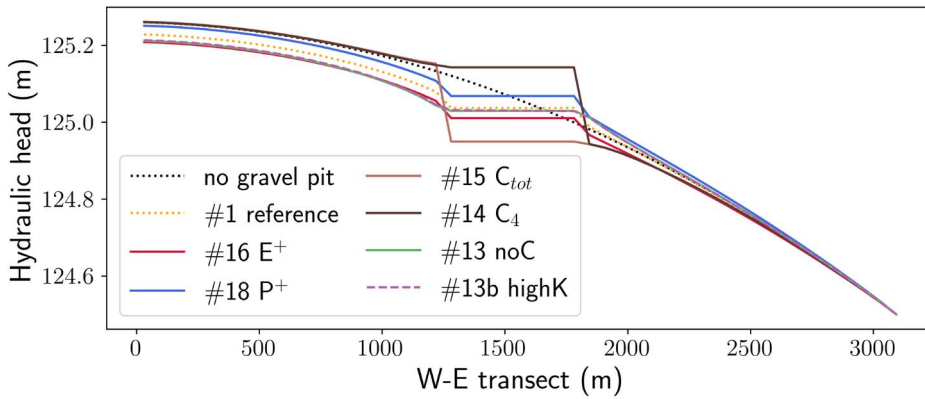
In the chosen configuration, with a Dirichlet boundary condition downstream and a Neumann boundary condition upstream, the impact of the gravel pit lake on groundwater levels is mainly felt upstream, where the water table is generally lowered as compared to the pre-extraction levels, whereas the lake influence is rather limited downstream of the gravel pit lake (Figure 6). The intensity of the response of the aquifer system to the creation of the water body is also conditioned by the position of the lake along the flow line. Far from the low points of the valley, only a limited drop in the water table is simulated (#7, Figure 6b). It is even more pronounced the closer the gravel pit lake is to the river (#8), where the initial water table gradient is higher. This is also the case when the transmissivity of the system decreases (#11, for the chalk aquifer, Figure 6c). On the other hand, in a highly permeable aquifer (#10, for the alluvial aquifer), the change in water table from the

pre-extraction situation will be difficult to measure (see Table 5). The impact of the gravel pit on groundwater levels can finally be reduced by choosing a gravel pit layout perpendicular to the flow, which offers a larger exchange surface to the flow (#6 and #3), by favouring the number of water bodies (#5, Figure 6b) rather than a single large lake (#2), whose level could end up higher than the natural ground level downstream.

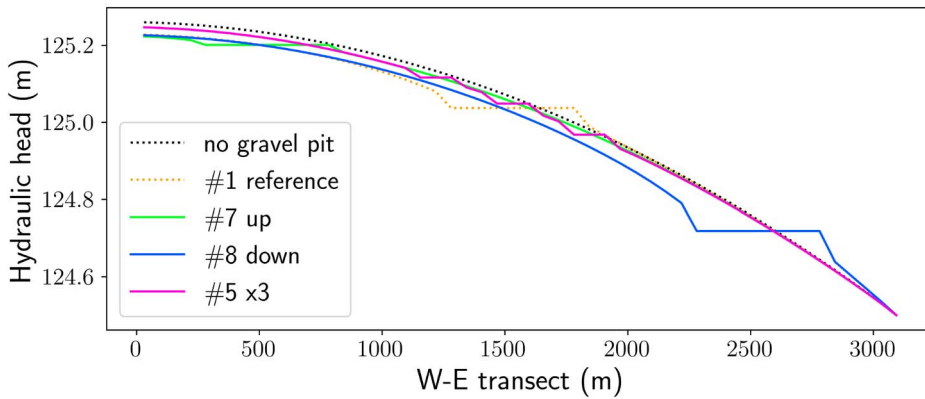
With regards to the water balance, groundwater exchange with the gravel pit lake represents, in the examples considered here, up to 20% of the total flow circulating in the aquifer system. It is enhanced when a higher upstream groundwater flow is intercepted by the water body (#19 and #8), especially if the gravel pit is larger or deeper (#2 and #4), and it is generally accompanied by a decrease in inflowing water evaporated from the lake. A greater number of water bodies also means an increased rate of water cycling (#5). Two parameters of the model again play a major role in controlling the lake seepage, namely the bed and banks conductance (#14 and #15) and the hydraulic conductivity of the superficial aquifer (#10). Indeed, the rate of groundwater input to the lake becomes a negligible component of the overall water balance when clogging is severe. With respect to the hydraulic conductivity of the alluvial aquifer, the model results indicate that there is a reduction in groundwater inflow to the lake, which is related to the lower hydraulic gradient, while the conductance of the lake/aquifer interface remains dominated by bank clogging.

For an isotropic and homogeneous porous media (#9), upward flux of deep groundwater into the gravel pit lake is predominant due to the larger area offered to the flow by the bottom of the gravel pit. However, if the medium is anisotropic and the hydraulic conductivity decreases with depth, results show that the majority (three quarters in the reference case) of the lake in seepage comes from the shallow aquifer (horizontal conductance of the interfaces one order of magnitude higher than the vertical conductance in simulation #1). The share of flow through the lake bottom may increase in case of clogging if it is homogeneous (see noC, ref. and  $C_{tot}$  simulations). The relative contributions of the two aquifers are indeed variable and depend on the hydraulic conductivity contrast between the two units (#9–11), the ratio of horizontal to vertical hydraulic conductivity (#12) and the extent

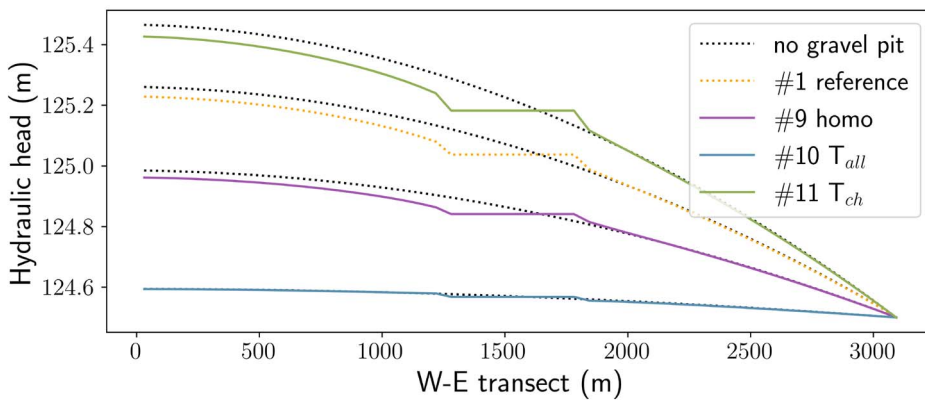




(a) Gravel pit lake meteorological factors & clogging: E<sup>+</sup> stands for an increased  $E - P$  deficit over the gravel pit lake,  $P - E > 0$  in P<sup>+</sup>, #13b uses the “high K” technique, there is no clogging in #13, partial clogging in #14 and total clogging in #15.

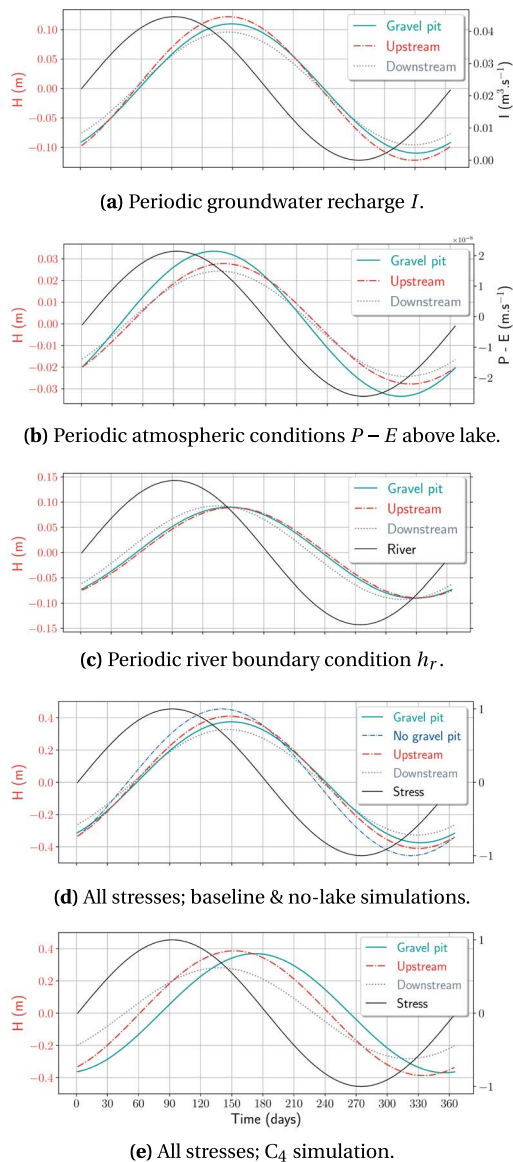


(b) Descriptive factors of the gravel pit lake: location of the gravel pit lake, #7 upstream, #1 in the centre and #8 downstream; lake cut into three pieces #5.



(c) Aquifer hydrodynamic factors: homogeneous aquifers in #9, more permeable alluvial aquifer in #10 and less permeable deeper aquifer in #11, as compared to the reference #1.

**Figure 6.** Selected results of the steady-state simulations along a west–east flow line and across the water body, for different configurations of the gravel pit (see location in Figure 3a).



**Figure 7.** Theoretical response of the gravel pit lake level, and of selected upstream and downstream water levels in the alluvial aquifer (see location in Figure 3a), here represented by their amplitude, to various simple sinusoidal hydraulic stresses applied to the lake and the aquifer (shown in black in the figure). Results correspond to baseline simulation, except (e), for a partially clogged gravel pit lake ( $C_4$  simulation). Also plotted in (d), the equivalent upstream groundwater level simulated in the absence of a gravel pit. The graphs share a common  $x$ -axis.

of lake sealing (#13–15). They are highly dependent on local groundwater flow conditions, according to whether the gravel pit is complete and rests on a low permeability layer (no exchange via the bottom of the gravel pit in the case tested here, #11), on the vertical heterogeneity of the bank clogging, or on whether the banks are more clogged than the bottom of the gravel pit due to the partial backfilling of the excavation with the overburden.

The computation of the water residence time  $\tau_r$  in the gravel pit lake summarises the previous observations concerning the level at which the gravel pit lake equilibrates and its exchanges with the aquifers. In this theoretical case study, the water retention time is around 3 years. In particular, it is variable along a flow path, here five times higher upstream than downstream (down versus up simulations). It becomes longer as the amount of seepage is reduced (e.g.,  $T_{\text{all}}$  simulation), mainly with the ageing of the gravel pit and the subsequent clogging (residence time more than one order of magnitude higher than the median).

#### 4.2.2. Transient analysis

The transient numerical models are used to investigate the dynamic changes in gravel pit lake water level as a result of variations in the balance between inputs and outputs of water. Estimating gravel pit lake response times provides information on the time required for the lake level to adjust to these changes. They are shown in Table 5 for each of the 19 cases. The calculated  $\tau_c$  values range from 2 to 3 months. It is dependent on the horizontal distance away from the divide, and is longer for the lake in the most up-gradient position. The lake response is also defined by the size of the gravel pit. The results mainly illustrate how the response for any gravel pit lake is determined by the hydrodynamics properties of the lake-groundwater system. The larger the hydraulic diffusivity of the shallow aquifer, the shorter the time to reach the equilibrium lake level and conversely when the gravel pit lies on a low permeability bedrock. Finally,  $\tau_c$  depends largely on the connectivity between the gravel pit lake and groundwater, i.e., on the conductance of their interfaces. Significantly slower responses, up to 3 years, are simulated as a result of sealing of the bed and banks of the gravel pit lake.

For small lakes such as those in gravel pits, however, relatively short response times are expected and

**Table 6.** Results of the baseline transient simulations under the action of each stress separately and then all together: amplitude  $H$  and phase shift  $\phi$  of the gravel pit lake level and of the alluvial groundwater level at about 100 m upstream and downstream of the lake

Stress	Upstream		Gravel pit lake		Downstream	
	$H$ (m)	$\phi$ (days)	$H$ (m)	$\phi$ (days)	$H$ (m)	$\phi$ (days)
$I$	0.122	52	0.110	56	0.096	51
$q_{\text{lat}}$	0.172	56	0.145	65	0.120	65
$h_r$	0.089	56	0.090	53	0.093	40
$P - E$	0.028	42	0.033	36	0.024	38
All	0.410	55	0.375	58	0.328	53
No gravel pit	0.455	48	—	—	0.336	44
All $C_4$	0.387	60	0.370	80	0.282	46
All $C_{\text{tot}}$	0.369	44	0.130	91	0.216	46

Also presented are the results of an equivalent simulation with no gravel pit lake, in response to the periodic variations of the first three hydraulic stresses, and two additional simulations involving respectively a partially and a totally clogged gravel pit lake.

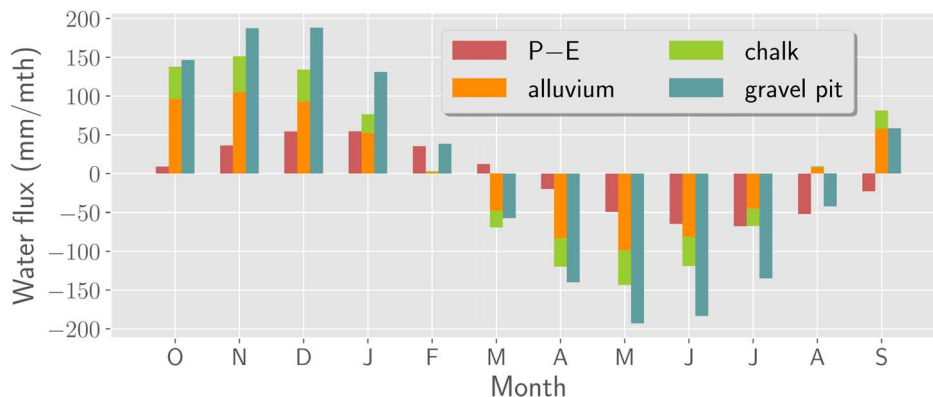
hence an ability to propagate forcings whose period of fluctuations is of the same order as the response time, i.e., also short. In this theoretical case study, the annual period is close to  $2\pi\tau_c$ . For lower frequency components, the gravel pit lake will remain approximately in equilibrium while high frequency signals will be attenuated independently of  $\tau_c$ , the lake acting as a low pass filter [Mason et al., 1994].

We now examine the one-year response of the gravel pit lake to such periodic forcing, once the lake level has reached a pseudo-sinusoidal steady-state. Figure 7 illustrates the rise and fall of water levels around the average in the gravel pit lake and in the shallow aquifer about 100 m upstream and downstream of the lake, in response to seasonal varying groundwater (Figure 7a for  $I$ ; not shown for recharge flux  $q_{\text{lat}}$  across the lateral boundary) and lake recharge (Figure 7b), river head (Figure 7c) and to all hydraulic stresses considered together (Figure 7d), for the base case. Water table fluctuations are greater near the groundwater divide than in the discharge area, except where changes in the river level are the main driver, for example during high-flow periods. Their simulated amplitude  $H$  and phase  $\phi$  values are summarised in Table 6. Relative to the no-lake case, the results show that the presence of the gravel pit lake dampens the groundwater wave propagation, as the water table exhibits smaller amplitude and increased phase lag (see also Figure 7d),

underlining the dominant role of the storage effect of the lake, as pointed out by Jazayeri et al. [2021]. Regarding the gravel pit lake's level itself, slower readjustments also occur in response to the sinusoidal variation in groundwater recharge, as compared to the aquifer, to be related to the response time  $\tau_c$  of the gravel pit lake. The influence of the varying precipitation inputs and evaporation outputs into/from the lake is felt beyond its physical extent on either side of the lake, as illustrated in plot Figure 7b, but is expected to be of limited extent, due to the small amplitude involved, and masked by the variability of groundwater seepage (Figure 7d).

How the sealing layer on the banks and at the bottom of the lake influences the attenuation and phase shift of the input signal within the lake-aquifer system is illustrated using the cases of a partially and a totally clogged gravel pits ( $C_4$  and  $C_{\text{tot}}$ ). Due to the longer response time of the water-filled pit, a further attenuation in amplitude and a significant shift in phase (from 58 to 81 or even 91 days) of the lake's response are simulated (shown for  $C_4$  case in Figure 7e), while contrasting effects on signal propagation are noted when comparing upstream and downstream of the alluvial aquifer on either side of the lake.

The water regime of the gravel pit lake is not only the result of precipitation inputs exceeding evaporation outputs or vice versa but is mainly determined



**Figure 8.** Water balance of the gravel pit lake: the monthly averages of net atmospheric inputs ( $P - E$ ), net seepage from the alluvial and chalk aquifers ( $Q_{in} - Q_{out}$ ), and variations in the gravel pit storage are shown for the reference transient simulation, under the effect of all applied stresses.

by the seasonal variability of net groundwater flow  $Q_{in} - Q_{out}$ . Assessing groundwater–lake exchanges is therefore a key element in the overall understanding of gravel pit lake hydrology. The simulated seasonal distribution of the lake water balance components over a hydrological year supposed to start in October is presented in Figure 8, where positive values indicate fluxes recharging the pit lake. During the wet season, water storage in the lake is provided by net influx of groundwater and direct precipitations, in proportions that will depend on local climate, hydrogeological conditions and gravel pit lake hydraulic properties. During the dry season, the surface water evaporation is mainly compensated by the release of water stored during the preceding high-flow period, with an additional groundwater supply, at the end of the dry period in the baseline case or in the case of a completely clogged gravel pit ( $C_4$ ).

## 5. Conclusion

The theoretical numerical case studies, conducted in this work under steady-state and transient regimes for a single gravel pit lake, illustrate the diversity of the situations that can be encountered under real-world conditions. They are useful for understanding the mechanisms that govern the temporal and spatial occurrence of groundwater–gravel pit lake interactions. The intensity of their exchanges is one of the particularities of these man-made lakes, which distinguishes them from natural lakes, although the

importance of this process is increasingly recognised even for the latter [e.g., Hokanson *et al.*, 2022]. The gravel pit lake water level and its changes are primarily dependent on the variability of the groundwater contribution, while the balance between precipitation and evaporation plays a secondary role. The significant amount of groundwater through-flow, combined with the small size of the open pits, results in short residence times. Low values, ranging from a few days to one and a half years, are indeed generally compiled for gravel pits [Schanen, 1998, Weilhartner *et al.*, 2012]. This has implications for the nutrient budget of the gravel pit lake and hence, their ecology and the effective management of these aquatic ecosystems. Accurate spatial and temporal quantification of groundwater inflow and outflow to the lake is a prerequisite for estimating the mass balance of its dissolved components.

Despite the interannual variability due to surface water evaporation is generally outweighed by the through-flow of groundwater, it is still a potential sink term for the aquifer system. More investigation is needed to obtain valid estimations of the magnitude and seasonal distribution of open water evaporation in the particular case of gravel pit lakes. In this respect, state-of-the-art thermodynamic lake models [Ottlé *et al.*, 2020] or isotope mass balance models [Gibson, 2002] should prove useful in estimating the evaporation losses. A generalisation of studies on gravel pit lakes to any type of climatic context is also awaited, as well as a broad assessment

of how groundwater-gravel pit lake interactions may be modified as a result of climate change, following the pioneer work of Mollema and Antonellini [2016]. Given their small size, it is expected that gravel pit lakes will have a more sensitive short-term response to seasonal events than to low-frequency climatic events such as droughts, but this needs to be checked against the geographical context of each site.

Assessing the response of the gravel pit lake-aquifer system to changes in climate conditions and land use may require further modelling developments to take into account a wider range of processes such as a more comprehensive description of vertical flows in the vadose zone, whether to better estimate infiltration and evapotranspiration processes or additional delays in water transfer to the aquifer. Also, there is less understanding of how extreme events, such as flash floods or droughts, may influence this behaviour [Cross *et al.*, 2014] and additional coupling of the subsurface to the surface would be useful in this respect. This is indeed one of the limitations of the simple modelling approach we have developed within the CAWAQS platform, as well as the simplification of solving the flow equation only for a confined aquifer for the moment. In this latter case, this leads in particular to a slight overestimation of the flows exchanged with the shallow aquifer (of the order of +5%) at the expense of the deep aquifer (−10%), as compared to the results of an additional baseline run carried out using LAK-MODFLOW in unconfined mode. Greater numerical stability and ease of deployment at the scale of a real case study with multiple water bodies are however expected.

The next step is to use this mathematical tool for modelling purposes in such a regional case study. Although gravel pit lakes are becoming an increasingly common type of freshwater and their local impacts are regularly examined, they are more rarely studied on a regional scale. To fill this gap, we are currently developing an application in the alluvial plain of the Seine River, upstream of Paris. In the so-called Bassée region, a thousand small water bodies resulting from post-war sand and gravel extraction are scattered across the floodplain. The existence of numerous water level records on well-maintained sites, future accurate remote sensing of their temporal variations by the SWOT satellite, and the present numerical code will be of great benefit in deciphering the cumulative effect of a multitude of gravel pits

of different shape and size in a heterogeneous hydrogeological environment. Such a combination of tools will also be useful in monitoring water resources in strategic areas where water and aggregate resources coexist.

### Conflicts of interest

The authors declare no competing financial interest.

### Dedication

The manuscript was written by AJ. SW developed the first version of the code and TV updated it, under the supervision of AJ. FC participated in the sensitivity analysis conducted by AJ. NF provided access to CAWAQS and expertise. All authors have given approval to the final version of the manuscript.

### Acknowledgements

This work was funded by the Centre National d'Etudes Spatiales (CNES) through the TOSCA-SPAWET programme, with additional support from the PIREN-Seine programme. The authors warmly thank the two referees whose valuable comments helped to clarify the manuscript.

### References

- Apaydin, A. (2012). Dual impact on the groundwater aquifer in the Kazan Plain (Ankara, Turkey): Sand-gravel mining and over-abstraction. *Environ. Earth Sci.*, 65(1), 241–255.
- Bendixen, M., Best, J., Hackney, C., and Iversen, L. L. (2019). Time is running out for sand. *Nature*, 571(7763), 29–31.
- Bendixen, M., Iversen, L. L., Best, J., Franks, D. M., Hackney, C. R., Latrubesse, E. M., and Tusting, L. S. (2021). Sand, gravel, and UN Sustainable Development Goals: Conflicts, synergies, and pathways forward. *One Earth*, 4(8), 1095–1111.
- Bessière, H., Klinka, T., and Saplairoles, M. (2013). Evaluation de l'impact sur la ressource en eau souterraine de l'exploitation de granulat dans le milieu alluvionnaire de l'Ariège. Technical Report RP-61982-FR, BRGM, Orléans, France, <http://ficheinfoterre.brgm.fr/document/RP-61982-FR>.

- Blanchard, F., Keime, M.-P., and Lansiard, M. (1991). Etude du colmatage et de sa cinétique de formation à l'interface gravière/nappe. Technical Report RR-32248-FR, BRGM, Orléans, France, <http://ficheinfoterre.brgm.fr/document/RR-32248-FR>.
- Botta, S., Comoglio, C., Quaglino, A., and Torchia, A. (2009). Implementation of environmental management systems in the extraction of construction aggregates from gravel pit lakes. *Am. J. Environ. Sci.*, 5(4), 526–535.
- Cheng, X. and Anderson, M. P. (1994). Simulating the influence of lake position on groundwater fluxes. *Water Resour. Res.*, 30(7), 2041–2049.
- Cross, I. D., McGowan, S., Needham, T., and Pointer, C. M. (2014). The effects of hydrological extremes on former gravel pit lake ecology: Management implications. *Fundam. Appl. Limnol.*, 185(1), 71–90.
- Cuthbert, M. O. (2014). Straight thinking about groundwater recession. *Water Resour. Res.*, 50(3), 2407–2424.
- Czernichowski-Lauriol, I. (1998). Etude bibliographique sur l'impact des gravières sur les crues de rivières. Technical Report RR-40022-FR, BRGM, Orléans, France, <http://ficheinfoterre.brgm.fr/document/RR-40022-FR>.
- Darmendrail, D. (1986). Colmatage et effet filtre des berges lors de l'alimentation des nappes alluviales par les cours d'eau (synthèse bibliographique). Technical Report 86-SGN-371-EAU, BRGM, Orléans, France, <http://ficheinfoterre.brgm.fr/document/86-SGN-371-EAU>.
- de Marsily, G. and Tardieu, B. (2018). Stratégie d'utilisation des ressources du sous-sol pour la transition énergétique française : les métaux rares. Technical report, Académie des Sciences, Académie des Technologies, Paris, France, <https://www.academie-sciences.fr/fr/Rapports-ouvrages-avis-et-recommandations-de-l-Academie/utilisation-des-ressources-du-sous-sol-pour-la-transition-energetique.html>.
- Durbec, A. (1986). *Sectorisation des berges des ballastières en eaux : application à l'étude des échanges hydrodynamiques avec la nappe phréatique d'Alsace, au nord de Strasbourg*. PhD thesis, Université de Strasbourg 1, <http://www.theses.fr/1986STR13082>.
- Eberentz, P. and Rinck, G. (1987). Impact qualitatif des carrières en eau sur les nappes souterraines. Rapport de synthèse. Technical Report 87-SGN-567-HNO-ALS, BRGM, Orléans, France, <http://ficheinfoterre.brgm.fr/document/87-SGN-567-HNO-ALS>.
- El-Zehairy, A. A., Lubczynski, M. W., and Gurwin, J. (2018). Interactions of artificial lakes with groundwater applying an integrated MODFLOW solution. *Hydrogeol. J.*, 26(1), 109–132.
- Flipo, N. (2005). *Modélisation intégrée des transferts d'azote dans les aquifères et les rivières : application au bassin du Grand Morin*. PhD thesis, Ecole Nationale des Mines de Paris, <http://www.theses.fr/2005ENMP1300>.
- Fouché, O., Brun, J., and Nasri, B. (2020). Urban lakes: From lack of regard to smart deal? In Gourbesville, P. and Caignaert, G., editors, *Advances in Hydroinformatics*, pages 685–699. Springer Singapore, Singapore.
- Genereux, D. and Bandopadhyay, I. (2001). Numerical investigation of lake bed seepage patterns: Effects of porous medium and lake properties. *J. Hydrol.*, 241(3), 286–303.
- Gibson, J. J. (2002). Short-term evaporation and water budget comparisons in shallow Arctic lakes using non-steady isotope mass balance. *J. Hydrol.*, 264(1), 242–261.
- Gravost, M. (1988). Interactions entre les carrières et les eaux souterraines et superficielles. Bilan des connaissances techniques. Technical Report 87-SGN-391-PAC, BRGM, Orléans, France, <http://ficheinfoterre.brgm.fr/document/87-SGN-391-PAC>.
- Green, J. A., Pavlish, J. A., Merritt, R. G., and Leete, J. L. (2005). Hydraulic Impacts of Quarries and Gravel Pits. Technical report, Minnesota Department of Natural Resources, Division of Waters, Minnesota, USA, [https://www.researchgate.net/profile/Jeffrey-Green-7/publication/349379109\\_Hydraulic\\_Impacts\\_of\\_Quarries\\_and\\_Gravel\\_Pits/links/602d590d4585158939b05b73/Hydraulic-Impacts-of-Quarries-and-Gravel-Pits.pdf](https://www.researchgate.net/profile/Jeffrey-Green-7/publication/349379109_Hydraulic_Impacts_of_Quarries_and_Gravel_Pits/links/602d590d4585158939b05b73/Hydraulic-Impacts-of-Quarries-and-Gravel-Pits.pdf).
- Harbaugh, A. W. (2005). MODFLOW-2005, the US Geological Survey Modular Ground-Water Model: the Ground-Water Flow Process. Technical Report 6-A16, US Department of the Interior, US Geological Survey, Reston, VA, USA, <http://pubs.er.usgs.gov/publication/tm6A16>.
- Hatva, T. (1994). Effect of gravel extraction on groundwater. In *IAHS Publications-Series of Pro-*

- ceedings and Reports-Intern. Assoc. Hydrological Sciences*, volume 222, pages 427–434.
- Hayashi, M. and van der Kamp, G. (2021). Chapter 9 – Water level changes in ponds and lakes: The hydrological processes. In Johnson, E. A. and Miyanishi, K., editors, *Plant Disturbance Ecology*, pages 321–351. Academic Press, Cambridge, MA, USA, 2nd edition.
- Hindák, F. and Hindáková, A. (2003). Diversity of cyanobacteria and algae of urban gravel pit lakes in Bratislava, Slovakia: a survey. *Hydrobiologia*, 506(1), 155–162.
- Hokanson, K. J., Rostron, B. J., Devito, K. J., Hopkinson, C., and Mendoza, C. A. (2022). Landscape controls of surface-water/groundwater interactions on shallow outwash lakes: How the long-term groundwater signal overrides interannual variability due to evaporative effects. *Hydrogeol. J.*, 30(1), 251–264.
- Ioannidou, D., Sonnemann, G., and Suh, S. (2020). Do we have enough natural sand for low-carbon infrastructure? *J. Ind. Ecol.*, 24(5), 1004–1015.
- Jazayeri, A., Werner, A. D., and Cartwright, N. (2021). Partially penetrating lake-aquifer interaction in a laboratory-scale tidal setting. *J. Hydrol.*, 603, article no. 127080.
- Kattner, E., Schwarz, D., and Maier, G. (2000). Eutrophication of gravel pit lakes which are situated in close vicinity to the river Donau: Water and nutrient transport. *Limnologica*, 30(3), 261–270.
- Kidmose, J., Engesgaard, P., Nilsson, B., Laier, T., and Looms, M. C. (2011). Spatial distribution of seepage at a flow-through lake: Lake Hampen, Western Denmark. *Vadose Zone J.*, 10(1), 110–124.
- Koehnken, L. and Rintoul, M. (2018). Impacts of Sand Mining on Ecosystem Structure, Process and Biodiversity in Rivers. Technical report, World Wildlife Fund International, Gland, Switzerland, [https://wwfeu.awsassets.panda.org/downloads/sand\\_mining\\_impacts\\_on\\_world\\_rivers\\_final.pdf](https://wwfeu.awsassets.panda.org/downloads/sand_mining_impacts_on_world_rivers_final.pdf).
- Kondolf, G. M. (1997). Hungry water: Effects of dams and gravel mining on river channels. *Environ. Manage.*, 21(4), 533–551.
- Kuchovský, T., Říčka, A., and Červenková, J. (2008). Impact of gravel pits on ground water: Case study of gravel pits near the Mohelnice City, Czech Republic. In *Proceedings of 10th International Mine Water Assoc Congress, Karlovy Vary, Czech Republic*, pages 69–72. Technical University of Ostrava, Czech Republic, [https://www.imwa.info/docs/imwa\\_2008/IMWA2008\\_069\\_Kuchovsky.pdf](https://www.imwa.info/docs/imwa_2008/IMWA2008_069_Kuchovsky.pdf).
- Labarthe, B. (2016). *Quantification des échanges nappe-rivière au sein de l'hydrosystème Seine par modélisation multi-échelle*. PhD thesis, Université Paris Sciences et Lettres, <http://www.theses.fr/2016PSLEM090>.
- Lu, C., Zhang, B., He, X., Cao, G., Sun, Q., Yan, L., Qin, T., Li, T., and Li, Z. (2021). Simulation of lake-groundwater interaction under steady-state flow. *Groundwater*, 59(1), 90–99.
- Maliva, R. G., Coulibaly, K., Guo, W., and Missimer, T. M. (2010). Simulations of impacts of sand and rock mining on Florida coastal plain water resources. *Mine Water Environ.*, 29(4), 294–300.
- Marsland, P. A. and Hall, D. H. (1989). Gravel extraction and water resources management of the Denge Gravel Aquifer, Kent, England. In *Groundwater Management: Quantity and Quality (Proceedings of the Benidorm Symposium, October 1989)*, IAHS Publication No. 188. IAHS Publ.
- Mason, I. M., Guzkowska, M. A. J., Rapley, C. G., and Street-Perrott, F. A. (1994). The response of lake levels and areas to climatic change. *Clim. Change*, 27(2), 161–197.
- Masse-Dufresne, J., Barbecot, F., Baudron, P., and Gibson, J. (2021). Quantifying floodwater impacts on a lake water budget via volume-dependent transient stable isotope mass balance. *Hydrol. Earth Syst. Sci.*, 25(6), 3731–3757.
- Mazenc, B., Jean, P., and Roussel, P. (1990). Modélisation des impacts hydrauliques des gravières en amont de La Fère. Technical Report RR-30626-FR, BRGM, Orléans, France, <http://ficheinfoterre.brgm.fr/document/RR-30626-FR>.
- Mead, R. (1995). The Direct and Cumulative Effects of Gravel Mining on Groundwater with Thurston County, Washington. Technical report, Environmental Health Division, Groundwater Management Program, Thurston County Public Health and Social Services Department, Olympia, WA, USA.
- Merritt, M. L. and Konikow, L. F. (2000). Documentation of a computer program to simulate lake-aquifer interaction using the MODFLOW groundwater flow model and the MOC3D solute-transport model. Technical Report 4167, US Department of the Interior, US Geological Survey, Reston, VA, USA.
- Michalek, T. (2001). *Quantifying impacts to a floodplain groundwater system from large-scale*

- sand and gravel mining*. Master's thesis, University of Montana, <https://scholarworks.umt.edu/etd/8154>.
- Mimoun, D. (2004). *Spatialisation de l'information : une aide à l'analyse hydraulique et paysagère développée lors de la réhabilitation des sites post-industriels - Cas des réaménagements des gravières en eau en milieu alluvionnaire*. PhD thesis, Ecole Nationale Supérieure des Mines de Saint-Etienne, Université Jean Monnet Saint-Etienne, <http://www.theses.fr/2004EMSE0022>.
- Mollema, P. N. and Antonellini, M. (2016). Water and (bio)chemical cycling in gravel pit lakes: A review and outlook. *Earth-Sci. Rev.*, 159, 247–270.
- Morgan-Jones, M., Bennett, S., and Kinsella, J. V. (1984). The hydrological effects of gravel winning in an area west of London, United Kingdom. *Groundwater*, 22(2), 154–161.
- Muellegger, C., Weilhartner, A., Battin, T. J., and Hofmann, T. (2013). Positive and negative impacts of five Austrian gravel pit lakes on groundwater quality. *Sci. Total Environ.*, 443, 14–23.
- Nadeau, S., Rosa, E., Cloutier, V., Daigneault, R.-A., and Veillette, J. (2015). A GIS-based approach for supporting groundwater protection in eskers: Application to sand and gravel extraction activities in Abitibi-Témiscamingue, Québec, Canada. *J. Hydrol. Reg. Stud.*, 4, 535–549.
- OECD (2019). *Global Material Resources Outlook to 2060: Economic Drivers and Environmental Consequences*. OECD Publishing, Paris, France.
- Otlé, C., Bernus, A., Verbeke, T., Pétrus, K., Yin, Z., Biancamaria, S., Jost, A., Desroches, D., Pottier, C., Perrin, C., de Lavenne, A., Flipo, N., and Rivière, A. (2020). Characterization of SWOT water level errors on Seine reservoirs and La Bassée gravel pits: Impacts on water surface energy budget modeling. *Remote Sens.*, 12(18), article no. 2911.
- Panel, L. (1991). *Modélisation hydrodynamique de l'influence des exploitations de granulats sur la nappe du Perthois (Marne)*. Technical Report RR-32543-FR, BRGM, Orléans, France, <http://ficheinfoterre.brgm.fr/document/RR-32543-FR>.
- Peaudecerf, P. (1975). Effets des gravières sur le comportement hydrodynamique des nappes d'eau souterraines. *La Houille Blanche*, 2–3, 133–140.
- Peckenham, J. M., Thornton, T., and Whalen, B. (2009). Sand and gravel mining: Effects on ground water resources in Hancock county, Maine, USA. *Environ. Geol.*, 56(6), 1103–1114.
- Peduzzi, P. (2014). Sand, rarer than one thinks. *Environ. Dev.*, 11, 208–218.
- Sandercocock, P. and Ladson, A. (2014). *Review of Floodplain Mining Impacts and Risks*. Technical report, Jacobs and Moroka, Naremburn, Australia, [https://www.gbcma.vic.gov.au/downloads/Quarries\\_on\\_the\\_Goulburn\\_Floodplain/Jacobs\\_and\\_Moroka\\_2014\\_Review\\_of\\_floodplain\\_mining\\_and\\_risks.pdf](https://www.gbcma.vic.gov.au/downloads/Quarries_on_the_Goulburn_Floodplain/Jacobs_and_Moroka_2014_Review_of_floodplain_mining_and_risks.pdf).
- Saplaïroles, M., Desprats, J.-E., Delpont, G., and Bourguignon, A. (2007). *Evaluation de l'impact sur la ressource en eaux souterraines de l'exploitation de granulats dans le milieu alluvionnaire de la Garonne (Haute-Garonne)*. Technical Report RP-55673-FR, BRGM, Orléans, France, <http://ficheinfoterre.brgm.fr/document/RP-55673-FR>.
- Schanen, O. (1998). *Analyse et modélisation de l'impact hydrodynamique et biogéochimique des lacs de gravières sur la nappe alluviale du val de Seine*. PhD thesis, Université Pierre et Marie Curie Paris 6, <http://www.theses.fr/1998PA066321>.
- Schanen, O., Bendjoudi, H., Levassor, A., and Fustec, E. (1998). Estimation of groundwater gravel pit lake interface flows by water balance optimization. *C. R. Acad. Sci. Ser. 2 a Sci. Terre Planet./Earth Planet. Sci.*, 326, 107–112.
- Seelen, L. M. S., Teurlincx, S., Bruinsma, J., Huijsmans, T. M. F., van Donk, E., Lürling, M., and de Senerpont Domis, L. N. (2021). The value of novel ecosystems: Disclosing the ecological quality of quarry lakes. *Sci. Total Environ.*, 769, article no. 144294.
- Shrestha, N., Mittelstet, A. R., Gilmore, T. E., Zlotnik, V., and Neale, C. M. (2021). Effects of drought on groundwater-fed lake areas in the Nebraska Sand Hills. *J. Hydrol. Reg. Stud.*, 36, article no. 100877.
- Sinoquet, C. (1987). *Impacts d'une ballastière en eau sur la qualité des eaux souterraines : cas de deux ballastières alsaciennes et modélisation mathématique appliquée aux échanges hydrodynamiques et hydrothermiques*. PhD thesis, Université de Strasbourg 1, <http://www.theses.fr/1987STR13111>.
- Smerdon, B. D., Mendoza, C. A., and Devito, K. J. (2012). The impact of gravel extraction on groundwater dependent wetlands and lakes in the Boreal Plains, Canada. *Environ. Earth Sci.*, 67(5), 1249–1259.



- Søndergaard, M., Lauridsen, T. L., Johansson, L. S., and Jeppesen, E. (2018). Gravel pit lakes in Denmark: Chemical and biological state. *Sci. Total Environ.*, 612, 9–17.
- Sverdrup, H. U., Koca, D., and Schlyter, P. (2017). A simple system dynamics model for the global production rate of sand, gravel, crushed rock and stone, market prices and long-term supply embedded into the WORLD6 model. *BioPhys. Econ. Res. Qual.*, 2(2), 8.
- Vandenbeusch, M. (1975). Nappe alluviale de la basse plaine de la Garonne dans la région toulousaine. Influence gravières-captages. Technical Report 75-SGN-132-MPY, BRGM, Orléans, France, <http://ficheinfoterre.brgm.fr/document/75-SGN-132-MPY>.
- Vucic, J. M., Cohen, R. S., Gray, D. K., Murdoch, A. D., Shuvo, A., and Sharma, S. (2019). Young gravel-pit lakes along Canada's Dempster Highway: How do they compare with natural lakes? *Arct. Antarct. Alp. Res.*, 51(1), 25–39.
- Weilhartner, A., Muellegger, C., Kainz, M., Mathieu, E., Hofmann, T., and Battin, T. J. (2012). Gravel pit lake ecosystems reduce nitrate and phosphate concentrations in the outflowing groundwater. *Sci. Total Environ.*, 420, 222–228.
- Wilson, I. G. (1984). *The effects of gravel extraction on groundwater hydrology*. PhD thesis, University of Oxford, <https://ora.ox.ac.uk/objects/uuid:1147ef12-7cac-4a2c-a63b-7fd54ed5cc11>.
- Winter, T. C. (1976). Numerical Simulation Analysis of the Interaction of Lakes and Ground Water. Technical Report 1001, US Government Printing Office, Washington, DC, USA.
- Wrobel, J. (1980). Wechselbeziehungen zwischen Baggerseen und Grundwasser in gut durchlässigen Schottern. *GWF Wasser/Abwasser*, 121(4), 165–173.
- Zhan, S., Song, C., Wang, J., Sheng, Y., and Quan, J. (2019). A global assessment of terrestrial evapotranspiration increase due to surface water area change. *Earth Future*, 7(3), 266–282.
- Zhang, B., Zheng, X., Zheng, T., Xin, J., Sui, S., and Zhang, D. (2019). The influence of slope collapse on water exchange between a pit lake and a heterogeneous aquifer. *Front. Environ. Sci. Eng.*, 13(2), article no. 20.
- Zhao, G., Li, Y., Zhou, L., and Gao, H. (2022). Evaporative water loss of 1.42 million global lakes. *Nat. Commun.*, 13(1), article no. 3686.





Research article

## Geo-hydrological Data & Models

# An underground view of surface hydrology: what can piezometers tell us about river floods and droughts?

Antoine Pelletier<sup>Ⓢ,\*,a,b</sup> and Vazken Andréassian<sup>Ⓢ,b</sup>

<sup>a</sup> École des Ponts, Marne-la-Vallée, France

<sup>b</sup> Université Paris-Saclay, UR HYCAR, INRAE, Antony, France

*E-mails:* antoine.pelletier@inrae.fr (A. Pelletier), vazken.andreassian@inrae.fr (V. Andréassian)

**Abstract.** The role of aquifers in extreme hydrological events has been highlighted in numerous different contexts, both for river high flows and low flows. Many aquifers are monitored by networks of piezometers that measure groundwater levels: their long-term records allow for a joint analysis of contemporaneous streamflow time series. Yet, groundwater level data are rarely used in surface hydrological modelling approaches, because of the huge complexity of river–aquifer interactions. We propose a simple correlation analysis between yearly extrema of groundwater level and streamflow time series, performed on a large set of 107 catchments and 355 piezometers, located throughout mainland France. We show that it is possible to discriminate between aquifers that correlate only with high-flow events, only with low flows, with both types of events or with none of them. Our typology presents new opportunities to characterise the surface–groundwater exchanges, which can be very important for understanding the occurrence of river extremes.

**Keywords.** Hydrological processes, Groundwater, Hydrological hazards, Aquifer–river interaction, Piezometers.

*Manuscript received 5 April 2022, revised 15 November 2022, accepted 28 November 2022.*

## 1. Introduction

The special IPCC report on climate extremes [Seneviratne et al., 2012] and its 6th assessment report [Ara Begum et al., 2022] highlighted the need to improve our understanding of processes and space–time dynamics of hydrological extremes, to be able to forecast and limit their impacts within the broader challenge of climate change mitigation. Although the role of groundwater in hydrological extremes is well-documented, surface–groundwater interactions are rarely identified well enough for a given watershed to be taken into account while trying to properly understand the hydrological behaviour of a catchment,

and this is particularly true for extremes. Aquifers are often described as the buffering or storage elements of the catchments, able to store water and to release it with a delay [Van Loon, 2015]; in other words, they have the slowest response, among water cycle elements, to climate events, with a filtering effect of the climatic signal [Lo and Famiglietti, 2010, Wossenyeleh et al., 2021]. During droughts, this phenomenon results in aquifers supporting low flows by releasing water: this has been observed in numerous studies in many contexts [Tague and Grant, 2009, Käser and Hunkeler, 2016, Hayashi, 2020, Tobin and Schwartz, 2020, Anderson et al., 2021]. As for floods, they are more rapid phenomena in which the role of aquifers is more ambiguous: the buffering effect can consist in storing excess water and mitigating high

\* Corresponding author.

flows [Chinnasamy, 2016], but Kirchner [2003] highlighted that in many catchments, the flood flows are often mainly composed of groundwater that is mobilised by the excess recharge of aquifers, which was confirmed by local studies [Guérin et al., 2019, Wittenberg et al., 2019]. Flood events can even be caused by a rapid increase of water table levels in aquifers, which is known as *groundwater flooding* [Abboud et al., 2018]. The flood event in the Somme basin in 2000–2001 was concomitant to a rise in the groundwater table of more than 10 m [Pointet et al., 2003], with a large part of flood water carrying the isotopic signature of the chalk aquifer [Négrel and Petelet-Giraud, 2005].

From a hydrological modelling perspective, taking into account the behaviour of aquifers is crucial for simulating and forecasting extremes. But this behaviour is heavily dependent on the local geological context [Eltahir and Yeh, 1999, Soulsby et al., 2006, Bloomfield et al., 2015]. Existing coupled modelling approaches often involve complex distributed models to take into account these idiosyncrasies, with difficult parametrisation and costly computation [see, e.g. Barthel and Banzhaf, 2015, for a review]. To develop simpler approaches that can be easily used in data-scarce or knowledge-scarce contexts, a prior assessment of the nature of the relationship between groundwater and surface water, in terms of hydrological events, must be made. When and where they are available, physical descriptors of the aquifer can provide indicators [Carlier et al., 2018, Wirth et al., 2020]. However, groundwater level time series, measured by piezometer networks, are real-time images of the state of aquifers: along with a minimal analysis of hydrogeological maps, they should be useful in assessing the nature of the river–aquifer interaction. Once the latter is established, these groundwater level time series could help to constrain a hydrological model, thereby possibly improving its performance, its robustness facing new conditions, its ability to represent flood peaks and low-flow recessions and its realism from a process representation point of view.

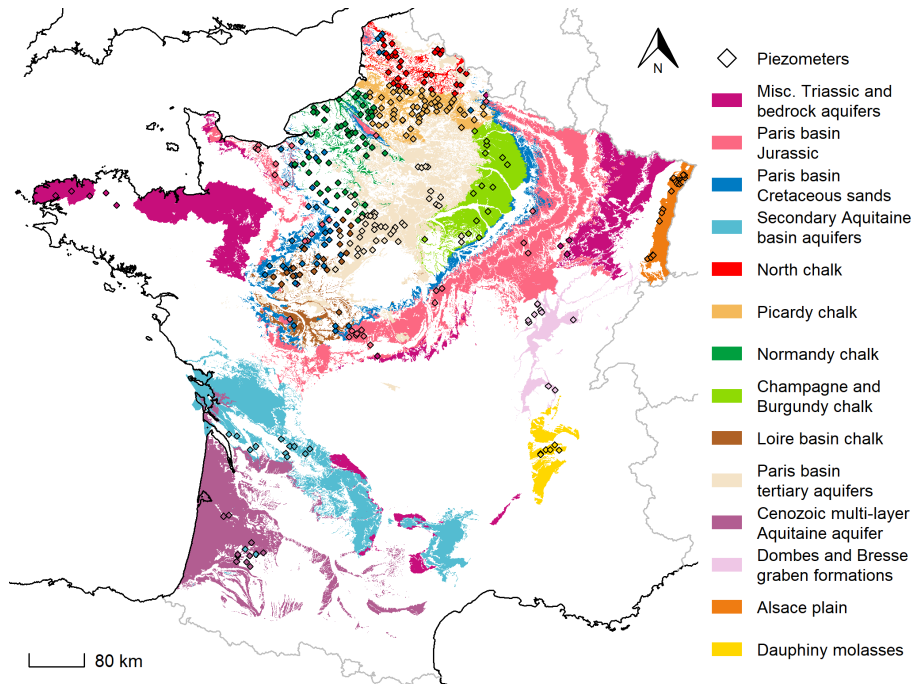
The development of a surface–groundwater joint modelling approach is conditioned by the choice of the right groundwater level measurement points—i.e. piezometers—according to the aim of the modeller. Indeed, to improve streamflow simulation and forecast, the additional information brought to

the model should be relevant. For this work, we created a database of streamflow and groundwater level time series from 107 catchments and 355 piezometers, spread throughout mainland France and we computed annual low-flow and high-flow statistics. Through a simple correlation study, we assessed whether groundwater signal is representative of low flows, high flows or none of them. The dataset was then split into 14 groups, with homogeneous geological characteristics, to obtain synthetic regional information that can be used both by hydrologists and hydrogeologists in their particular context.

## 2. Presentation of the dataset

The French mainland territory hosts a wide diversity of climatic and geological contexts, which induces a large variety of hydrological and hydrogeological configurations and behaviours, as shown in Figure 1. The largest aquifers are located in two major sedimentary basins, the Paris basin and the Aquitaine basin, which host piles of secondary and tertiary formations. The Paris basin Late Cretaceous chalk aquifer and the tertiary limestone Beauce aquifer are among the largest ones in Europe, with major concerns about water use conflicts during droughts or groundwater-induced floods [Pinault et al., 2005, Lalot et al., 2015]. On the outer parts of the Paris basin, smaller aquifers such as Perche Cenomanian sands and Vosges Triassic sandstones are also crucial resources for drinking water supply or biodiversity support in rivers. These regional sedimentary aquifers are generally monitored by a network of long-term piezometers. Alluvial plains also host relevant and monitored aquifers, such as the Alsace alluvium or the Bresse graben gravels. Finally, regions in which the geology is dominated by metamorphic and plutonic formations, such as Brittany, host small local aquifers in fractured bedrock, some of which are monitored and can therefore be taken into account for the present study.

Data were taken from two national databases managed by the *Ministère de la Transition Écologique* (MTE): Banque Hydro [Leleu et al., 2014, Brigode et al., 2020, Delaigue et al., 2021] for streamflow and ADES [Chery and Cattani, 2003, Chery et al., 2008, BRGM, 2021] for groundwater level. Thousands of measurement points are available in these databases,



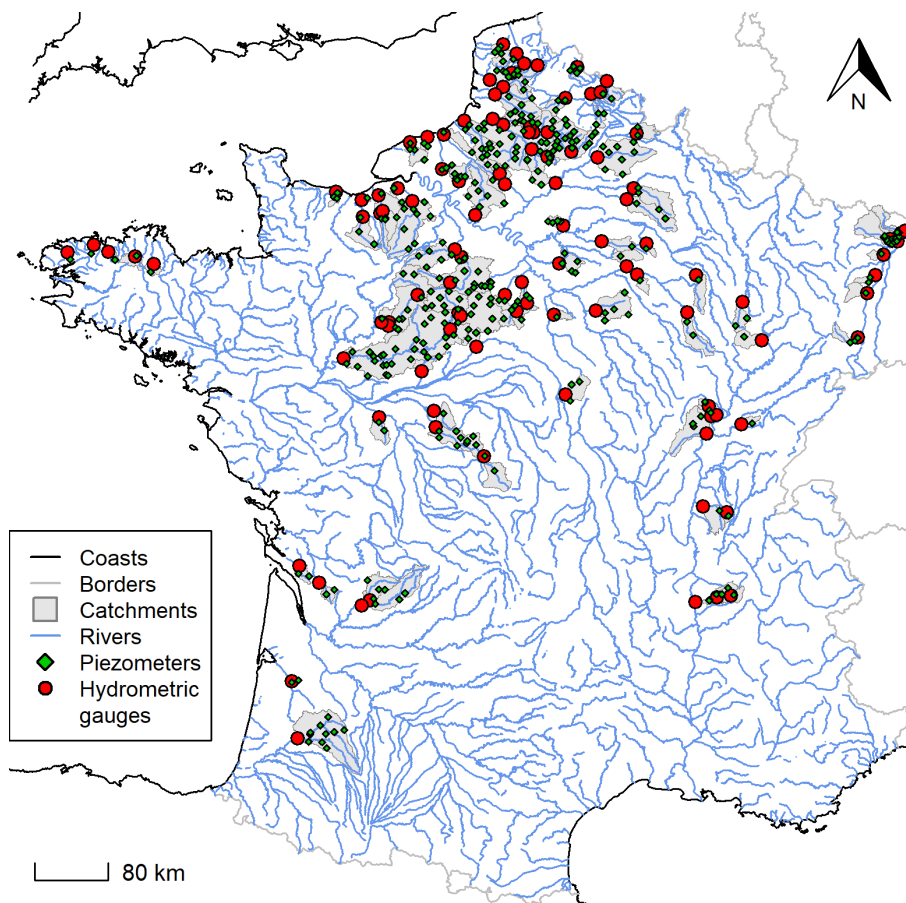
**Figure 1.** The 14 groups of catchment/piezometers pairs with the spatial extents of associated aquifers, according to the BDLISA [Brugeron et al., 2018] national reference map. Only formations regarded as permeable by BDLISA are shown, which can hide some local aquifer formations, especially bedrock formations.

with significant differences in data quality and time extent: an initial selection through visual inspection of time series was made to eliminate points with less than 20 years of available data and affected by anthropogenic influence, such as dams or major withdrawals for catchments and known significant pumping for piezometers—modelling the influence of human activities on catchments is outside the scope of the present study. Catchments in which more than 10% of precipitations fall as snow were also discarded, to avoid taking into account the memory effect of the snow cover.

After this first selection, catchment/piezometer pairs were created. First, each piezometer was assigned to a formation of the French national hydrogeological reference database, BDLISA [Brugeron et al., 2018]. Then, for each catchment, only piezometers whose hydrogeological formation is outcropping within the catchment borders were selected. Along with the criterion of 20 years of available data with less than 10% of missing points, a contemporaneity criterion was added: streamflow and groundwater

level time series must have at least 10 years in common of available data, with less than 10% of missing points. Streamflow data are available at daily time step and groundwater level are available with variable time steps—daily, 10 days or monthly; for the study, all data were aggregated at monthly time steps.

Figure 2 shows the resulting dataset, comprising 107 catchments and 355 piezometers throughout the French territory, with a higher concentration in the Paris basin due to data availability. Each catchment is associated with 1 to 50 piezometers and one piezometer can be associated with 1 to 5 catchments, which results in 456 catchment/piezometer pairs to be evaluated. In a given catchment, several piezometers can monitor the same aquifer—the most relevant piezometer to represent this aquifer must be selected, through criteria like the quality of data or the local hydrogeological context—or different aquifers—the question is, in this case, more about which aquifer is relevant for hydrological modelling in this catchment. Similarly, the same



**Figure 2.** Map of database points: 107 catchments and hydrometric gauges and 355 piezometers.

piezometer can be paired with several catchments, because of nested catchments or because they are located on a watershed.

In order to reach regionally relevant conclusions, 14 groups of catchment/piezometers pairs were created, with homogeneous geological characteristics. Aggregated formations of BDLISA were used as grouping factors, but larger groups had to be created to avoid groups that are too small. In particular, Triassic and bedrock formations, despite having major differences in geological characteristics, were merged as a miscellaneous group of old rocks aquifers. The chalk aquifer of the Paris basin, which represents the majority of catchment/piezometer pairs due to the exceptional availability of monitoring data on its outcropping area, was split into 5 regional groups with homogeneous geological descriptors. The resulting groups are shown in Figure 1.

### 3. Methods

To assess the relationship between streamflow and groundwater level time series from the perspective of drought and flood events, we studied the correlation between yearly indicators of extreme events. We used the annual statistics *QMNA* and *zMNA* for droughts and *QMXA* and *zMXA* for floods [Allier et al., 2019]. *QMNA* is defined, for every year, as the minimum value of monthly means of streamflow; *zMNA* is the same indicator for groundwater level. *QMXA* and *zMXA* represent the maximum values of respective monthly means of the same time series. These statistics are preferred to the yearly minima or maxima of daily values because they are representative of the long-term dynamics of the time series, more than an isolated high-flow or low-flow event. Furthermore, the monthly aggregation makes the indi-

cators less likely to highlight mistaken measurements or abnormal data. As control indicators, yearly mean streamflow  $Q_m$  and groundwater level  $z_m$  were also computed.

However, the calendar year is not the most adapted cut for computing these yearly indicators. Indeed, December and January can host both the end of the drought season and the start of the flood season—see, e.g. January 2018 with both a decadal flood event in the Seine basin and a persisting drought in the upper Loire basin [MTE, 2018]. To compute the flood indicators  $QMNA$  and  $zMNA$ , the classic hydrological year from October 1st to September 30th was used [Linton, 1959], since very few flood events happen at the end of summer. For droughts, a symmetrical *dry* hydrological year was used, from April 1st to March 31st, with the corresponding statement: most French rivers have above-average flows at the end of winter. Yearly means  $Q_m$  and  $z_m$  were computed using calendar years.

After computing these yearly extrema indicators, a correlation study was performed: are low and high groundwater levels correlated, respectively, with low and high streamflow? To assess this, for each catchment/piezometer pair, three values of Spearman correlation [Spearman, 1907] were computed: between  $QMNA$  and  $zMNA$ , between  $QMNA$  and  $zMXA$  and between  $Q_m$  and  $z_m$ . These values were then grouped into the 14 groups of piezometers described in Section 2 to perform a regional analysis. Spearman correlation was preferred to other computations, such as Pearson correlation, since the groundwater level–streamflow relationship is generally non-linear [Eltahir and Yeh, 1999]. To produce maps, for piezometers that are used in several catchments, the maximum value of each indicator was taken among the corresponding catchment/piezometer pairs.

#### 4. Results and discussion

The geographical distribution of results can be seen in Figure 3. It appears that most of the correlation values are high, above 0.7. The correlation between mean streamflow and mean groundwater level is, in most cases, slightly higher than the extrema correlations. Several regions host the lowest values: in the western part of the Paris basin, several points show non-significant negative correlations for means

and droughts, but not for floods. Low values can be also observed in the Saone basin, in the upper Meuse catchment and in the Aquitaine basin. All these regions present a large variability in the correlation between points, which is caused by the overlap of several aquifers monitored by different piezometers: this highlights the necessity of taking into account the geological context and the various properties of the aquifers present in the catchment. In other words, the geographical distribution of the results does not highlight any clear regionally variable trend that could be extended to other neighbouring piezometers or catchments.

This is why the analysis of the geologically homogeneous groups, shown in Figure 4, appears to be more relevant to reaching a more general conclusion. First, the ensemble distribution depicted in grey on the top of the plot shows that the correlation distributions are wide and similar for the three indicators. More than three out of four catchment/piezometer pairs have a correlation above 0.5, with a median of approx. 0.75, while several values below 0 are observed for each indicator distribution. The 14 groups can be classified according to the typology below, in line with their correlation distribution: *high* if most values are above 0.8; *low* if most values are under 0.6; *uncertain* if the distribution is in-between or too wide to fit into these categories.

*Groups with high correlations for droughts and floods:* Dauphiny molasses and Picardy chalk;

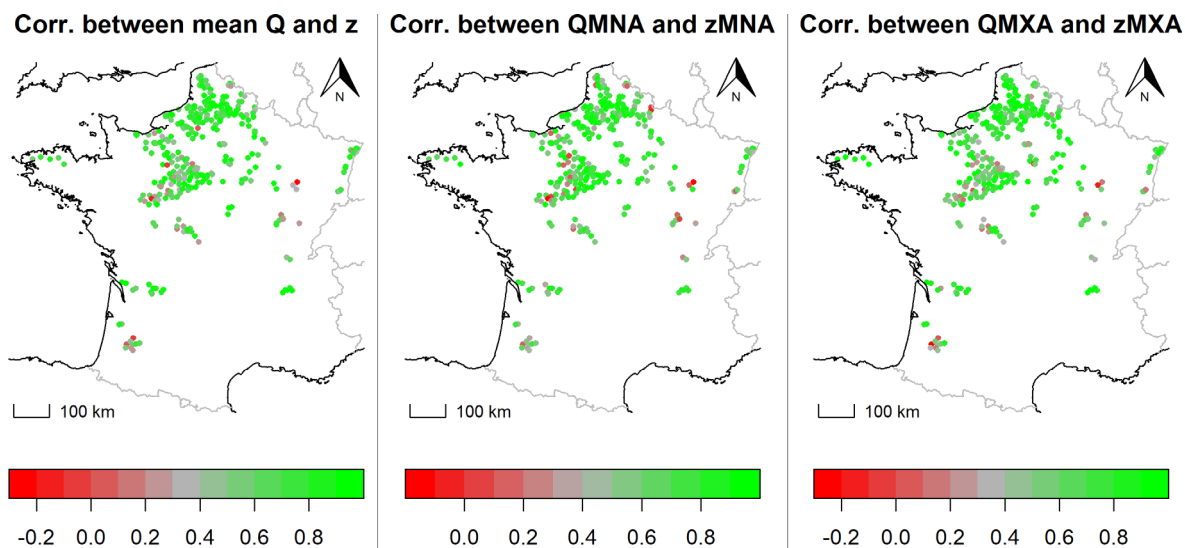
*Groups with high correlations for floods and uncertain behaviour for droughts:* Alsace plain, Champagne and Burgundy chalk, North chalk, Paris basin Jurassic;

*Groups with high correlations for droughts and uncertain behaviour for floods:* Paris basin tertiary aquifers, Normandy chalk, Secondary Aquitaine basin aquifers;

*Group with low correlations for droughts and floods:* Dombes and Bresse graben gravels;

*Groups with uncertain behaviour for both floods and droughts:* Paris basin Cretaceous sands, Triassic and bedrock aquifers, Cenozoic multi-layer Aquitaine aquifer, Loire basin chalk.

These group results can be linked to certain already known properties of aquifers. In particular, the large Paris basin chalk aquifer is known to be



**Figure 3.** Maps of the resulting correlations between droughts, floods and average statistics of streamflow and groundwater level, shown at piezometer locations. For piezometers compared to several catchments, the maximum correlation was conserved.

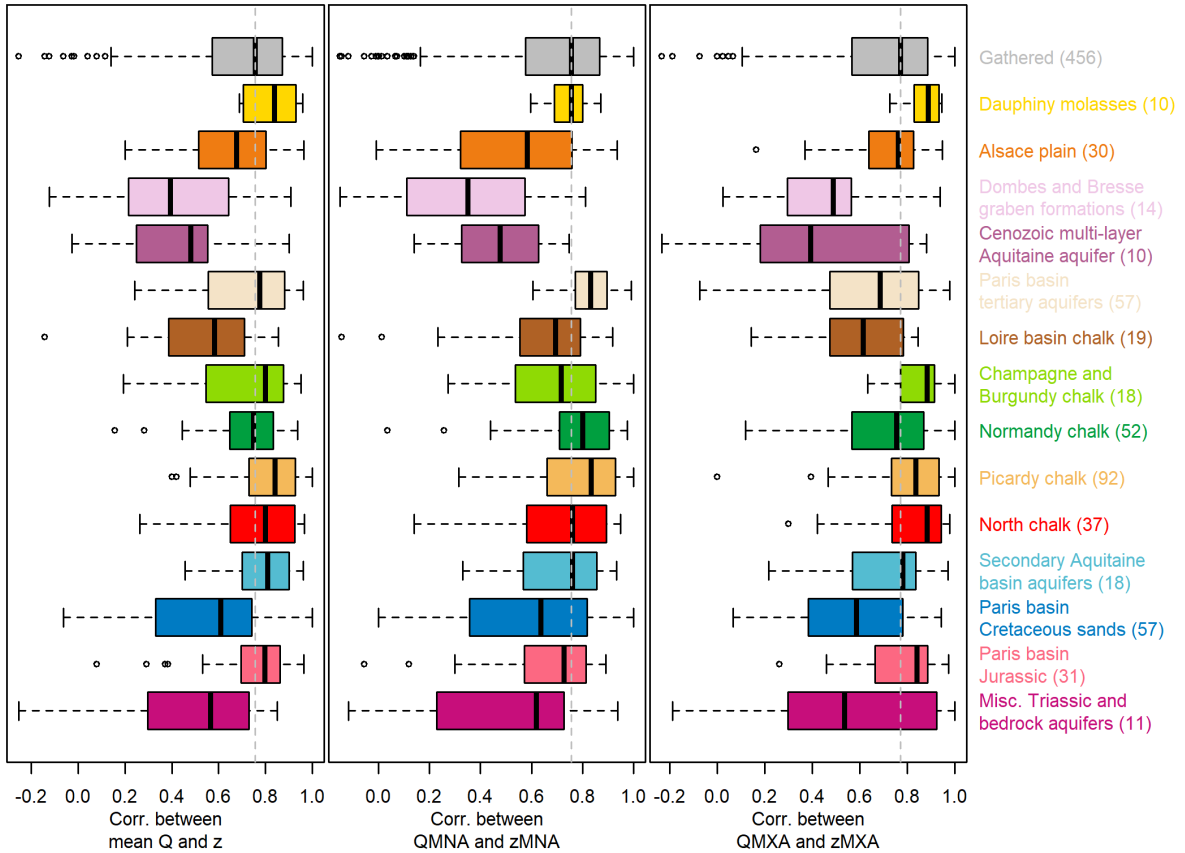
far from homogeneous for its permeability or karstification [Roux and Martin, 2006]: the group analysis highlights this diversity with very different behaviours between the Loire basin and Picardy, for instance. Most of the Dombes and Bresse graben gravels are often covered by less permeable formations, such as silts—see the cartography in Brugeron et al. [2018]—which limits their interactions with rivers, especially in small catchments like the ones studied in the database. The Dauphiny molasses aquifer has an exceptionally high permeability [Dubar et al., 1994], which makes it reach the highest correlation values of the groups studied. The Paris basin Jurassic aquifers are mostly karstified in their outcropping areas [Nicod, 1995], with a quick response to flood events and little ability to store water and support low flows, which explains their better correlation with floods. The Beauce aquifer, which represents the majority of points in the Paris basin tertiary group, has a known role in supporting low flows [Monteil, 2011], which explains the exceptionally high correlation observed for drought indicators, but a more ambiguous role for floods, with a broader distribution correlation for *QMXA* and *zMXA*.

Despite being based on the aquifer division and categorisation in the BDLISA database [Brugeron et al., 2018], the group classification used in the

present study is imperfect, since the group choice was influenced by data availability. It is clear that groups with a too-wide diversity of geological contexts, such as the two Aquitaine basin groups or the miscellaneous Triassic and bedrock aquifers, have the widest distributions of correlation values. No clear trend can be highlighted for these groups, while for more homogeneous ones—such as the regional chalk groups—less equivocal conclusions can be drawn. Data availability also has an influence on the correlation values: even though the dataset creation criterion of 10 years of contemporaneous data between streamflow and groundwater level, cited in Section 2, avoids computing correlations on a too-small number of points, higher values of Spearman correlations are more likely to be reached on 10 points than on longer time series. But this does not seem to be a systematic bias in our analysis, since regions that have been monitored for a long time—Beauce, Burgundy chalk—reach higher correlation values than regions for which data are scarcer—such as the Aquitaine basin or the Bresse graben.

To study its spatial variations, we assumed the river–aquifer relationship to be constant through time. However, this connection is also dependent on local contexts, such as the nature of the river streambed and the related hydrodynamical param-





**Figure 4.** Resulting correlations arranged according to geological and regional group of piezometers. Colours correspond to those used in Figure 1. The grey dashed vertical line represents the median of the ensemble, visible in the *gathered* group on the top.

eters [Baratelli et al., 2016]. Even though we deal with relatively small time extents—in the majority of cases, correlations are computed on less than 20 years—a particular flood or drought event can modify the riverbed properties and, therefore, have an influence on the river–aquifer relationship. The present results can be considered valid as long as no exceptional morphogenic event, likely to break the stationarity hypothesis, is observed in the catchment.

**5. Conclusion**

In this article, we used a national database of streamflow and groundwater level data to assess the relationship between these two physical quantities, through a correlation study between yearly extrema

indicators. Results show that in most cases, the intensity of groundwater level events is well correlated with the magnitude of droughts and floods, with large spatial heterogeneity. Subgroup analysis according to the geological context shows clearer patterns: some regional-scale aquifers have a univocal response to both low-flow and high-flow events, while others respond only to one type of event, and some do not seem to follow a trajectory that can be linked to streamflow events.

Although it is very simple, without any modelling hypothesis, the analysis provides indications as to whether groundwater level data are likely to add useful information to a hydrological model, using a relatively small number of data. Indeed, the implementation of a coupled river/groundwater modelling is generally a complex and substantial task and this kind of prior analysis helps in assessing the relevance

of such an establishment for operational modelling purposes, such as flood and drought forecasting. For instance, the remarkable correlation between low groundwater levels and hydrological drought events in the Beauce region and its tertiary aquifer shows the necessity of including in a model the underground component of the water cycle—see, e.g. Flipo et al. [2012] for an example of coupled modelling. Another interesting case is the chalk aquifer, because of the spatial variability of its behaviour—it was identified as the main habitat of *hydrological monsters* [Le Moine and Andréassian, 2008], i.e. catchments in which streamflow is particularly difficult to simulate and forecast for hydrological models. The chalk area may also benefit from modelling approaches using groundwater level information, but not in all regions. It is also noteworthy that several aquifers, despite being of major importance for regional water resources, exhibit equivocal relationships with surface water—for instance Cenomanian Perche sands and Vosges Triassic sandstones.

The database and data gathering represents the largest part of the present work and we intend to share it as openly as possible to facilitate more thorough analyses. In particular, the correlation computation of yearly indicators hides the lag between groundwater level and streamflow extrema. Preliminary results that are not detailed here show that most piezometers are late with respect to streamflow, i.e. droughts and floods happen before low and high groundwater levels. The database may also be used to conduct a large sample evaluation of groundwater/river coupled modelling approaches, in order to develop generalisable tools for operational hydrological simulations and forecasts. As a final recommendation, we encourage the development of long-term and high-frequency groundwater level measurements in gauged catchments, for surface hydrologists to be able to conduct thorough coupled modelling studies.

### Conflicts of interest

The authors declare no competing financial interest.

### Dedication

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

### Acknowledgments

This work is part of the CIPRHES project, funded by the French national research agency ANR (grant #ANR-20-CE04-0009), and a PhD funded by the French Ministry of Environmental Transition (MTE). We would like to thank Delphine ALLIER (BRGM) for the help in selecting the test dataset and Jean-Baptiste BOISSONNAT (INRAE Antony) and Benoît GÉNOT (INRAE Antony, now at U.R.B.S.) for the database maintenance. We extend our warmest thanks to Dr. Léonard SANTOS (INRAE Antony) and Dr. Paul ROYER-GASPARD for their careful proofreading that considerably improved the article. Last but not least, maps could not have been produced without the precious expertise of Olivier DELAIGUE (INRAE Antony).

### Supplementary data

Supporting information for this article is available on the journal's website under <https://doi.org/10.5802/crgeos.195> or from the author.

### References

- Abboud, J., Ryan, M., and Osborn, G. (2018). Groundwater flooding in a river-connected alluvial aquifer. *J. Flood Risk Manag.*, 11(4), article no. e12334.
- Allier, D., Buscarlet, E., Manceau, J.-C., Klinka, T., and Marly, X. (2019). Sécheresse hydrogéologique : mise en place d'indicateurs piézométriques adaptés à la dynamique des eaux souterraines. In *Sécheresses, étages et déficits en eau*. Paris, France, <https://hal-brgm.archives-ouvertes.fr/hal-02337510>.
- Anderson, T. T., Bestland, E. A., Wallis, I., Kretschmer, P. J. C., Soloninka, L., Banks, E. W., Werner, A. D., Cendón, D. I., Pichler, M. M., and Guan, H. (2021). Catchment-scale groundwater-flow and recharge paradox revealed from base flow analysis during the Australian millennium drought (mt lofty ranges, south Australia). *Hydrogeol. J.*, 29(3), 963–983.
- Ara Begum, R., Lempert, R., Ali, E., Benjaminsen, T., Bernauer, T., Cramer, W., Cui, X., Mach, K., Nagy, G., Stenseth, N., Sukumar, R., and Wester, P. (2022). Point of departure and key concepts. In

- Climate Change 2022: Impacts, Adaptation, and Vulnerability—Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 121–196. Intergovernmental Panel on Climate Change – Cambridge University Press, Cambridge, UK and New York, NY, USA.
- Baratelli, F., Flipo, N., and Moatar, F. (2016). Estimation of stream-aquifer exchanges at regional scale using a distributed model: Sensitivity to in-stream water level fluctuations, riverbed elevation and roughness. *J. Hydrol.*, 542, 686–703.
- Barthel, R. and Banzhaf, S. (2015). Groundwater and surface water interaction at the regional-scale—a review with focus on regional integrated models. *Water Resour. Manage.*, 30(1), 1–32.
- Bloomfield, J. P., Marchant, B. P., Bricker, S. H., and Morgan, R. B. (2015). Regional analysis of groundwater droughts using hydrograph classification. *Hydrol. Earth Syst. Sci.*, 19(10), 4327–4344.
- BRGM (2021). Ades: portail national d'accès aux données sur les eaux souterraines. <https://ades.eaufrance.fr/>. last consulted on 2022-12-07.
- Brigode, P., Génot, B., Lobligois, F., and Delaigue, O. (2020). Summary sheets of watershed-scale hydroclimatic observed data for France. Recherche Data Gouv, V1, last consulted on 2022-12-07.
- Brugeron, A., Paroissien, J., and Tillier, L. (2018). Référentiel hydrogéologique bdlisa version 2 : Principes de construction et évolutions. <http://infoterre.brgm.fr/rapports/RP-67489-FR.pdf>. Rapport final RP-67489-FR, BRGM.
- Carlier, C., Wirth, S. B., Cochand, F., Hunkeler, D., and Brunner, P. (2018). Geology controls streamflow dynamics. *J. Hydrol.*, 566, 756–769.
- Chery, L. and Cattani, A. (2003). ADES : la banque de données nationale sur les eaux souterraines. *La Houille Blanche*, 89(2), 115–119.
- Chery, L., Martin, A., and Legrand, H. (2008). Les systèmes de diffusion nationaux de données: exemple du portail ades. *Tech. Sci. Méthodes*, 1(2), 85–92.
- Chinnasamy, P. (2016). Depleting groundwater—an opportunity for flood storage? a case study from part of the ganges river basin, India. *Hydrol. Res.*, 48(2), 431–441.
- Delaigue, O., Génot, B., Lebecherel, L., Brigode, P., and Bourgin, P.-Y. (2021). Base de données hydroclimatiques observées à l'échelle de la France. <https://webgr.inrae.fr/base-de-donnees>. INRAE, UR HYCAR.
- Dubar, C., Fournier, I., Berthier, M., Beguin, D., Coulet, G., Fiquet, M., and Rieux, J. (1994). Synthèse hydrologique du bassin de Bièvre-Valloire—phase 1. <http://ficheinfoterre.brgm.fr/document/RN-01678-FR>. Report RN-01678-FR, Bureau de recherches géologiques et minières.
- Eltahir, E. A. B. and Yeh, P. J.-F. (1999). On the asymmetric response of aquifer water level to floods and droughts in illinois. *Water Resour. Res.*, 35(4), 1199–1217.
- Flipo, N., Monteil, C., Poulin, M., de Fouquet, C., and Krimissa, M. (2012). Hybrid fitting of a hydrosystem model: Long-term insight into the Beauce aquifer functioning (France). *Water Resour. Res.*, 48(5), article no. W05509.
- Guérin, A., Devauchelle, O., Robert, V., Kitou, T., Dessert, C., Quiquerez, A., Allemand, P., and Lajeunesse, E. (2019). Stream-discharge surges generated by groundwater flow. *Geophys. Res. Lett.*, 46(13), 7447–7455.
- Hayashi, M. (2020). Alpine hydrogeology: The critical role of groundwater in sourcing the headwaters of the world. *Groundwater*, 58(4), 498–510.
- Käser, D. and Hunkeler, D. (2016). Contribution of alluvial groundwater to the outflow of mountainous catchments. *Water Resour. Res.*, 52(2), 680–697.
- Kirchner, J. W. (2003). A double paradox in catchment hydrology and geochemistry. *Hydrol. Process.*, 17(4), 871–874.
- Lalot, E., Curie, F., Wawrzyniak, V., Baratelli, F., Schomburgk, S., Flipo, N., Piegay, H., and Moatar, F. (2015). Quantification of the contribution of the Beauce groundwater aquifer to the discharge of the loire river using thermal infrared satellite imaging. *Hydrol. Earth Syst. Sci.*, 19(11), 4479–4492.
- Le Moine, N. and Andréassian, V. (2008). When hydrological monstrosity stems from a bad initialization of rainfall-runoff simulations. In *The Court of Miracles of Hydrology*. Cemagref, Paris, France, [https://fresno.cemagref.fr/hydromiracles/Download/POSTER\\_LEMOINE.pdf](https://fresno.cemagref.fr/hydromiracles/Download/POSTER_LEMOINE.pdf).
- Leleu, I., Tonnelier, I., Puechberty, R., Gouin, P., Viquendi, I., Cobos, L., Foray, A., Baillon, M., and Ndima, P.-O. (2014). La refonte du système d'information national pour la gestion et la mise à disposition des données hydrométriques. *La*

- Houille Blanche*, 100(1), 25–32.
- Linton, D. L. (1959). River flow in great Britain 1955–56. *Nature*, 183(4663), 714–716.
- Lo, M.-H. and Famiglietti, J. S. (2010). Effect of water table dynamics on land surface hydrologic memory. *J. Geophys. Res.*, 115(D22), article no. D22118.
- Monteil, C. (2011). *Estimation de la contribution des principaux aquifères du bassin versant de la Loire au fonctionnement hydrologique du fleuve à l'étiage*. PhD thesis, École Nationale Supérieure des Mines de Paris.
- MTE (2018). Bulletin national de situation hydrologique du 15 janvier 2018. Technical report, Agence Française pour la Biodiversité—Système d'Information sur l'eau, <http://www.eaufrance.fr/docs/bsh/2018/01>.
- Négre, P. and Petelet-Giraud, E. (2005). Strontium isotopes as tracers of groundwater-induced floods: the somme case study (France). *J. Hydrol.*, 305(1–4), 99–119.
- Nicod, J. (1995). Carte géomorphologique des karsts de France. Notice d'une maquette au 1/3 000 000 environ. *Karstologia*, 25(1), 21–34.
- Pinault, J.-L., Amraoui, N., and Golaz, C. (2005). Groundwater-induced flooding in macropore-dominated hydrological system in the context of climate changes. *Water Resour. Res.*, 41(5), article no. W05001.
- Pointet, T., Amraoui, N., Golaz, C., Mardhel, V., Negrel, P., Pennequin, D., and Pinault, J.-L. (2003). La contribution des eaux souterraines aux crues exceptionnelles de la somme en 2001 observations, hypothèses, modélisation. *Houille Blanche*, 89(6), 112–122.
- Roux, J.-C. and Martin, Y. (2006). *Aquifères et eaux souterraines en France*. BRGM, Orléans.
- Seneviratne, S., Nicholls, N., Easterling, D., Goodess, C., Kanae, S., Kossin, J., Luo, Y., Marengo, J., McInnes, K., Rahimi, M., et al. (2012). *Changes in Climate Extremes and their Impacts on the Natural Physical Environment. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC)*. Intergovernmental Panel on Climate Change, Cambridge, UK and New York, NY, USA.
- Soulsby, C., Tetzlaff, D., Rodgers, P., Dunn, S., and Waldron, S. (2006). Runoff processes, stream water residence times and controlling landscape characteristics in a mesoscale catchment: An initial evaluation. *J. Hydrol.*, 325(1), 197–221.
- Spearman, C. (1907). Demonstration of formulae for true measurement of correlation. *Am. J. Psychol.*, 18(2), 161–169.
- Tague, C. and Grant, G. E. (2009). Groundwater dynamics mediate low-flow response to global warming in snow-dominated alpine regions. *Water Resour. Res.*, 45(7), article no. W07421.
- Tobin, B. W. and Schwartz, B. F. (2020). Quantifying the role of karstic groundwater in a snowmelt-dominated hydrologic system. *Hydrol. Process.*, 34(16), 3439–3447.
- Van Loon, A. F. (2015). Hydrological drought explained. *WIREs Water*, 2(4), 359–392.
- Wirth, S. B., Carlier, C., Cochand, F., Hunkeler, D., and Brunner, P. (2020). Lithological and tectonic control on groundwater contribution to stream discharge during low-flow conditions. *Water*, 12(3), article no. 821.
- Wittenberg, H., Aksoy, H., and Miegel, K. (2019). Fast response of groundwater to heavy rainfall. *J. Hydrol.*, 571, 837–842.
- Wossenyeleh, B. K., Worku, K. A., Verbeiren, B., and Huysmans, M. (2021). Drought propagation and its impact on groundwater hydrology of wetlands: a case study on the doode bemde nature reserve (Belgium). *Nat. Hazards Earth Syst. Sci.*, 21(1), 39–51.



Research article

Geo-hydrological Data & Models

# Dealing with hydrologic data scarcity: the case of the Tindouf basin

Julio Gonçalves<sup>\*,a</sup>, Alexis Nutz<sup>\*,a</sup>, Pierre Séraphin<sup>\*,a</sup>, Amine Chekireb<sup>\*,a</sup>, Lahcen Kabiri<sup>\*,b</sup> and Pierre Deschamps<sup>\*,a</sup>

<sup>a</sup> Aix Marseille Univ, CNRS, IRD, INRAE, Coll France, CEREGE, Aix-en-Provence, France

<sup>b</sup> Département de Géologie, Faculty of Science and Technology of Errachidia (FSTE), Université Moulay Ismail (UMI), Errachidia, Morocco

E-mails: [goncalves@cerge.fr](mailto:goncalves@cerge.fr) (J. Gonçalves), [nutz@cerge.fr](mailto:nutz@cerge.fr) (A. Nutz), [seraphin@cerge.fr](mailto:seraphin@cerge.fr) (P. Séraphin), [chekireb@cerge.fr](mailto:chekireb@cerge.fr) (A. Chekireb), [l.kabiri@fste.umi.ac.ma](mailto:l.kabiri@fste.umi.ac.ma) (L. Kabiri), [deschamps@cerge.fr](mailto:deschamps@cerge.fr) (P. Deschamps)

**Abstract.** This study proposes a first conceptual model for the hydrogeological functioning of the Tindouf Aquifer System (TAS) building upon geomorphological, geological and hydrological remotely sensed and ground observations. Exploratory sketches of the piezometric contours for the main aquifers of the TAS identified (i) recharge areas on the Anti-Atlas relief and the Hamada plateau, (ii) and discharge areas mainly located in the southeastern region where countless salt pans associated with groundwater evaporation were observed. The latter confirms a chiefly endorheic functioning of the TAS. A basin-scale groundwater balance based on gravity data from the Gravity Recovery And Climate Experiment (GRACE) was established to estimate values of the different terms of the hydrogeological water budget of the TAS. A recharge at  $0.8 \pm 0.2 \text{ mm}\cdot\text{yr}^{-1}$ , an evaporation at  $0.9 \pm 0.2 \text{ mm}\cdot\text{yr}^{-1}$  and GRACE-derived groundwater depletion at  $-0.7 \pm 0.1 \text{ mm}\cdot\text{yr}^{-1}$  due to pumping and evaporation in excess of lateral groundwater flow are obtained.

**Keywords.** Groundwater budget, Satellite gravity data, Transboundary aquifers, Phreatic evaporation, Recharge.

*Manuscript received 31 March 2022, revised 21 October 2022 and 27 January 2023, accepted 27 January 2023.*

## 1. Introduction

While elegant methods such as those reported by de Marsily et al. [2005] enable *dealing with spatial heterogeneity*, they are unfortunately of no use when basic hydrological information is missing. This is sometimes the case for some of the world's

major aquifers, especially in regions facing groundwater resources sustainability issues. Large transboundary aquifers have been attracting special attention, and as stated by de Marsily et al. [1978] “*one reason is that their water resources are very large indeed, but another one is that many problems become apparent when one aims at an intensive exploitation*”. Two decades of groundwater resources studies based on satellite hydrology through the Gravity Recovery and Climate Experiment (GRACE) led to a definitive conclusion of world-wide aquifer depletion [see e.g. Richey et al., 2015]. In arid areas such as the Sahara where groundwater is almost the

\* Corresponding author.

only accessible resource, the unsustainable exploitation of large transboundary aquifers is due to a low, sporadic, and therefore difficult to detect recharge which does not compensate the withdrawals. This causes a piezometric decline and therefore major groundwater depletion for Saharan aquifers [Frappart, 2020, Gonçalves *et al.*, 2013]. Four major regional-scale aquifers with surface areas larger than 100,000 km<sup>2</sup> mostly extending within the Saharan belt are concerned: the Northwestern Sahara Aquifer System (NWSAS, Algeria, Tunisia and Libya), the Nubian Aquifer System (NAS, Egypt, Sudan, Chad, Libya), the Murzuk aquifer (Libya) and the Tindouf aquifer system (TAS, see below for basin extension). The first two aquifer systems have been particularly well studied especially to support the well-developed oasis agriculture [Gonçalves *et al.*, 2013, 2020, Heintz and Brinkmann, 1989, Mohamed *et al.*, 2017, Mohamed and Gonçalves, 2021]. The basin of Murzuk has been the subject of some studies especially in the context of the Great Man Made River project to supply water in the Tripoli area [Libya, Pizzi and Sartori, 1984]. The last of these four large Saharan aquifers, the TAS, is undoubtedly the least well-known in terms of groundwater budget, available resources, and overall hydrogeological functioning. Territorial conflicts, the low population and the hyper-arid climate prevailing in the eastern half of the area very likely explain this relative lack of hydrogeological information.

The TAS covers three countries (Algeria, Morocco, Mauritania) and one conflict territory (Western Sahara) over an area of more than 200,000 km<sup>2</sup>. The geological structures as well as the stratigraphy of this basin are relatively well known due to petroleum and mining (iron) economic interests for this basin. Yet, very few and almost only qualitative hydrogeological studies are available [García *et al.*, 2021, Heiß *et al.*, 2020, Lamouroux and Hani, 2006, Mahia *et al.*, 2017]. They are based on geochemical approaches (stable isotopes of water, major element, conductivity) to identify relationships between aquifers and overall groundwater flow directions. Using GRACE data, Frappart [2020] showed temporally stable groundwater resources for the TAS. However, the author used a very approximate geographical limit of the basin which does not correspond to the real geological extension of the TAS. Thus, the global groundwater balance really applying to the TAS re-

mains to be established.

The TAS is considered as an endorheic basin in which outlets correspond to groundwater exfiltration (springs), pumping, and evaporation, at the *sebkhas* (salt pans, see below) of which the largest is that of Tindouf [Figure 3; Lamouroux and Hani, 2006]. If the recharge of aquifers does not compensate for groundwater pumping and evaporation that occurred in *sebkhas*, this should lead to a decrease in groundwater resources and a decline in piezometric levels. Compared to the other large Saharan basins, the TAS is not subjected to excessive pumping. An old estimate of questionable validity now points to groundwater withdrawals in the TAS of about 0.1 km<sup>3</sup>·yr<sup>-1</sup> [UNESCO, 2006], whereas they represent several km<sup>3</sup>·yr<sup>-1</sup> for the other basins.

In arid environments, direct evaporation from the free surface of aquifers can be expected but its intensity is variable and poorly assessed. One hundred km east of Tindouf, in the Beni Abbès region (NWSAS, Western Algeria) where the Average Annual Rainfall (AAR) is 32 mm·yr<sup>-1</sup>, Fontes *et al.* [1986] established experimentally using geochemical profiles (stable isotopes of water and chlorides) the possibility of a net “negative recharge” or more rigorously a diffuse discharge through a thick unsaturated zone. A net negative recharge occurs when infiltration is seasonally or permanently lower than the groundwater evaporation [Lekula and Lubczynski, 2019]. Therefore, this flux corresponds to a groundwater uptake by evaporation for a water table depth up to 200 m according to Kamai and Assouline [2018]. Estimates of direct evaporation from the phreatic surface of aquifers are relatively rare and difficult to perform, either by using geochemical tracers or by direct measurement with surface evaporation flow sensors for point values [Schulz *et al.*, 2015]. For more regional values, mathematical expressions for evaporation as a function of the water table depth were used [Coudrain-Ribstein *et al.*, 1998]. Hydrogeological modeling [Baba-Sy, 2005, Zammouri and Ribeiro, 2017] and remote sensing methods are also useful. For the latter, the satellite measurement of surface temperature is introduced in ground heat balance models in order to identify the contribution of the latent heat of vaporization [Brunner *et al.*, 2008]. However, these models are highly uncertain, particularly in the Saharan region [Weerasinghe *et al.*, 2020]. When this evaporation is very high, it leads to

the development of salt pans at the ground surface in topographical depressions that can be seasonally filled by shallow lakes. In the Maghreb, these areas are referred to as *sebkhas*, the largest of which is the Chott Djerid in the west of Tunisia with an area of 5360 km<sup>2</sup> [Bryant *et al.*, 1994]. In this area, hydrogeological modeling leads to an estimate of the order of 6 m<sup>3</sup>·s<sup>-1</sup> (or 35 mm·yr<sup>-1</sup>) of evaporation flux [Baba-Sy, 2005]. While such diffuse discharge was identified at the local scale, the Tindouf basin offers the opportunity to describe this flux at a regional scale.

In this study, the scarce piezometric data were supplemented with some geomorphological and hydrological features specific to the arid domain and indicative of piezometric levels. This enabled us first to propose a conceptual model for the hydrodynamic functioning of the basin and then to establish a first set of water level contour sketches for the multi-layer aquifers of Tindouf basin, not available hitherto. The regional groundwater budget over the TAS domain was also obtained using GRACE data and compared to the conceptual hydrodynamic model and to the piezometric analysis. This satellite information (gravity and geographic distribution of *sebkhas*) made it possible to propose an estimate for the unknown direct phreatic surface evaporation through a groundwater mass balance.

## 2. The Tindouf basin: physical context

### 2.1. Geomorphology

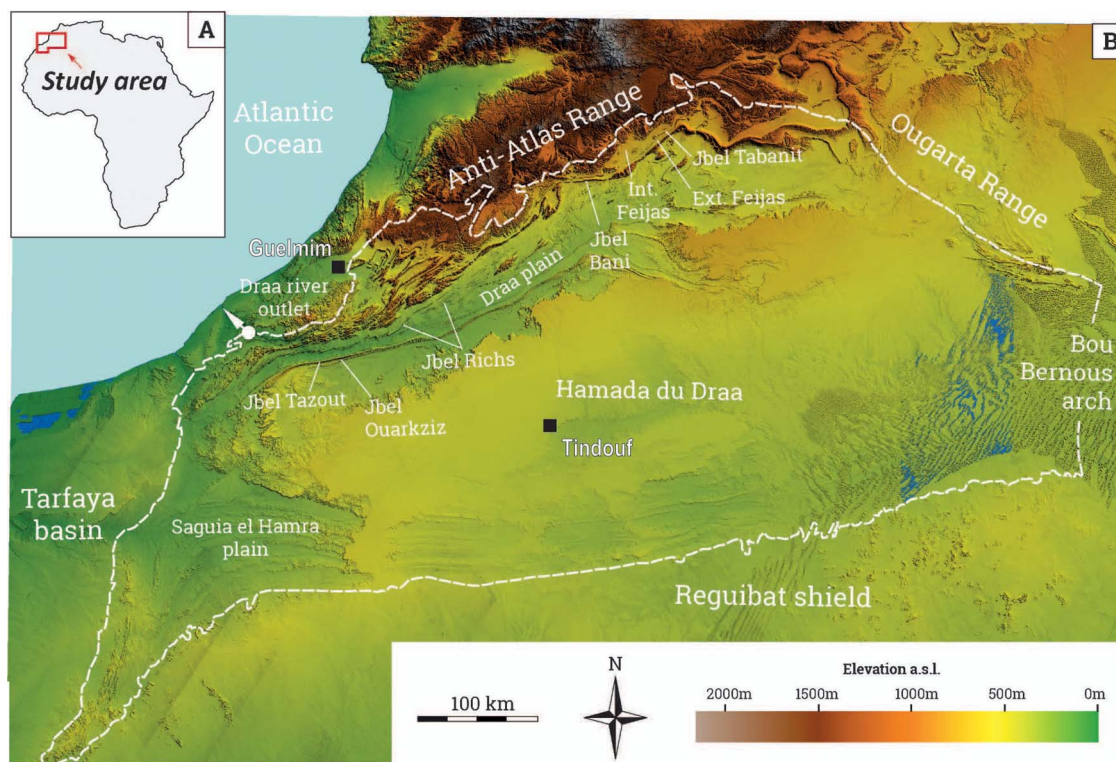
The Tindouf basin coincides with a 800 km long and 350 km wide NE-SW oriented low-relief area in North-West Africa (Figure 1A). It is bounded by the Anti-Atlas and the Ougarta ranges to the North, by the Tarfaya basin to the West, by the Bou Bernous arch to the East and by the Reguibat Shield to the South (Figure 1B). The Anti-Atlas range forms an up to 2000 m high relief laterally connected to the lower relief of the Ougarta range. The northern part of the Tindouf basin shows successive southward-dipping *cuestas*. They originate from the differential erosion of clay-dominated Middle Cambrian, Lower Ordovician, Silurian and Late Devonian formations forming elongated depressions referred to as the internal and external Feijas [Nutz *et al.*, 2019, Thorp *et al.*, 2002] and the Drâa Plain, respectively, and more resistant sandstone- and limestone-dominated Lower Cambrian, upper Middle Cambrian, Upper Ordovician,

Early Devonian and upper Early Carboniferous rocks [Michard *et al.*, 2008]. More resistant rocks form the relief of the Jbels Tabanit, Bani, Richs, Tazout and Ouarkiz (Figure 1B). *Cuestas* and anticlines are crosscut by *wadis* in narrow gorges (locally called *foums*) that progressively connect to form the Drâa River flowing in the Drâa plain, which corresponds to the valley located between the Jbel Bani and the Jbel Ouarkiz (Figure 1B). The Drâa valley is the lowest area of the Tindouf basin. In the central portion of the Tindouf basin, a slightly concave plateau referred to as the Hamada du Drâa slowly dips southeastward. On top of the Hamada du Drâa lie abundant *sebkhas*, especially in the southeastern and eastern part. Eastward and southward, transitions from the Tindouf basin to both the Bou Bernous arch and the Reguibat shield are not expressed by a significant topography. To the west, from the Hamada du Drâa to the Tarfaya basin, the topography abruptly decreases at the edge of the Hamada du Drâa plateau entering the catchment of the Saguia el Hamra wadi (Saguia el Hamra plain) and then progressively decreases westward until the Tarfaya basin.

### 2.2. History of the Tindouf basin and lithostratigraphy

Initiation of the Tindouf basin dates back to Paleozoic times, during which a large wave-length flexural deformation has affected the Saharan platform. This deformation resulted in several intracratonic basins separated by anticlines, referred to as arches [Peron, 2019]. The Tindouf basin is the western-most of these intracratonic basins; it is a 800 km elongated asymmetrical syncline (Figure 2A) oriented ENE to WSW characterized by a southern flank gently dipping to the north and a northern flank steeply dipping to the south (Figure 2B). In places, the Tindouf basin is thought to reach a maximum depth of about 8 km [Craig *et al.*, 2008] while it commonly reaches 2–3 km (Figure 2B).

The lithostratigraphy consists of km-scale Paleozoic formations from Cambrian to Late Carboniferous. Unconformably overlaying, hundreds of meters-scale Cretaceous, Tertiary and Quaternary deposits form the Hamadian Formation. All together, these formations constitute the Tindouf Aquifer System



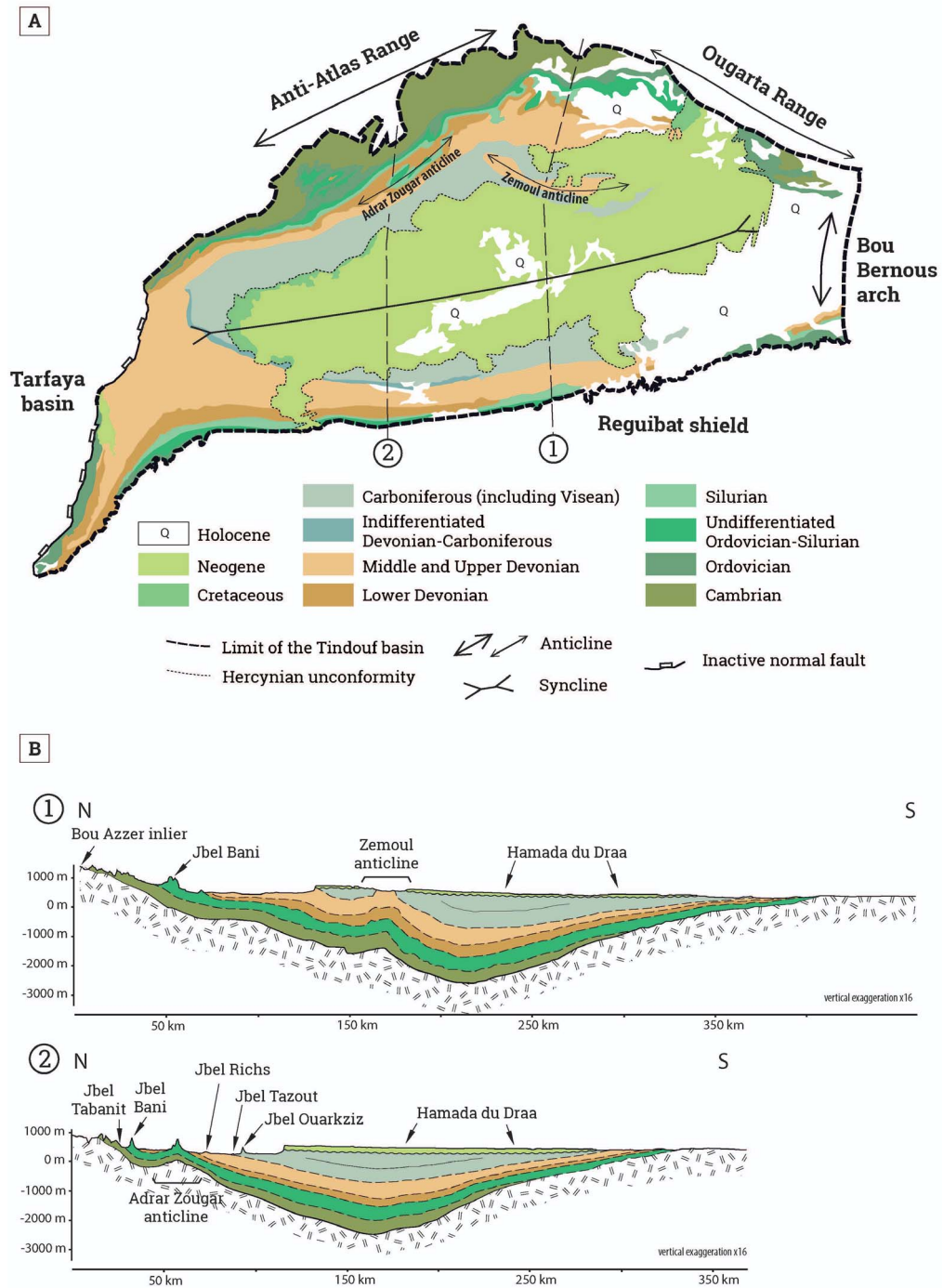
**Figure 1.** Digital Elevation Model of northwest Africa (SRTM30) showing the Tindouf basin and surrounding areas. Estimated limits of the basin are presented as well as major sub-regions of the basin.

(TAS). The Paleozoic sedimentary succession originates from the long-term evolution of a several hundreds of km long embayment connected northward to the Tethyan sea where global eustatic fluctuations led to successive transgressions and regressions from the Cambrian to the Late Carboniferous Guiraud *et al.* [2005]. From the Late Carboniferous, uplift of the northern Sahara platform, as a consequence of the Hercynian orogeny, marks the onset of the Anti-Atlas and the Ougarta range uplift. As a consequence, the Tindouf basin was enclosed and then inverted, interrupting sedimentation. The area likely remained stable in a terrestrial context until the Lower Cretaceous, during which the Sahara Platform and the Tindouf basin were intermittently flooded until the Paleogene, the maximum submersion being coeval to the global Cenomanian–Turonian transgression. During the Middle to Upper Eocene, the first Atlasic tectonic event [*sensu* Frizon de Lamotte *et al.*, 2009] slightly uplifted the basin and sedimentation stopped, until another subsidence phase during the Neogene, likely

during the Mio-Pliocene, allowed sedimentation to restart. Afterwards, the Tindouf basin was definitely inverted [Frizon de Lamotte *et al.*, 2009]; since that time, eolian deflation and fluvial incision have been the predominant processes modifying the Tindouf basin surface.

Even if the stratigraphic pile of the Tindouf basin is complex due to strong variations in lateral facies and thickness, as a simplification, from base to top, the Tindouf basin succession starts with Cambrian limestones [Maloof *et al.*, 2010]. In places, the Cambrian limestones overlay a basal clastic interval [Michard *et al.*, 2008]. In the northern part of the Tindouf basin and in the Anti-Atlas, the Cambrian interval is several kilometers thick, progressively thinning southward, pinching out in depth in the Tindouf basin (Figure 2B). Above, Lower Ordovician rocks consist of shales recording offshore marine conditions [Loi *et al.*, 2010]. Overlying are sandstones of the Upper Ordovician; they originate from a major global eustatic sea-level fall due to the end-Ordovician





**Figure 2.** Simplified geological map of the Tindouf basin and regional-scale cross sections showing the TAS geometry [compiled from Boote *et al.*, 1998, Leprêtre, 2015, Perron, 2019].

glaciations [Beuf *et al.*, 1971, Ghienne *et al.*, 2007, Loi *et al.*, 2010]. Interruption of this glacial event is recorded by an abrupt transition to Silurian shales [Lüning *et al.*, 2000, Nutz *et al.*, 2013]; in the Tindouf basin, Silurian shales are 100–200 m thick [Guerrak, 1989]. Overlying the Silurian shales, lower to middle Devonian rocks are shallow marine shales while during the Middle to Upper Devonian a global marine regression led to the deposition of marine fine sandstones [Guerrak, 1989]. Above, Lower Carboniferous rocks (Visean) include shallow marine limestones and occasional continental sandstones. These deposits are overlain by marine shales which are Middle Carboniferous in age. Above, continental sandstones terminate the Paleozoic succession as a new global regression occurred during the Upper Carboniferous. Overlying the Paleozoic succession, the Hamadian package includes Lower Cretaceous, Cenomanian and Turonian marine and continental deposits, unconformably overlain by Mio-Pliocene continental sandstones.

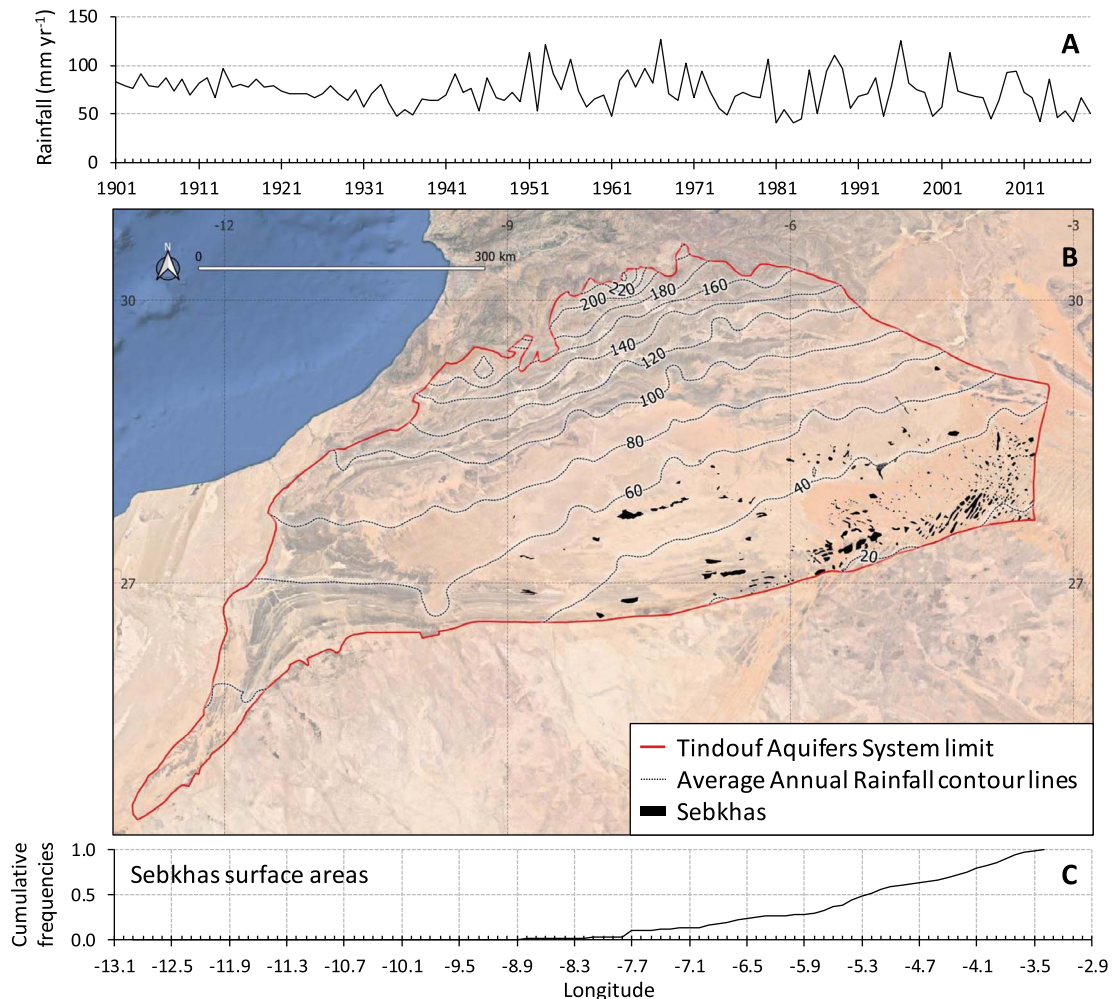
### 2.3. Hydrology

Using the geomorphological, geological and structural information presented in Sections 2.1 and 2.2, the hydrogeological limits for the TAS (contacts with the basement, hydraulic thresholds associated with folding) shown in Figures 1 and 2 were identified in a much more realistic way than those presented in IGRAC (International Groundwater Resources Assessment Centre) [2015]. Over this domain of 255,000 km<sup>2</sup> in surface area, four aquifer formations were identified: the Upper Ordovician sandstones, the Middle to Upper Devonian fine sandstones, the Carboniferous mainly made of the Visean limestones and sandstones, and the Hamadian platform (Cretaceous, Tertiary and Quaternary). The syncline morphology of the Tindouf basin suggests a topographically driven groundwater flow regime which is clear for the Paleozoic aquifers affected by major folding and more subtle for the Hamadian, where lower stratal dips are observed.

Rainfall gridded data beginning in 1901 were retrieved from the CRU data center (Climatic Research Unit, <http://wps-web1.ceda.ac.uk>) at a resolution of 0.5° × 0.5°. It should be noted, however, that the number of rainfall land stations (some of which were available after 1901) is very limited in the Sahara [Harris *et al.*, 2020] and specifically in the TAS area

(7 stations). Precipitation shows a NW–SE linear gradient with maximum precipitation onto the Anti-Atlas relief with an average annual rainfall reaching 200 mm·yr<sup>-1</sup>, while minimum precipitation locates to the southeastern part of the basin with an average value of 20 mm·yr<sup>-1</sup> (Figure 3). This rainfall pattern with higher values in the Anti-Atlas greatly influences the hydrographic network. Only one river of regional importance is present in the TAS domain but only in its lower and mainly dry part, the Drâa river (Figure 1). In its upper and middle valley, the Drâa river episodically flows mainly on the bedrock and not on the sedimentary formations of the TAS. In its upstream part, flood waters are stored in dams to supplement groundwater for human activities, especially agriculture. Moreover, the larger average annual rainfall (AAR) values located on the Anti-Atlas localize the likely active recharge area. In opposition, *sebkhas* mostly situated in the south-east of the basin (Figure 3) correspond to areas of high water evaporation (groundwater and/or wadi floods) and thus certainly localize some of the main outlets of the TAS. However, the real hydrodynamic functioning (recharge–discharge areas) of the Tindouf basin is difficult to decipher due to the scarcity of the piezometric data.

The few available water levels (31 values) mostly date back to 1980 and are localized around the city of Tindouf for the Hamadian and Visean aquifers, and south of Tindouf city for Devonian and Cambro-Ordovician aquifers [Lamouroux and Hani, 2006]. However, a few additional piezometric levels measured in the 2010–2020 decade were provided by Ron Martin [2020]. By comparing these recent levels with the 1980 data, the temporal evolution of aquifer water levels was estimated, at least in the western part. A decline of a few meters per decade for the Hamadian aquifer is observed while a relative stability of piezometric levels characterizes other aquifers. According to Lamouroux and Hani [2006], the Hamadian and Visean aquifers are hydraulically connected at the southern border of the basin, as evidenced by very similar piezometric levels at about 400 m in dedicated piezometers for the two aquifers in the area of Tindouf city. The largest piezometric level data set for the Hamadian aquifer demonstrates a groundwater flow from the west-northwest (Hamada plateau) towards the *sebkha* of Tindouf, which is therefore a major discharge area by exfiltra-



**Figure 3.** (A) Domain-averaged time series of the Average Annual Rainfall (AAR) for the TAS, (B) distribution of the AAR and location of the sebkhas (salt pans), and (C) Cumulative distribution of the sebkhas surface areas.

tion and evaporation. For the Visean aquifer, Lamouroux and Hani [2006] reported unpublished (unavailable technical reports) hydraulic head measurement campaigns suggesting north-eastern streamlines converging to the sebkha of Tindouf and therefore pointing to an active recharge at the northern Anti-Atlas piedmont. This latter recharge area is consistent with the fact that in the northern part of the basin, the Zemoul anticline (Figure 2) exhumes the Devonian and the Visean, thus establishing a preferential recharge zone. For the aquifers of the Hamadian and Visean, recharge probably occurs by direct infiltration (diffuse recharge) of rainfall and infiltration of floodwater (focused recharge) on the

Hamada plateau and at the Zemoul anticline while discharge occurs at the southern sebkhas.

Along the southern flank of the Tindouf basin, in the Devonian aquifer, the piezometric data of Lamouroux and Hani [2006] point to a down-dip (towards the northeast) local hydraulic head gradient of  $10^{-3}$ . Despite scarce and poorly distributed data (three water level values), a similar gradient can be identified for the Cambro-Ordovician aquifer at the same location. No piezometric data are available in the literature to our knowledge for the Paleozoic formations on the northern flank of the Tindouf basin. Given the lack of an effective hydrodynamic conceptual model for the TAS, little is known about

recharge and discharge areas for the Devonian and Cambro-Ordovician aquifers which can only be postulated regarding the overall topographically-driven conceptual model for groundwater flow. Higher piezometric levels in the north of the basin compared to the south can be reasonably assumed, resulting in a southward gradient. Therefore, the general hydrodynamic functioning for the Devonian and Cambro-Ordovician aquifers could correspond at first order to a recharge on the northwest flank and a discharge on the southern flank. Local discharge areas can also be suspected. For instance, the absolute topographic low area of the basin corresponds to the Drâa valley to the northwest of the domain which deeply incised transversally the Devonian and Cambro-Ordovician strata making this area a possible outlet for these aquifers.

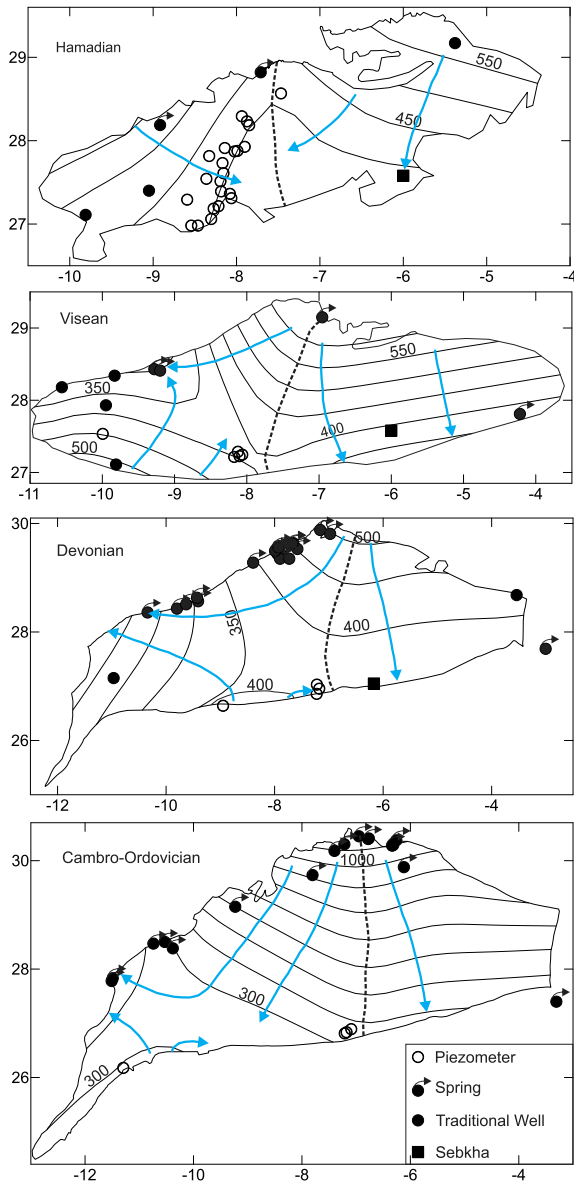
Average transmissivity values of  $6.5 \times 10^{-6}$ ,  $1 \times 10^{-4}$ ,  $9 \times 10^{-5}$  and  $6 \times 10^{-4} \text{ m}^2 \cdot \text{s}^{-1}$  were reported by Mahia *et al.* [2017] and García *et al.* [2021] for the Hamadian, Visean, Devonian, and Cambro-Ordovician aquifers, respectively. Storativity values are only available for the confined parts of the Devonian in the order of  $10^{-3}$  and for the unconfined areas of the Hamadian and Visean aquifers at 0.06 and 0.1 respectively. Considering the hydraulic head gradients suggested by the scarce piezometric data in the order of  $10^{-3}$ , the transmissivity values point to a low productivity of the aquifers. Groundwater exploitation mainly occurs in the Morocco fraction of the TAS with more than  $100 \times 10^6 \text{ m}^3 \cdot \text{yr}^{-1}$ , much larger than the  $3 \times 10^6 \text{ m}^3 \cdot \text{yr}^{-1}$  withdrawn at Algerian Tindouf city [UNESCO, 2006]. However, beyond these rather old estimates, pumping data are extremely scarce. Only global values for the Moroccan basins are available [Hssaisoune *et al.*, 2020] with an estimate of  $323 \times 10^6 \text{ m}^3 \cdot \text{yr}^{-1}$  for the Drâa valley, of which only the lower part is in the TAS.

### 3. Material and methods

#### 3.1. Water levels mapping

The piezometric contour map, which is the basic tool for identifying groundwater flow, is lacking here for almost all the aquifers due to water level data scarcity, with the exception of a very restricted area around the city of Tindouf for the Hamadian formation. However, some indirect information makes

it possible to compensate for the lack of hydraulic head data. Geomorphological and hydrological features indicative of the piezometric level in this arid context can be comprehensively collected. The locations of springs were primarily considered here since their topographical elevation gives the piezometric levels of the aquifers from which such groundwater outflows. The corresponding aquifer was simply identified by considering the exposed geological strata at the location of the springs. Additional information was parsimoniously used to supplement the spring data, i.e. only to fill sampling gaps, especially in the eastern and western parts of the basin. This additional information is made of traditional or shallow water wells (hassi in local arabic) and sebkhas. Water levels very close to the ground surface for hand dug traditional wells mean that a piezometric level almost equal to the topographic elevation can be assumed. Elevations are obtained here using the Digital Elevation Model (DEM) SRTM (1-arcsec resolution). The Khetarra systems, which are gently dipping tunnels drilled to reach the water table of an aquifer (indicative of the elevation of the water table) provided by Heiß *et al.* [2020], were not used here since they are redundant with numerous springs. Spring and well locations and names are available from the GEOnet Names Server (GNS) of the US National Geospatial Intelligence Agency (NGA, <https://geonames.nga.mil/gns/html>). The sebkhas are flat depressions collecting water from wadi floods and/or aquifers where water evaporates, causing salt precipitation. At the sebkhas, and especially where they are exclusively fed by groundwater, the water table is nearly at the ground surface [Lamouroux and Hani, 2006]. It is assumed here that the small surface water bodies observed on the hyper-arid southeastern and eastern parts of the basin are sebkhas exclusively fed by aquifers whose free surface at these outcrops is very close to the ground surface. The few sebkhas (two) used here were obtained from the surface water bodies mapping provided by Pekel *et al.* [2016, data available at <https://global-surface-water.appspot.com>] using satellite imagery. The indirect water level information collected here was associated with the available piezometric data (Figure 4 and Supplementary Information) in an attempt to improve our understanding of the overall hydrodynamic functioning of the aquifers system.

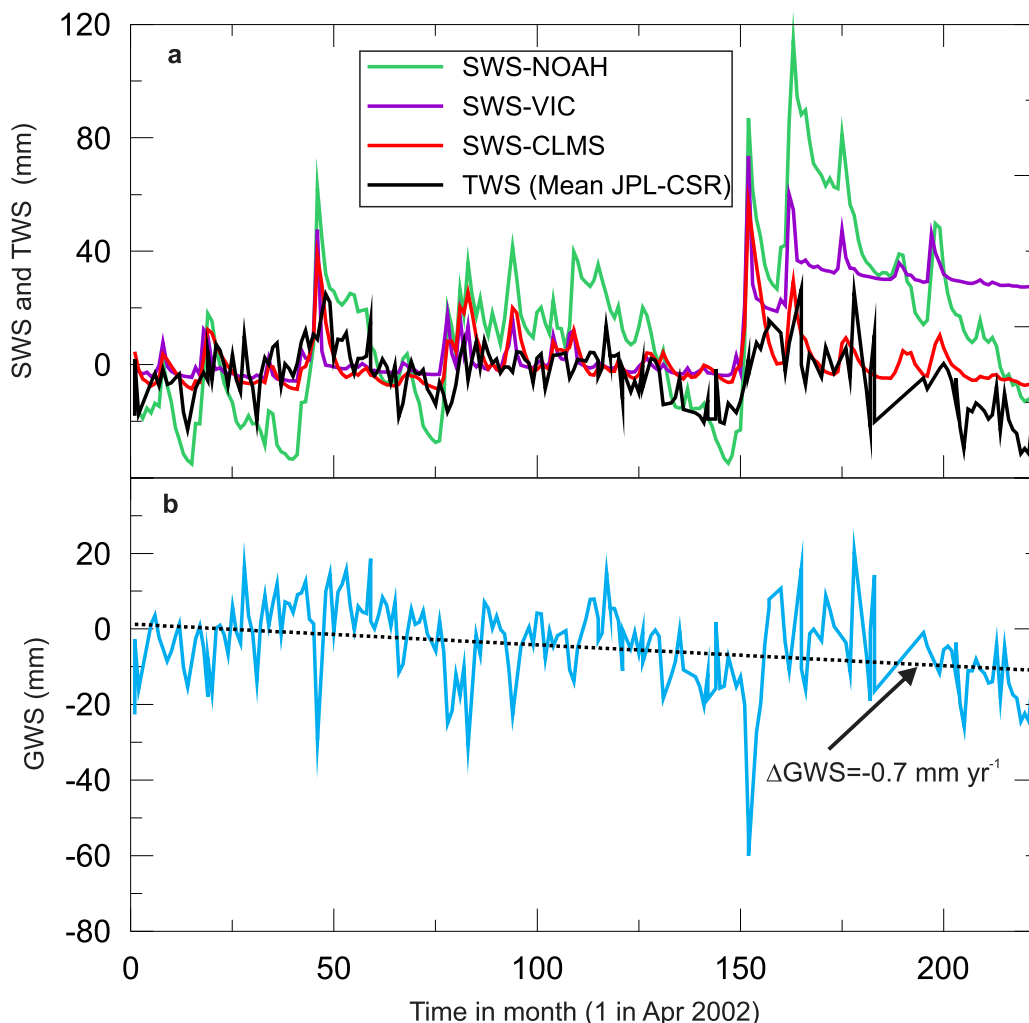


**Figure 4.** Piezometric sketches for the four main aquifers. The blue arrows indicate main groundwater streamlines and the black dashed line groundwater divides (Cambro-Ordovician, Devonian, and Visean) or convergence to a median discharge area (Hamadian). Springs beyond the limits (e.g. groundwater divide at Bou Bernous) were considered to constrain the interpolation.

### 3.2. Groundwater mass balance analysis using satellite-based gravity data

The GRACE twin-satellite system launched in 2002 by the NASA and the German Aerospace Center (DLR) monitors the earth's gravity on a monthly basis [Tapley *et al.*, 2004]. The gravity variations monitored by the GRACE satellites are attributed to mass variations of the surface water bodies (groundwater, soil water, and surface water). Given the obvious hydrological interest of this kind of measurement, for example for water balance purposes at the regional scale, the data are expressed in water height. The more recent version of the GRACE data is based on an interpretation of the gravity signal using discrete regional mass concentration blocks (spherical caps) called Mascons [Watkins *et al.*, 2015]. Monthly GRACE-derived Terrestrial Water Storage (*TWS*) anomalies (difference between a *TWS* value and its average value calculated for the period January 2004–December 2009) were retrieved from the Jet Propulsion Laboratory (JPL data center, <http://grace.jpl.nasa.gov>). The JPL R06M solution filtered from the contribution of oceans [CRI filter, Wiese *et al.*, 2016], therefore giving the *TWS* time-series corresponding only to the land part of a Mascon, was used here. The scaling factors proposed by Wiese *et al.* [2016], which correspond in fact to downscaling factors [Scanlon *et al.*, 2016], were used to downscale the data from their native resolution ( $3^\circ \times 3^\circ$ ) to a  $1^\circ \times 1^\circ$  resolution. These downscaling factors are based on the mass distribution calculated by a land surface model (soil-atmosphere water balance) which is considered suitable to the Tindouf basin where evaporation through the soil might be a prominent discharge. The relevance of these factors in the case where groundwater abstraction is the dominant term was recently questioned by Gonçalves *et al.* [2021]. The Center for Space Research (CSR) solution produced at University of Texas (Austin, USA) and involving inversion of the gravity signal at a  $1^\circ$  resolution was also retrieved. In this study, 223 monthly solutions for the period between April 2002 and October 2020 were used (Figure 5). The trend of a *TWS* time-series is interpreted as water variation (e.g. in  $\text{mm}\cdot\text{yr}^{-1}$ ) over a land surface and thus water fluxes over the TAS.

As outlined above, groundwater, surface water, and soil moisture variations are accounted for in *TWS*



**Figure 5.** (a) Domain-averaged time-series for the different Soil Water Storage (SWS) models GRACE (mean CSR-JPL) from GLDAS (NOAH, VIC, CLSM) and average Terrestrial Water Storage (TWS) between JPL and CSR solutions., and (b) Groundwater Storage (GWS) times series obtained by subtracting the retained SWS (CLMS) from the averaged TWS.

time series. Therefore, and in the absence of substantial surface water in the TAS domain, Soil Water Storage (SWS) has to be subtracted from TWS to identify Groundwater Storage (GWS). Soil water storage variations ( $\Delta SWS$ ) are usually small due to almost constant and low soil moisture in arid areas and may represent a small or even negligible component of TWS [Gonçalves *et al.*, 2013, Nimmo *et al.*, 1994]. Despite the questionable geographical limit used, the soil moisture time-series proposed by Frappart [2020] for the TAS suggest an almost zero SWS interannual variation. Soil Water Storage (SWS) time-series are

available from the Global Land Data Assimilation System (GLDAS, <http://disc.sci.gsfc.nasa.gov/>) combining ground-based and satellite data and the results of hydrological surface models [Rodell *et al.*, 2004]. The  $1^\circ$  spatial resolution data was retrieved for the same period as the GRACE data and for the 3 available land surface models (LSM):

- the Variable Infiltration Capacity model (VIC)
- the Community Land Surface Model (CLSM)

- the National Oceanic and Atmospheric administration model (NOAH).

The Soil Water Storage (SWS) was expressed in anomalies in the same way as for the GRACE TWS anomalies, i.e. by subtracting the 2004–2009 average from each term of the time series. Upon verification of the consistency of the alternative LSM outputs with the GRACE data, GLDAS data yielded  $GWS = TWS - SWS$  time-series. The groundwater storage variations ( $\Delta GWS$ ) identified from this  $GWS$  time-series were then used in regional-scale groundwater balance calculations to identify in particular two unknown terms: the recharge and the evaporation.

## 4. Results

### 4.1. Hydrodynamic conceptual model for the TAS based on piezometric information

Piezometric sketches not available hitherto were obtained here using the direct and indirect water levels collected (see Section 3.1 and Supplementary Information). The main objective was to identify recharge and discharge areas for the different aquifers of the TAS as well as the main flow directions. For this purpose, we used the local polynomial interpolation package in R which is not an exact interpolator, thus producing smooth surfaces given the fact that data of different origins and dates (despite low temporal variations, see Section 2.3) were used.

With this approach, the conceptual model for the hydrogeology of the TAS was largely updated. The piezometric levels shown in Figure 4 point to a similar hydrodynamic behavior between Cambro-Ordovician, Devonian and Viséan aquifers on the one hand and Hamadian aquifer on the other hand consistent with the stratigraphic unconformity. Regarding first the deeper Paleozoic aquifers, the highest piezometric levels (above 500 m) are observed at the northwest boundary of the basin, making this area a major recharge zone from which groundwater flow diverges towards two discharge areas: the southeastern border and the northwestern boundary close to the Atlantic coastal area (Figure 4). The southeastern discharge zone ought to be a major outlet of the Paleozoic aquifers in accordance with the presence of numerous sebkhas especially aligned along the outcrops of Paleozoic formations in this

hyper-arid area [Figures 2 and 3, and Pekel *et al.*, 2016]. At the location of the sebkhas, the discharge occurs by exfiltration and evaporation due to the proximity of the water table to the ground surface [Lamouroux and Hani, 2006]. Southeastern groundwater discharge is also supported by reconnaissance operations carried out by the French army in Algeria in the 1930's which identified numerous springs and possibly downstream sebkhas [Martin, 1930]. An area of likely modest recharge can also be identified in the southwest of the TAS for deep aquifers that corresponds to small topographic highs with relatively higher rainfall than the remaining southern boundary. Regarding now the Hamadian aquifer, the same recharge area at the Anti-Atlas piedmont as for the Paleozoic aquifers can be clearly identified in Figure 4. The likely presence of this recharge area was commented on but not evidenced by Lamouroux and Hani [2006]. Conversely to the other aquifers, there is only one discharge area for the Hamadian aquifer. It corresponds closely to the Sebkhha of Tindouf in central and southern position, making the area east of the  $-8^{\circ}W$  and south of  $28^{\circ}N$  certainly the main discharge area of the whole TAS. The other western recharge area for the Hamadian aquifer, i.e. the Hamada du Drâa plateau already described is also clearly identified in Figure 4.

Piezometric sketches suggest a common feature for all aquifers, i.e. the presence of groundwater divides or a convergence area between the parallels  $-8W$  and  $-7W$  (dashed lines in Figure 4) separating the TAS into two sub-basins. For all aquifers these groundwater divides correspond to a zero-flow boundary except for the Hamadian for which the central part is a discharge area that will be taken into account in the groundwater budget equations (see section below).

### 4.2. GRACE-derived groundwater balance of the TAS

Before correcting the GRACE-derived TWS data from soil water storage (SWS, see Section 3.2), the compatibility of the different soil models with the GRACE data needs to be verified for the region of interest [Gonçalves *et al.*, 2013, Scanlon *et al.*, 2019]. From this point of view, Scanlon *et al.* [2019] recently reiterated that the consistency of SWS fluctuations (seasonality) with that displayed by the GRACE signal

(SWS amplitudes expected to be less than or equal to that of TWS) should be verified. For this purpose, domain average terrestrial and soil water storage time-series were considered first for comparison purposes (Figure 5). The VIC and NOAH models show a sharp rise of about 100 mm in soil water storage at the end of 2014 then a roughly stable or declining plateau that does not appear as much in the GRACE data. Consequently, these two models that are inconsistent with the GRACE data (gravity observations) were rejected and only the CLSM model was considered in the following analysis. Despite a sharp increase also for the CLSM model in 2014, the latter shows the best overall consistency with the GRACE TWS time-series. It can be noted that the CLSM model is consistent with the soil model considered by Frappart [2020] showing a weak interannual storage variation (low trend). Using the CLSM model to identify SWS and to produce the GWS time series ( $TWS - SWS$ ) leads to an overall groundwater depletion of the TAS  $\Delta GWS = -0.7 \pm 0.1 \text{ mm}\cdot\text{yr}^{-1}$  (Figure 5).

The regional distribution of GRACE-derived values of  $\Delta GWS$  shows a western part in quasi-equilibrium regarding groundwater mass variations while the eastern part is in clear depletion (Figure 6). Moreover, the largest groundwater depletion of the TAS is identified at the southeastern boundary of the basin over a limited area of 16,000 km<sup>2</sup> and at a rate of about  $-5 \text{ mm}\cdot\text{yr}^{-1}$ . Interestingly, this large depletion zone corresponds exactly to the areas of high-density regarding salt pans at downstream Paleozoic aquifer outcrops where high evaporation is expected (see Figures 2 and 3). Therefore, evaporation may be a prominent outflow for the TAS without being precisely quantified at this stage.

The overall water balance for the TAS involves two unknowns: the recharge and the natural groundwater discharge, here dominated by evaporation at the sebkhas mostly in south-east quarter of the domain (sebkha of Tindouf and numerous sebkhas along the Viséan, Devonian and Cambro-Ordovician outcrops). Therefore, writing a single water balance for the entire basin as previously done e.g. by Gonçalves *et al.* [2013] for the NWSAS cannot provide independent estimates for these two quantities. Considering the results of the water level analysis (Section 4.1) and the spatial distribution of the sebkhas where this evaporation occurs (Figure 3), the basin is subdivided into two sub-domains separated by the 8th

parallel: an eastern, more arid, part where the sebkhas are mainly located (main discharge areas) and a western part, slightly more humid on average. Two distinct groundwater-balance equations were written over these two domains assuming negligible east-west groundwater exchanges based on the piezometric sketches of Section 4.1. It should be noted that for the Hamadian, groundwater converges towards this central area and do not pass from one sub-basin to another one. This sub-division is also largely supported by the GRACE data which show a western half-domain in near equilibrium with respect to the water budget with  $\Delta GWS^W = 0.1 \text{ mm}\cdot\text{yr}^{-1}$ , while the eastern area is characterized by a groundwater depletion with  $\Delta GWS^E = -1.3 \pm 0.1 \text{ mm}\cdot\text{yr}^{-1}$ . The GRACE-derived water budgets for the western and eastern sub-domains of the TAS are written:

$$\Delta GWS^W = -Q_W^W - Q_D^W + R_n^W + R_a^W, \quad (1)$$

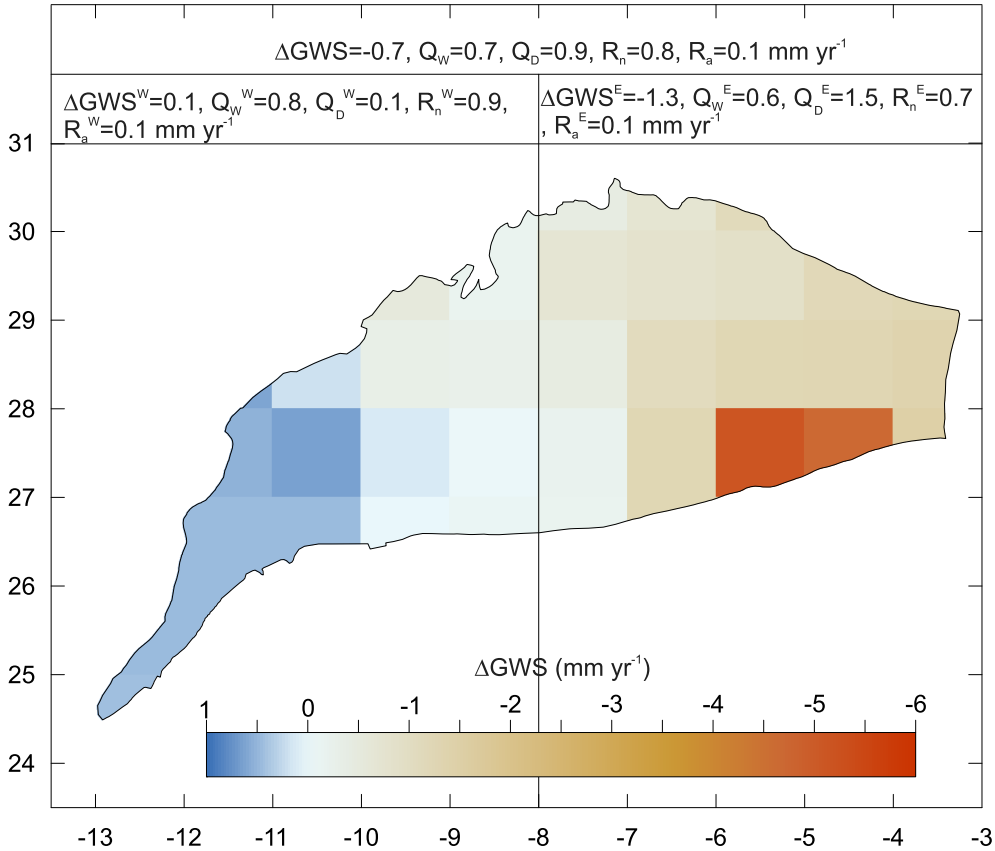
and:

$$\Delta GWS^E = -Q_W^E - Q_D^E + R_n^E + R_a^E \quad (2)$$

respectively, where  $R_n^i$  and  $R_a^i$  (with  $i = W$  or  $E$ ) are the natural recharge and artificial recharge due to irrigation return flow,  $Q_W^i$  the groundwater withdrawals expressed in water height ( $\text{mm}\cdot\text{yr}^{-1}$ ), and  $Q_D^i$  ( $\text{mm}\cdot\text{yr}^{-1}$ ) is the natural discharge (mostly evaporation).

Solving the TAS groundwater balance requires the identification of pumping rates for which only the global value is known. Groundwater extraction for the Algerian part [a few million cubic meters per year; García *et al.*, 2021] is negligible compared to the exploitation in Morocco mainly for irrigation (see Section 2.3). Global values at  $323 \times 10^6 \text{ m}^3\cdot\text{yr}^{-1}$  are given by Hssaisoune *et al.* [2020] for the Drâa basin and at  $22 \times 10^6 \text{ m}^3\cdot\text{yr}^{-1}$  for the Guelmim plain [Figure 1; Touzani, 2012] in Morocco. The methodology proposed here involves redistributing the global value to obtain pumping values  $Q_W^i$  for the two sub-basins. The partitioning of pumping between the two sub-basins was made using two alternative approaches: (i) using the spatial distribution of irrigated surfaces from FAO AQUASTAT (data version 5 available at <https://data.apps.fao.org/aquamaps/>), and (ii) the spatial distribution of pumping wells available from the Moroccan Souss-Massa-Drâa hydrological basin





**Figure 6.** Spatial distribution of Groundwater storage variations  $\Delta GWS$  (18 years average) of the Tindouf aquifer system and Groundwater budget fluxes (see Equations (1) and (2)) expressed in water heights.

agency (ABHSM, <http://www.abhsm.ma>). The irrigated surface areas within the TAS domain represent 25% east of  $-8^\circ W$  and 19% west of  $-8^\circ W$  of the global irrigation areas in the Drâa basin, and 15% of those in the Guelmim plain (western sub-basin). Assuming a proportionality between cultivated surfaces and groundwater abstractions yields  $Q_W^W = 65 \times 10^6 \text{ m}^3 \cdot \text{yr}^{-1}$  (19% of  $323 \times 10^6 \text{ m}^3 \cdot \text{yr}^{-1}$  plus 15% of  $22 \times 10^6 \text{ m}^3 \cdot \text{yr}^{-1}$ ) and  $Q_W^E = 82 \times 10^6 \text{ m}^3 \cdot \text{yr}^{-1}$ . The eastern and western sub-basins each contain 30% of the Drâa basin pumping wells while 50% of the Guelmim plain boreholes are located in the western part. Assuming again a linear relation between borehole density and groundwater abstractions yields  $Q_W^W = 108 \times 10^6 \text{ m}^3 \cdot \text{yr}^{-1}$  and  $Q_W^E = 97 \times 10^6 \text{ m}^3 \cdot \text{yr}^{-1}$ . The different values obtained for the pumping in the TAS are due to contrasting relative proportions of pumping and agricultural areas. Considering now average values between the two ap-

proaches leads to  $Q_W^W = 87 \pm 30 \times 10^6 \text{ m}^3 \cdot \text{yr}^{-1}$  or  $0.8 \pm 0.3 \text{ mm} \cdot \text{yr}^{-1}$  and  $Q_W^E = 90 \pm 10 \times 10^6 \text{ m}^3 \cdot \text{yr}^{-1}$  or  $0.6 \pm 0.1 \text{ mm} \cdot \text{yr}^{-1}$  over the  $112,000 \text{ km}^2$  and  $143,000 \text{ km}^2$  surface area of the western and eastern sub-domains, respectively. Following Gonçalves et al. [2013] artificial recharge (irrigation return flow) represents a fraction  $f = 15 \pm 15\%$  of withdrawals, leading to  $R_a^i = f \times Q_W^i$ . Regarding the last two terms on the right-hand side of Equations (1) and (2), linear relations between natural recharge and AAR ( $R_n^i = \alpha \times AAR^i$ ,  $AAR^E = 50 \text{ mm} \cdot \text{yr}^{-1}$  and  $AAR^W = 62 \text{ mm} \cdot \text{yr}^{-1}$ ) and between natural discharge (evaporation) and salt pans surface  $S_{SP}^i$  ( $Q_D^i = \beta \times S_{SP}^i$ ,  $S_S^W = 150 \text{ km}^2$  and  $S_{SP}^E = 3500 \text{ km}^2$ ) were assumed. By introducing these relations together with the values for  $\Delta GWS^W$ ,  $\Delta GWS^E$ ,  $Q_W^W$ ,  $Q_W^E$  and the corresponding artificial recharge values in Equations (1) and (2),  $\alpha$  ( $1.39 \pm 0.46 \times 10^{-2}$ ) and  $\beta$  ( $4.48 \pm 0.88 \times 10^{-4} \text{ mm} \cdot \text{km}^{-2} \cdot \text{yr}^{-1}$ ) and then the natural recharge and

discharge in each sub-domain given in Figure 6 were identified. The domain-averaging of these values using sub-domain surface areas yielded  $R_n = 0.8 \pm 0.2$  mm·yr<sup>-1</sup> ( $196 \pm 51 \times 10^6$  m<sup>3</sup>·yr<sup>-1</sup>) and an evaporation  $Q_D = 0.9 \pm 0.2$  mm·yr<sup>-1</sup> ( $225 \pm 51 \times 10^6$  m<sup>3</sup>·yr<sup>-1</sup>) for the entire TAS domain.

## 5. Discussion

In this study, the first step was to identify the hydrogeological limits of the basin as accurately as possible. Geological features such as arches or highs which likely constitute hydraulic divides were looked for. From this standpoint, the numerous geological studies available giving the structure and lithological descriptions provided valuable information for hydrogeological analysis. Low-extension anticlines were included in the domain in the north-northwestern zone where TAS limits were defined by the contact between the Cambrian and the crystalline basement. These anticlines can lead to local flow structures (e.g., piezometric highs, streamlines developing around the anticline) that disrupt the overall topographically driven groundwater flow pattern. The precise nature and geometry of these hydrodynamic disturbances cannot be characterized using piezometric information, which is too sparse and poorly distributed geographically. However, these anticline areas of moderate surfaces (few thousand km<sup>2</sup>) were integrated into the domain (i) for overall geological consistency, and (ii) for water balance purposes based on satellite data on a multi-layer system whose extension is ultimately given by that of the Cambrian formation. The geographical extension presented here can therefore be used to update the TAS limits in the international synthesis of aquifers by IGRAC (International Groundwater Resources Assessment Centre) [2015].

The physiographic information (DEM, geomorphology) and satellite images of earth surface to identify the salt pans were then considered in a second step in order to draw a conceptual model for the hydrogeology corresponding to an endorheic topographically-driven basin. This interpretation is consistent with the generally relatively high salinity reported for the groundwater of the different aquifers which is often in the order of a few g/l, requiring reverse osmosis treatment plants for water consumption [García *et al.*, 2021, Lamouroux and Hani, 2006].

Here, this conceptual model was supported by piezometric sketches obtained by compensating the lack of water level data with indirect information of varying quality: DEM-derived elevations of springs, sebkhas, and traditional wells. The maximum value of the DEM uncertainty (SRTM 1 arcsec) at the global scale is 16 m. However, when possible, this uncertainty should be identified locally by comparing the DEM values with reliable elevation data. To our knowledge, this comparison has not been carried out on the present study area but on a nearby area in Morocco [Khal *et al.*, 2020]. The authors identified an average error of about 2 m for SRTM 1 arcsec. This error is thus larger than the accuracy of the piezometry but remains acceptable given the several hundred meters range of variation of piezometric levels. Of the indirect piezometric information, the sebkhas are the most uncertain. Their use assumes a preferential groundwater supply, which is verified if the free water exists long enough after the rain. The validity of such condition is difficult to verify. The satellite imagery interpretation proposed by Pekel *et al.* [2016] can generally identify only one to three years of water bodies occurrence in the period 1984–2021 in the area of interest for the use of sebkhas (southeast of the isohyets 30 mm/year, Figure 3). Generally, the sebkhas in this area contain water the year following a particularly rainy year (according to the CRU data in Figure 3), which suggests, but does not formally establish, the occurrence of surface water bodies with some time lag. Only the Hamadian could be really impacted by this uncertainty on the two sebkhas considered here since no other data is available in the southeast part of this aquifer, unlike the other three aquifers. However, the overall interpretation in two hydraulic sub-basins used as a basis for the water budget calculations using GRACE is little impacted by this uncertainty, including for the Hamadian since the piezometric data and the well (Figure 4) for the latter already suggests the considered structuring. The piezometric spatialization proposed here is certainly approximate but the use of more robust interpolation methods was difficult to consider due to the scarcity and heterogeneity of the data as well as their sparse spatial distribution. Although approximate, these piezometric sketches are not only illustrative but provide a valuable scheme for the hydrodynamic functioning of the TAS (main streamlines, recharge and discharge areas) consistent with what was known in the literature and with the GRACE

data. Therefore, the proposed piezometric sketches, which represent a first attempt, should be confirmed by a consistent water level data set, even if this seems difficult at this time. However, the good consistency between (i) the conceptual model proposed here for the TAS, (ii) the piezometric sketches for the main aquifers of the TAS, and (iii) the regional distribution of  $\Delta GWS$  derived from the GRACE-GLDAS analysis reinforces the confidence in the simple piezometric analysis proposed here.

Regarding the overall groundwater budget using GRACE data, only an interannual analysis aiming at identifying long-term average flows such as recharge is proposed here. This is adapted to arid environments characterized by a lag time of sometimes several decades between rainfall and recharge related to thick vadose zones as discussed in Gonçalves *et al.* [2021] or Séraphin *et al.* [2022]. However, in climatic and/or geological contexts where recharge is rapid, a seasonal analysis using GRACE data can be conducted [e.g., Pascal *et al.*, 2022]. In addition, only one GLDAS soil model (CLMS) could be retained following the comparative analysis with the *TWS* signal, and uncertainties at each step (linear regression on *GWS* time series, pumping estimation, return flow) were proposed to quantify the uncertainty of each water budget terms. Similarly to Frappart [2020], this study indicates a very small groundwater storage change for the TAS. However, a slight decrease in *GWS* of  $0.7 \text{ mm}\cdot\text{yr}^{-1}$  was obtained here and not a slight increase of about  $0.2 \text{ mm}\cdot\text{yr}^{-1}$  as previously proposed. This increase can be explained by an erroneous circular domain considered in Frappart [2020] rather centered on the west of the actual TAS basin, which was indeed found in slight gain in the present study. Finally, the discharge of the Paleozoic aquifers at the northwestern limit in the downstream valley of the Drâa River suggested by the piezometric maps was not considered. However, a possible loss of this natural outflow caused by pumping is discussed below.

An almost homogeneous land use and climate enables a direct comparison between the recharge of  $0.8 \pm 0.2 \text{ mm}\cdot\text{yr}^{-1}$  identified here for the TAS during the period 2002–2020 for a long-term value of AAR at  $55 \text{ mm}\cdot\text{yr}^{-1}$  with the values previously obtained for the other regional-scale aquifers of the Saharan belt. These aquifers are the Nubian Sandstone Aquifer System (NSAS,  $2.2 \times 10^6 \text{ km}^2$  of surface area over Egypt, Libya, Chad, and Sudan), the

Murzuk aquifer ( $0.64 \times 10^6 \text{ km}^2$  of surface area over Libya, Niger and Algeria) and the Northwestern Sahara aquifer system (NWSAS,  $10^6 \text{ km}^2$  of extension over Algeria, Tunisia and Libya). For the NSAS, a previous study by Mohamed *et al.* [2017] yielded a recharge over the entire domain of  $0.48 \pm 0.63 \text{ mm}\cdot\text{yr}^{-1}$  for an AAR of  $36 \text{ mm}\cdot\text{yr}^{-1}$ . The recharge of the Murzuk aquifer was estimated at  $0.25 \pm 0.40 \text{ mm}\cdot\text{yr}^{-1}$  by Mohamed and Gonçalves [2021] for an AAR of  $25 \text{ mm}\cdot\text{yr}^{-1}$ . Here it is assumed that GRACE-derived recharge values are good estimates of long term values for the three aquifers. Using GRACE data and a global groundwater mass balance, a recharge of  $1.76 \pm 0.44 \text{ mm}\cdot\text{yr}^{-1}$  was derived for the NWSAS by Mohamed and Gonçalves [2021] for a long-term value of AAR at  $75 \text{ mm}\cdot\text{yr}^{-1}$ . When plotted against the AAR, recharge values for the different transboundary aquifers of the African Saharan belt show a good consistency and represent a small fraction of AAR, i.e. about  $1.5 \pm 0.6\%$ .

Here a value of  $225 \pm 51 \times 10^6 \text{ m}^3\cdot\text{yr}^{-1}$  for the overall groundwater evaporation was obtained. Distributed over the  $3550 \text{ km}^2$  of surface area for the salt pans this represents an evaporation rate of  $63 \pm 14 \text{ mm}\cdot\text{yr}^{-1}$  and a basin-scale average of  $0.9 \pm 0.2 \text{ mm}\cdot\text{yr}^{-1}$ . Local values were obtained by Fontes *et al.* [1986] by means of stable isotopes of water profiles through the vadose zone in southwestern Algerian Sahara close to the TAS. An evaporation value of  $2.3 \text{ mm}\cdot\text{yr}^{-1}$  was identified for a water table depth at 8 m. An intermediate scale study by Schulz *et al.* [2015] for a *sebkha* system covering a surface area of  $3600 \text{ km}^2$  in southwestern Saudi Arabia yielded a value of  $39 \text{ mm}\cdot\text{yr}^{-1}$  using soil column experiments to obtain a relation between evaporation rate and temperature, relative humidity and water level for spatial distribution and domain averaging purposes. Large scale values are also available for two arid domains. The first one is the Central Kalahari basin in Southern Africa where a regional scale value of  $1.87 \text{ mm}\cdot\text{yr}^{-1}$  was obtained using a calibrated hydrogeological model of this  $195,000 \text{ km}^2$  surface area domain. Hydrogeological modeling of the NWSAS [Baba-Sy, 2005] led to an estimate on the order of  $190 \times 10^6 \text{ m}^3\cdot\text{yr}^{-1}$  ( $35 \text{ mm}\cdot\text{yr}^{-1}$ ) for the evaporation flux in the Chott Djérid.

Recharge and evaporation were identified on the two presumably independent sub-basins on which groundwater budgets were established based on GRACE-GLDAS data. These two budgets must now

be separately interpreted to explain possible causes of imbalance which are also valid for the entire TAS whose groundwater balance cannot be directly estimated. An incautious analysis of the reconstructed overall water balance of the TAS could lead to misinterpretation such as a groundwater depletion caused only by pumping, which is inconsistent with the minimum  $\Delta GWS$  values (largest depletion) in the non-pumping southeastern area. The western part of the TAS is at hydrodynamic equilibrium despite pumping, which is only possible if some groundwater discharge that naturally balanced the recharge has dried up. This natural discharge may correspond to the topographical low in the lower Drâa valley (Section 2.3), i.e. within the pumping area. The overall groundwater depletion of the TAS is therefore entirely attributed to the eastern part. Regardless of pumping, the groundwater budget of the eastern sub-basin is “naturally” unbalanced since recharge is lower than evaporation ( $0.7$  and  $1.5 \text{ mm}\cdot\text{yr}^{-1}$ , see Figure 6). This contributes  $-0.8 \text{ mm}\cdot\text{yr}^{-1}$  ( $\approx 60\%$ ) to the overall balance of  $-1.3 \text{ mm}\cdot\text{yr}^{-1}$  in this sub-domain. A recent increase in evaporation e.g. associated with the ongoing global warming or a past higher recharge and thus lateral groundwater drainage in equilibrium with evaporation then a reduction of the former can be invoked to explain the unbalanced evaporation. Evaporation from the water table is controlled by the depth of the free surface but also by climatologic variables [soil surface atmospheric air temperature and relative humidity, see e.g. Kamai and Assouline, 2018] that can vary on interannual or even long-term time scales. Regarding recharge variations, the few available values for the hydraulic parameters (Section 2.3) together with recharge–discharge distances of several hundreds of km suggest characteristic times [Rousseau-Gueutin *et al.*, 2013] for the TAS aquifers of several thousand years and point to the Holocene period for such variations. The remaining 40% of the eastern sub-basin negative groundwater budget is due to pumping. These contributions of evaporation and pumping to the negative groundwater balance also apply to the entire TAS upon surface averaging since the western subbasin is budget neutral.

## 6. Conclusions

The combined geological, geomorphological and land surface analysis conducted here concluded to

a topographically-driven conceptual model for the groundwater hydrodynamics of the TAS involving at least one recharge area, the Anti-Atlas piedmont, and one major discharge area through groundwater evaporation located on the southeastern boundary of the basin where the highest density of salt pans (sebkhas) was found. Based on abundant geological information notably related to petroleum engineering studies, more realistic geographical limits than those considered so far in IGRAC (International Groundwater Resources Assessment Centre) [2015], and therefore with a relevant hydrodynamic significance, were proposed. The piezometric sketches, in good general agreement with our conceptual model, provide however more valuable insights. A second recharge area, i.e. the Hamada plateau for the Hamadian and the Carboniferous aquifers and the southwestern border for the Devonian and Cambro-Ordovician aquifers, was identified. In addition, a second discharge area on the northwest boundary of the TAS (downstream valley of the Drâa), which may be drying up, is suggested by the water levels for the Paleozoic aquifers. However, these results could be further established if additional water level data were available to obtain more reliable piezometric contour maps. A quantitative analysis of the groundwater mass balance was conducted using GRACE-GLDAS data leading to the conclusion of an overall depletion of the basin of about  $-0.7 \text{ mm}\cdot\text{yr}^{-1}$  ( $178.5 \times 10^6 \text{ m}^3\cdot\text{yr}^{-1}$ ). More precise water budget calculations identified separately the domain-averaged evaporation at  $0.9 \pm 0.2 \text{ mm}\cdot\text{yr}^{-1}$  ( $225 \pm 51 \times 10^6 \text{ m}^3\cdot\text{yr}^{-1}$ ) and the recharge at  $0.8 \pm 0.2 \text{ mm}\cdot\text{yr}^{-1}$  ( $196 \pm 51 \times 10^6 \text{ m}^3\cdot\text{yr}^{-1}$ ), confirming the endorheic behavior of the TAS. The depletion of the TAS attributed here to evaporation and withdrawals is however more moderate than in the other basins of the Saharan belt due to a relatively low exploitation of the aquifers.

## Conflicts of interest

Authors have no conflict of interest to declare.

## Acknowledgments

GRACE land data are available at <http://grace.jpl.nasa.gov>, supported by the NASA MEaSURES Program. This study has been supported by the Labex

OT-Med project of the French program “Investissement d’Avenir.” The two anonymous reviewers are acknowledged for their comments which contributed to a major improvement of the manuscript.

## Supplementary data

Supporting information for this article is available on the journal’s website under <https://doi.org/10.5802/crgeos.202> or from the author.

## References

- Baba-Sy, J. (2005). *Recharge et paléorecharge du système aquifère du Sahara septentrional*. PhD thesis, Univ. Tunis, Tunisia.
- Beuf, S., Biju-Duval, B., De Charpal, O., Rognon, P., Gariel, O., and Bennacef, A. (1971). *Les grès du Paléozoïque inférieur au Sahara*, volume 18 of *Collection Science et Technique du Pétrole*. Institut Français du Pétrole, Rueil-Malmaison, France.
- Boote, D. R. D., Clark-Lowes, D. D., and Traut, M. W. (1998). Palaeozoic petroleum systems of North Africa. *Geol. Soc. Lond. Spec. Publ.*, 132, 7–68.
- Brunner, P., Li, H. T., Kinzelbach, W., Li, W. P., and Dong, X. G. (2008). Extracting phreatic evaporation from remotely sensed maps of evapotranspiration. *Water Resour. Res.*, 44, article no. W08428.
- Bryant, R. G., Sellwood, B. W., Millington, A., and Drake, N. (1994). Marine-like potash evaporite formation on a continental playa: case study from Chott el Djerid, southern Tunisia. *Sediment. Geol.*, 90, 269–291.
- Coudrain-Ribstein, A., Pratz, B., Talbi, A., and Jusserand, C. (1998). Is the evaporation from phreatic aquifers in arid zones independent of the soil characteristics? *C. R. Acad. Sci.—Ser. Ila Sci. Terre planètes*, 326, 159–165.
- Craig, J., Rizzi, C., Said, F., Thusu, B., Luning, S., Asbali, A. I., and Hamblett, C. (2008). Structural styles and prospectivity in the Precambrian and Palaeozoic hydrocarbon systems of North Africa. *Geol. East Libya*, 4, 51–122.
- de Marsily, Gh., Delay, F., Gonçalves, J., Renard, Ph., Teles, V., and Violette, S. (2005). Dealing with spatial heterogeneity. *Hydrogeol. J.*, 13, 161–183.
- de Marsily, Gh., Ledoux, E., Levassor, A., Poittrinal, D., and Salem, A. (1978). Modelling of large multilayered aquifer systems: theory and applications. *J. Hydrol.*, 36, 1–34.
- Fontes, J. C., Yousfi, M., and Allison, G. B. (1986). Estimation of long-term, diffuse groundwater discharge in the Northern Sahara using stable isotope profiles in soil water. *Hydrol. J.*, 86, 315–327.
- Frappart, F. (2020). Groundwater storage changes in the major North African transboundary aquifer systems during the GRACE Era (2003–2016). *Water*, 12, article no. 2669.
- Frizon de Lamotte, D., Leturmy, P., Missenard, Y., Khomsi, S., Ruiz, G., Saddiqi, O., Guillocheau, F., and Michard, A. (2009). Mesozoic and Cenozoic vertical movements in the Atlas system (Algeria, Morocco, Tunisia): An overview. *Tectonophysics*, 475, 9–28.
- García, R., Naves, A., Anta, J., Ron, M., and Molinero, J. (2021). Drinking water provision and quality at the Sahrawi refugee camps in Tindouf (Algeria) from 2006 to 2016. *Sci. Total Environ.*, 780, article no. 146504.
- Ghienne, J.-F., Le Heron, D., Moreau, J., Denis, M., and Deynoux, M. (2007). The Late Ordovician glacial sedimentary system of the North Gondwana platform. In Hambrey, M., Christoffersen, P., Glasser, N., Janssen, P., Hubbard, B., and Siegert, M., editors, *Glacial Sedimentary Processes and Products*, Special Publication no. 39, International Association of Sedimentologist, pages 295–319. Blackwell, Oxford, UK.
- Gonçalves, J., Deschamps, P., Hamelin, B., Vallet-Coulomb, C., Petersen, J., and Chekireb, A. (2020). Revisiting recharge and sustainability of the North-Western Sahara aquifers. *Reg. Environ. Change*, 20, article no. 47.
- Gonçalves, J., Petersen, J., Deschamps, P., Hamelin, B., and Baba-Sy, O. (2013). Quantifying the modern recharge of the “fossil” Sahara aquifers. *Geophys. Res. Lett.*, 40, 2673–2678.
- Gonçalves, J., Séraphin, P., Stieglitz, T., Chekireb, A., Hamelin, B., and Deschamps, P. (2021). Coastal aquifer recharge and groundwater–seawater exchanges using downscaled GRACE data: case study of the Djeffara plain (Libya–Tunisia). *C. R. Geosci.*, 353, 297–318.
- Guerrak, S. (1989). Time and space distribution of Paleozoic oolitic ironstones in the Tindouf Basin, Algerian Sahara. In Young, T. P. and Taylor, W. E. G., editors, *Phanerozoic Ironstones*, volume 46 of *Geological Society, London, Special Publications*, pages 197–212. Geological Society, London, UK.

- Guiraud, R., Bosworth, W., Thierry, J., and Delplanque, A. (2005). Phanerozoic geological evolution of Northern and Central Africa: an overview. *J. Afr. Earth Sci.*, 43, 83–143.
- Harris, I., Osborn, T. J., Jones, P., and Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci. Data*, 7, article no. 109.
- Heinl, M. and Brinkmann, P. J. (1989). A groundwater-model of the Nubian aquifer system. *Hydrol. Sci. J.*, 34, 425–447.
- Heiß, L., Bouchaou, L., Tadoumant, S., and Reichert, B. (2020). Multi-tracer approach for assessing complex aquifer systems under arid climate: Case study of the River Tata catchment in the Moroccan Anti-Atlas Mountains. *Appl. Geochem.*, 120, article no. 104671.
- Hssaisoune, M., Bouchaou, L., Sifeddine, A., Bouimetarhan, I., and Chehbouni, A. (2020). Moroccan groundwater resources and evolution with global climate changes. *Geosciences*, 10, article no. 81.
- IGRAC (International Groundwater Resources Assessment Centre) (2015). U.-I. (UNESCO I.H.P). Transboundary Aquifers of the World Map 2015 | International Groundwater Resources Assessment Centre. Available online: <https://www.un-igrac.org/resource/transboundary-aquifers-world-map-2015> (accessed on 1 December 2021).
- Kamai, T. and Assouline, S. (2018). Evaporation from deep aquifers in arid regions: Analytical model for combined liquid and vapor water fluxes. *Water Resour. Res.*, 54, 4805–4822.
- Khal, M., Algouti, A., Algouti, A., Akdim, N., Stankevich, S. A., and Menenti, M. (2020). Evaluation of open digital elevation models: estimation of topographic indices relevant to erosion risk in the Wadi M'Goun watershed, Morocco. *AIMS Geosci.*, 6, 231–257.
- Lamouroux, C. and Hani, A. (2006). Identification of groundwater flow paths in complex aquifer systems. *Hydrol. Process*, 20, 2971–2987.
- Lekula, M. and Lubczynski, M. W. (2019). Use of remote sensing and long-term in-situ time-series data in an integrated hydrological model of the Central Kalahari Basin, Southern Africa. *Hydrogeol. J.*, 27, 1541–1562.
- Leprêtre, R. (2015). *Evolution phanérozoïque du Craton Ouest Africain et de ses bordures Nord et Ouest*. PhD thesis, Université Paris Sud - Paris XI.
- Loi, A., Ghienne, J.-F., Dabard, M. P., Paris, F., Botquelen, A., Christ, N., and Destombes, J. (2010). The Late Ordovician glacio-eustatic record from a high-latitude storm-dominated shelf succession: The BouIngarf section (Anti-Atlas, Southern Morocco). *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, 296, 332–358.
- Lüning, S., Craig, J., Loydell, D. K., Štorch, P., and Fitches, B. (2000). Lower Silurian “hot shales” in North Africa and Arabia: regional distribution and depositional model. *Earth Sci. Rev.*, 49, 121–200.
- Mahia, M., Djorfi, S., and Hani, A. (2017). Alternatives for water resources management in the Tindouf Basin (Algeria). *J. Bio. Env.*, 11, 123–131.
- Maloof, A. C., Ramezani, J., Bowring, S. A., Fike, D. A., Porter, S. M., and Mazouad, M. (2010). Constraints on early Cambrian carbon cycling from the duration of the Nemakit-Daldynian-Tommotian boundary 13C shift, Morocco. *Geology*, 38, 623–626.
- Martin, C. (1930). Observations Géographiques sur les Eglab et le Haut-Plateau du Dra. In *HESPERIS, Archives Berbères et Bulletin de l'Institut des Hautes-Etudes Marocaines-Tome XI*. E. Larose, Paris, France.
- Michard, A., Hoepffner, C., Soulaïmani, A., and Baidder, L. (2008). The variscan belt. In Michard, A., Saddiqi, O., Chalouan, A., and de Frizon Lamotte, D., editors, *Continental Evolution: The Geology of Morocco*, volume 126 of *Lecture Notes in Earth Sciences*, pages 65–132. Springer, Berlin, Heidelberg, Germany.
- Mohamed, A. and Gonçalves, J. (2021). Hydrogeophysical monitoring of the North-Western Sahara Aquifer System's groundwater resources using gravity data. *J. Afr. Earth Sci.*, 178, article no. 104188.
- Mohamed, A., Sultan, M., Ahmed, M., Yan, E., and Ahmed, E. (2017). Aquifer recharge, depletion, and connectivity: inferences from GRACE, land surface models, and geochemical and geophysical data. *Bull. Geol. Soc. Am.*, 129, 534–546.
- Nimmo, J. R., Stonestrom, D. A., and Akstin, K. C. (1994). The feasibility of recharge rate measurements using the steady state centrifuge method. *Soil Sci. Soc. Am. J.*, 58, 49–56.
- Nutz, A., Ghienne, J.-F., and Storch, P. (2013). Circular, cryogenic structures from the Hirnantian deglaciation sequence (Anti-Atlas, Morocco). *J. Sediment.*

- Res.*, 83, 115–131.
- Nutz, A., Kwicien, O., Breitenbach, S. F. M., Cai, Y., Della Porta, G., Danisch, J., Kabiri, L., and Bodin, S. (2019). Fluvio-lacustrine sedimentation in the Agadir-Tissint Feija (Anti-Atlas, Morocco): A promising palaeoclimate archive for the last glacial cycle in northwest Africa. *Depos. Record*, 5, 362–387.
- Pascal, C., Ferrant, S., Selles, A., Maréchal, J.-C., Paswan, A., and Merlin, O. (2022). Evaluating downscaling methods of GRACE (gravity recovery and climate experiment) data: a case study over a fractured crystalline aquifer in southern India. *Hydrol. Earth Syst. Sci.*, 26, 4169–4186.
- Pekel, J.-F., Cottam, A., Gorelick, N., and Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, 540, 418–422.
- Perron, P. (2019). *Tectonique et architecture des bassins intracratoniques Paléozoïques : Impact sur l'enregistrement sédimentaire et la géométrie des réservoirs associés. Exemple de la marge Nord Gondwanienne*. PhD thesis, Dijon. 397 p.
- Pizzi, G. and Sartori, L. (1984). Interconnected groundwater systems simulation (IGROSS); description of the system and a case history application. *J. Hydrol.*, 75, 255–285.
- Richey, A. S., Thomas, B. F., Lo, M.-H., Reager, J. T., Famiglietti, J. S., Voss, K., Swenson, S., and Rodell, M. (2015). Quantifying renewable groundwater stress with GRACE. *Water Resour. Res.*, 51, 5217–5238.
- Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and Toll, D. (2004). The global land data assimilation system. *Bull. Am. Meteorol. Soc.*, 85, 381–394.
- Ron Martin, M. (2020). *Geología e hidrogeología de los asentamientos saharauís en la cuenca de Tindouf (Argelia y Sahara Occidental)*. PhD thesis, Universidad Complutense de Madrid, Madrid, Spain. Volumes I and II.
- Rousseau-Gueutin, P., Love, A. J., Vasseur, G., Robinson, N. I., Simmons, C. T., and de Marsily, G. (2013). Time to reach near-steady state in large aquifers. *Water Resour. Res.*, 49, 6893–6908.
- Scanlon, B. R., Zhang, Z., Rateb, A., Sun, A., Wiese, D., Save, H., Beaudoin, H., Lo, M. H., Müller-Schmied, H., Döll, P., Beek, R., Swenson, S., Lawrence, D., Croteau, M., and Reedy, R. C. (2019). Tracking seasonal fluctuations in land water storage using global models and GRACE satellites. *Geophys. Res. Lett.*, 46, 5254–5264.
- Scanlon, B. R., Zhang, Z., Save, H., Wiese, D. N., Landerer, F. W., Long, D., Longueuevigne, L., and Chen, J. (2016). Global evaluation of new GRACE mascon products for hydrologic applications. *Water Resour. Res.*, 52, 9412–9429.
- Schulz, S., Horovitz, M., Rausch, R., Michelsen, N., Mallast, U., Köhne, M., Siebert, C., Schüth, C., Al-Saud, M., and Merz, R. (2015). Groundwater evaporation from salt pans: Examples from the eastern Arabian Peninsula. *J. Hydrol.*, 531, 792–801.
- Séraphin, P., Gonçalves, J., Hamelin, B., Stieglitz, T., and Deschamps, P. (2022). Influence of intensive agriculture and geological heterogeneity on the recharge of an arid aquifer system (Saq-Ram, Arabian Peninsula) inferred from GRACE data. *Hydrol. Earth Syst. Sci.*, 26, 5757–5771.
- Tapley, B. D., Bettadpur, S., Watkins, M., and Reigber, C. (2004). The gravity recovery and climate experiment: Mission overview and early results. *Geophys. Res. Lett.*, 31, article no. L09607.
- Thorp, M., Glanville, P., Stokes, S., and Bailey, R. (2002). Preliminary optical and radiocarbon age determinations for Upper Pleistocene alluvial sediments in the southern Anti Atlas Mountains, Morocco. *C. R. Geosci.*, 334, 903–908.
- Touzani, M. (2012). Contribution à l'étude hydrogéologique de l'aquifère exploité dans la plaine de Guelmim. Master of Sc., University Cadi Ayyad, Marrakech, Morocco, 91 p.
- UNESCO (2006). Ressources en eau et gestion des aquifères transfrontaliers de l'Afrique du Nord et du Sahel, IHP-IV Series on Groundwater. 11, 134 p. Assessed at: [https://hydrologie.org/BIB/Publ\\_UNESCO/SOG11.pdf](https://hydrologie.org/BIB/Publ_UNESCO/SOG11.pdf).
- Watkins, M. M., Wiese, D. N., Yuan, D.-N., Boening, C., and Landerer, F. W. (2015). Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons. *J. Geophys. Res. Solid Earth*, 120, 2648–2671.
- Weerasinghe, I., Bastiaanssen, W., Mul, M., Jia, L., and van Griensven, A. (2020). Can we trust remote sensing evapotranspiration products over Africa? *Hydrol. Earth Syst. Sci.*, 24, 1565–1586.
- Wiese, D. N., Landerer, F. W., and Watkins, M. M. (2016). Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution. *Water*

- Resour. Res.*, 52, 7490–7502.
- Zammouri, M. and Ribeiro, L. (2017). Analyzing the effect of transmissivity uncertainty on the reliability of a model of the northwestern Sahara aquifer system. *J. Afr. Earth Sci.*, 129, 910–922.



# **Water in the anthropocene**





Research article

## Geo-hydrological Data & Models

# The water-agro-food system: upscaling from the Seine river basin to the global scale

Gilles Billen<sup>©,\*,a</sup> and Josette Garnier<sup>©,a</sup>

<sup>a</sup> Sorbonne Université, CNRS EPHE, Umr Metis 7619, 4 place Jussieu, 75005, Paris, France

E-mails: gilles.billen@upmc.fr (G. Billen), josette.garnier@upmc.fr (J. Garnier)

**Abstract.** Starting from the study of the major impacts of human activity in the Seine river basin on ground- and surface water quality, due to domestic and industrial wastewater and to intensive agricultural practices, a research framework was developed, combining the analysis of the agricultural systems, their connection to food requirement by local population and trade exchanges with other regions or countries, as well as the losses to the hydrosphere and atmosphere in terms of nitrogen, phosphorus, silica and carbon, and their associated environmental problems. From the Seine basin regional scale, the concept of water agro-food system was enlarged to the national, European and global scales. Coupling the GRAFS approach (Generalized Representation of Agro-Food Systems) to the Riverstrahler model (a biogeochemical drainage network model), allowed to make the link between the different issues of food production, river water quality and nutrient delivery to the coastal zone, as well as to develop an indicator of the risk for coastal water eutrophication (ICEP) now being included as an indicator for the United Nation Sustainable Development Goal (SDG) 14.

**Keywords.** River basin, Water quality, Agro-food system, Modelling, Nutrients, ICEP indicator.

*Manuscript received 24 April 2022, revised 29 June 2022, accepted 4 July 2022.*

## 1. Introduction

The PIREN-Seine program (<https://www.piren-seine.fr>), originating in 1989 as an applied interdisciplinary program devoted to the Seine river system and co-constructed with regional water managers [Billen, 2001], turned out to be a very fruitful source of inspiration for many research studies going well beyond the regional framework. The initial wide and long-term vision and the extreme open-mindedness of its founder, Professor Ghislain de Marsily, are certainly not unrelated to the unprecedented longevity of this still ongoing program.

The local stakes of river management, which were at the origin of the creation of the PIREN-Seine program were mainly related to water quality, endangered by pollution through untreated urban wastewater point release, particularly those of the huge Paris agglomeration (12 million inhabitants), as well as by diffuse sources of agricultural nutrients, pesticides and other toxic compounds [Meybeck et al., 1998, Garnier et al., 2022a]. Large sectors of the Seine River at that time were still oxygen deficient in summer, algal blooms were observed in the main tributaries, groundwater resources were threatened by nitrate and pesticides originating from one of the most intensive agricultural region of Europe, development of toxic algal species occurred in the Seine river plume in the English Channel [Cugier et al., 2005, Garnier et al., 2019b]. The PIREN Seine program, in close

\* Corresponding author.

connection with the Seine Water Agency and other stakeholders involved in the application of National and European directives, contributed to bring about efficient responses to several of these problems, although those linked to agriculture are still far from being solved.

At the same time, the awareness of global environmental issues has grown up. The concept of planetary boundaries [Rockström et al., 2009] defining the safe operating space for humanity with respect to the global Earth System, lead to identify the three major issues at stakes globally: increased greenhouse gas emission leading to climate change, loss of biodiversity and perturbation of the nitrogen cycle. Looking closely, the local stakes of river water management and the major planetary challenges appears as two sides of a same coin.

Here we aim to illustrate the development of the concept of water-agro-food system that gradually emerged from the diverse researches in the PIREN-Seine program. This concept offers a general framework for understanding the major challenges of environmental transition in an integrated view [Garnier et al., 2015, Billen and Garnier, 2021]. The essence of this approach is that of territorial biogeochemistry, derived from the territorial ecology concept [Buclet, 2015, Barles et al., 2011]. It is based on the conviction that most traits of the function of a territory (a geographical space appropriated by a society) are revealed by the analysis of the matter fluxes that cross it and ensure its operation [Billen et al., 2007]. Understanding and modeling these flows requires identifying their main control mechanisms and formalizing them. This can only be done based on a continuous confrontation of the calculated results with the observed field reality, which in turn allows increasing the predictive capacity of the model and its generalization to other situations and its applicability on larger scales. The Seine River system and its watershed have often served as the “demo-case” for conceiving original concepts and methods and to confront them to field reality, before being further applied to other systems and possibly enlarged in scope.

We report the gradual extension of our approach, starting from a renewed vision of river pollution issues established for understanding nutrient flows through the continuum of ecosystems from soil and agricultural systems (including human food require-

ment and wastes) to river systems and the coastal seas (Land-to-Sea continuums). This approach, although developed on the basis of confrontation with concrete observations at the local or watershed scale, could also be extended to regional and global scales. Further, we will illustrate how such an approach is of considerable help to conceive and explore future visions for the agro-food systems of France, Europe and the World, and specifically one scenario combining changes towards more healthy human diet and a sustainable agro-ecological agriculture, allowing to preserve the quality of water resources and halving greenhouse gas emissions.

## 2. Understanding river water quality

A river can be viewed from different perspectives, among which are (i) a receptacle and an easy way of disposal of wastewater; (ii) an aquatic ecosystem and habitat for fish; (iii) an active conveyor of water and nutrients originating from the watershed to the sea. The challenge of the early PIREN-Seine program was to consider these three perspectives together in a same modelling approach. The water quality models of the 1980's were dominated by the issue of oxygen deficits, a long lasting problem resulting from point wastewater release from Paris city [Gérardin, 1875, Streeter and Phelps, 1925, Lesouëf and André, 1982]. Taking into account the kinetics of microorganisms in the representation of the self-purification processes marked a great progress in the capacity to explain and predict oxygen deficits in the Seine river, not only immediately downstream from the large wastewater discharge from Paris agglomeration, but also 150–200 km downstream, in the Rouen estuarine sector, where ammonia was nitrified after a considerable lag required for the development of a large enough population of nitrifying bacteria [Chestérikoff et al., 1992, Brion et al., 2000, Brion and Billen, 2000, Garnier et al., 2007]. Also, in the 1990's algal development in large rivers became an issue for drinking water production [Garnier et al., 1995, 2005]. Taking into account phytoplankton and bacterial metabolisms in river water quality models opened the way to their transformation into complete aquatic ecosystem models, considering the dynamics of autotrophic and heterotrophic microorganisms, and the related major forms of carbon and nutrients. The RIVE module of the microbiological

processes involved in the ecological functioning of aquatic systems since then has grown considerably. It consists now of a generic representation of biogeochemical aquatic processes (Figure 1a). It has recently been fully coded in Python and published under GNU general public license (<https://gitlab.in2p3.fr/rive/pyrive>). RIVE is shared by a large number of modeling tools such as Riverstrahler [Billen et al., 1994, Garnier et al., 2002], Prose [Even et al., 1998], QualNET [Minaudo et al., 2018], and Barman [Garnier et al., 2000]. With a limited number of kinetic parameters, all of which have been experimentally determined in a variety of aquatic environments, subject to various hydrological and climatic constraints as well as to material inputs from the terrestrial systems, the biogeochemical functioning of these aquatic systems could be satisfactorily modelled.

A river system cannot be considered independently of its relationships with the terrestrial watershed, which controls not only the water flows into the tributaries, but also the quality of its surface water, defining the diffuse sources to the river system. The Riverstrahler model was developed within the PIREN-Seine program as a framework to formalize the relationships between an entire drainage network and the hydrological properties and land use of its watershed [Billen et al., 1994, Garnier et al., 2002] (Figure 1b). The same framework was still adopted in a Python version of the model, called PyNuts-Riverstrahler [Raimonet et al., 2018, Marescaux et al., 2020].

This Riverstrahler model is useful to describe quantitatively how nutrients brought from the watershed either as point or diffuse sources, are transferred, transformed, retained or eliminated during their travel through the drainage network down to the river outlet. Before reaching surface water however, flows of dissolved substances from the surface, sub-surface or groundwater runoff have to cross the riparian system, an often biogeochemically very active zone at the interface between land and river. The processes occurring in this riparian zone, including denitrification and N<sub>2</sub>O emissions have been explicitly taken into account in a recent version of the Riverstrahler model [Billen et al., 2018b, 2020a].

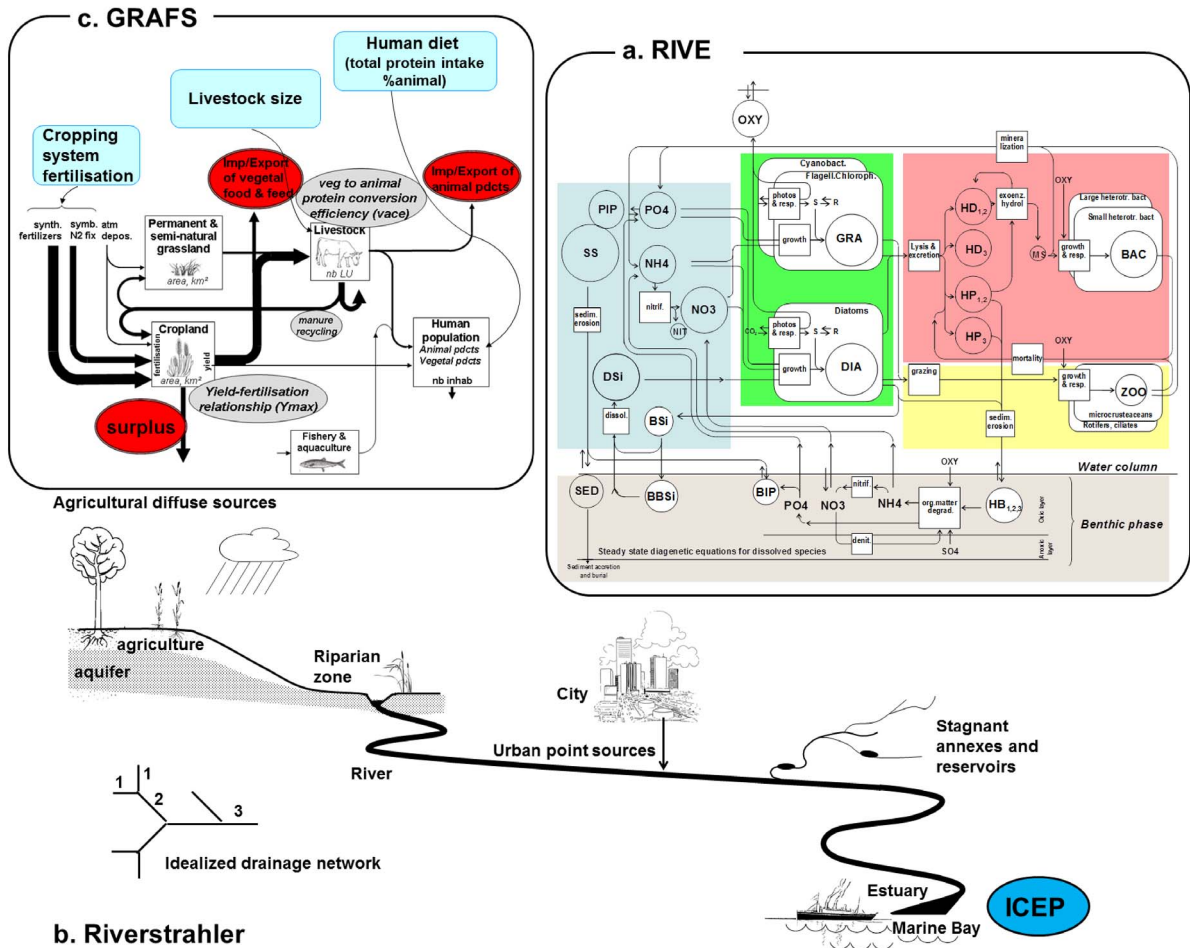
While most developments of the Riverstrahler model have been achieved and validated based on observations on the Seine River system, and often motivated by water management concerns in the

Seine Basin, the modelling approach was applied to a large number of river systems in France [Garnier et al., 2018b,a, 2019b], Europe [Garnier et al., 2002, Billen et al., 2005, Thieu et al., 2009, Desmit et al., 2018] and other rivers in the world with differing climate and land cover [Nordic rivers, Sferratore et al., 2008; the subtropical Red River, Le et al., 2015].

### 3. Modelling agro-food systems at territorial scale

Understanding the diffuse losses of nutrients to ground- and surface water requires description of how biological material is transformed within the mosaic of terrestrial systems that make up the watershed. Agriculture from that respect plays a major role. Very detailed agronomical models, such as the STICS or EPIC models, are able to calculate the growth of major crops of conventional systems and the associated flows of water and nutrients, including nitrogen leaching, at the plot scale [Williams et al., 1989, Brisson et al., 2003]. Application of such models at larger regional scale has been achieved, and coupled with hydrological models to predict the level of groundwater contamination from a detailed description of farming practices [Ledoux et al., 2007, Beaudoin et al., 2021]. Such a modelling approach, involving the coupling of the EPIC model to a river system model, is also at the core of the SWAT model of water quality in regional watersheds [Neitsch et al., 2001].

There is a need to understand the flows of nutrients involved in agriculture at territorial scale, in relation with structural features of the agro-food system, such as crop- and livestock farming orientations, human diet and trade exchanges of food and feed across the territory. For this purpose, the GRAFS approach (Generalized Representation of agro-food systems) was designed [Billen et al., 2013, 2021, Le Noë et al., 2017, 2018]. GRAFS considers cropland, permanent grassland, livestock and human as the four main compartments where fluxes of nitrogen, carbon and phosphorus are transformed and channeled through the system (Figure 1c). With regards to nitrogen, cropland receives inputs as synthetic fertilizers, manure, atmospheric deposition and symbiotic fixation and transforms them into harvested crops with a surplus at the origin of environmental losses. Livestock is fed partly from harvested crops, from grassland and/or from imported feed.



**Figure 1.** Schematic representation the different modelling approaches concurring to the model of the water-agro-food system of a watershed. (a) RIVE module of microbiological processes operating in aquatic ecosystems. (b) The Riverstrahler description of the interactions between the watershed and the river network. (c) The GRAFS model of the agro-food system. Direct connections can be established between the 3 components, leading to a full modelling of nutrient flows through the water-agro-food system.

Human diet, and its share of animal and vegetable food products defines the fate of crop and livestock production, as well as the amount of human excretion. GRAFS can be used as a simple framework for establishing a comprehensive budget of these material flows (N, P, C) within the agro-food system in a given territory based on available agricultural statistics such as fertilizer application, crop areas and production, livestock numbers and production, etc. Alternatively, once such a full budget of material flows has been established, a set of functional relationships linking those fluxes can be calibrated,

which then allows GRAFS to be used as a predictive model for the construction of scenarios based on some hypotheses regarding possible changes in the agro-food system.

The most important of these relationships is the one linking harvested yield of the cropping system ( $Y$ , expressed in nitrogen, kgN/ha/yr) to the total N soil input ( $F$ , kgN/ha/yr) whatever its form (synthetic fertilizers, manure, atmospheric deposition or symbiotic fixation by legume crops). It has been shown that a robust relationship exists between  $Y$  and  $F$ , integrated over the full crop rotation cycle under

given pedo-climatic conditions, whatever the cropping system, either organic or conventional [Lassaletta et al., 2014, Anglade et al., 2015, Le Noë et al., 2017, Billen et al., 2018a]. This yield-fertilization relationship follows a hyperbolic function with a single parameter ( $Y_{\max}$ , in kgN/ha/yr):

$$Y = Y_{\max} \cdot F / (F + Y_{\max}) \quad (1)$$

$Y_{\max}$  thus characterizes the general fertility of the territory. The level of fertilization, with respect to  $Y_{\max}$ , determines the intensity of fertilization (IF, dimensionless):

$$IF = F / Y_{\max} \quad (2)$$

Related to the shape of the  $Y$  versus  $F$  relationship, all fertilizing inputs to the soil are not retrieved in the harvested yield. The difference defines the surplus (S), or N soil balance. The value of the surplus, which is quadratically related to fertilizing intensity, is a good indicator of total N environmental losses.

$$S = F - Y = Y_{\max} \cdot IF^2 / (IF + 1) \quad (3)$$

The fate of the surplus includes denitrification, sequestration in the soil organic matter pool (when it is aggrading) and leaching. In the case of cropland, the latter is most often the larger share, except in situations where, owing e.g., to water logging, soil denitrification is particularly high, or, due to increased inputs of high C/N residues to the soil, significant sequestration of nitrogen occurs in the soil organic matter.

The second important set of relationships in the GRAFS model link the flow of material through livestock, relating edible production (edProd, in ktN/yr), excretion (Excr, in ktN/yr) and ingestion (Ingest, in ktN/yr), with only two dimensionless parameters, calibrated from observed data in real livestock system, the conversion efficiency of vegetal feed into edible animal products ( $cveff$ ) and the non-edible to edible production ratio ( $ned:ed$ ):

$$eProd = cveff \cdot Ingest$$

$$Ingest = edProd + Excr + nedProd$$

$$ned:edr = nedProd / edProd.$$

The GRAFS model has been widely used to describe the current agro-food system at different scales, such as the small catchment scale [Garnier et al., 2016], the Seine basin [Billen et al., 2020b], the French agricultural regions [Le Noë et al., 2017, 2018], Europe [Billen et al., 2021, 2022], or the world [Billen et al.,

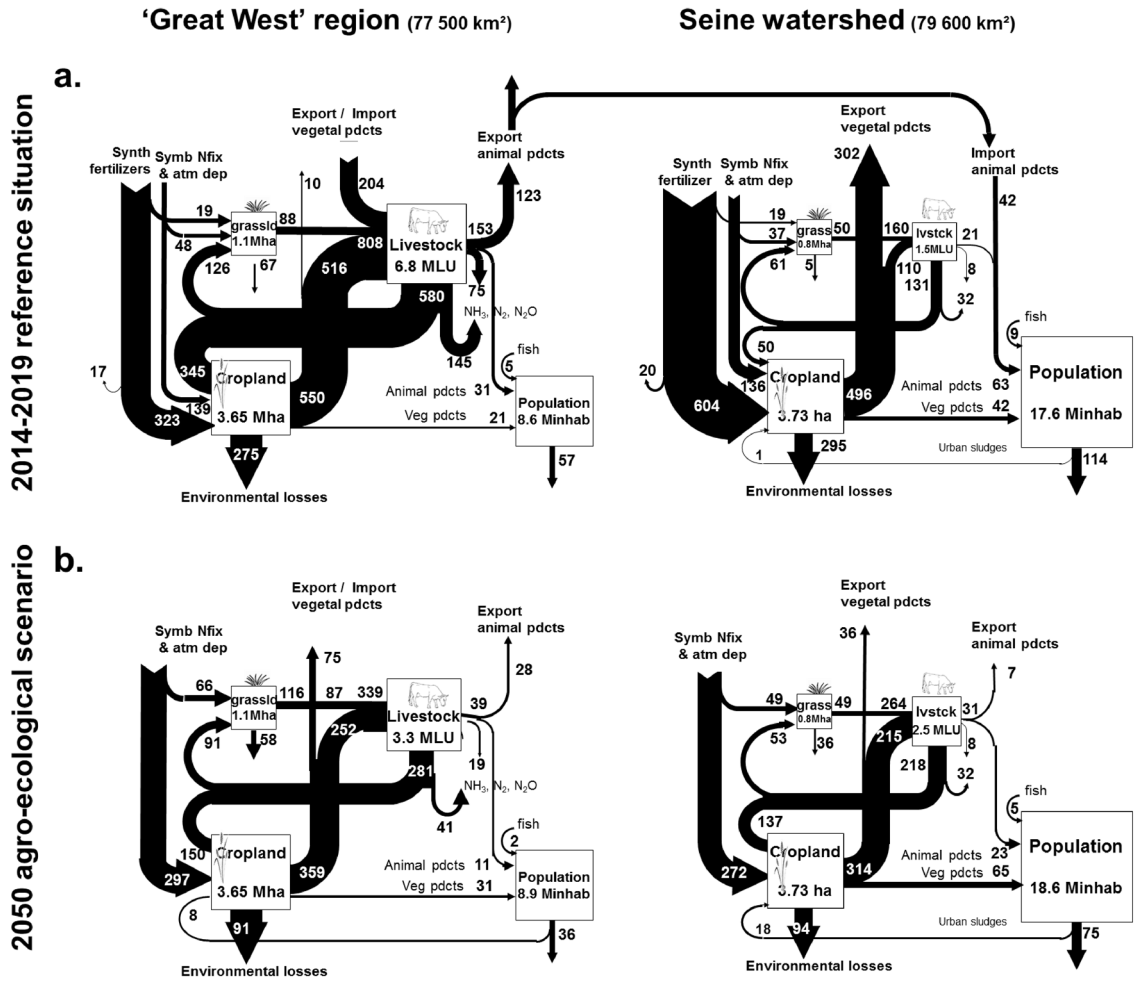
2013, Lassaletta et al., 2016]. It has also been used to interpret the long term past trajectory of the agro-food system of territories. In the case of the Seine watershed and France, the gradual territorial specialization into either intensive cereal cropping systems without livestock farming or intensive specialized livestock systems importing a large share of the feed had been identified as the cause of the opening of the nutrient cycles and increasing N losses to the environment (Figure 2). Similar conclusions have been drawn for other regions [Goyette et al., 2016] and at the global scale [Lassaletta et al., 2016].

GRAFS also calculates the C flows through the agro-food system, including those associated with the non-harvested aerial or underground residues that are returned to the soil and feed the organic carbon pool of the soil, making it possible to address the issue of long term soil C sequestration by coupling GRAFS with a soil C dynamics model such as the AMG model [Le Noë et al., 2019b,a, Garnier et al., 2022a].

For gaseous N losses, empirical relationships with farming practices and climate variables have been established for calculating  $NH_3$  volatilization [Sanz-Cobena et al., 2014] and  $N_2O$  emission [Garnier et al., 2019a]. As GRAFS calculates N surplus and leaching in all land use of a territory, it can be coupled with the Riverstrahler model for calculating the water quality resulting from a given operation of the agro-food system, allowing retrospective reconstruction of the past water quality [Billen et al., 2007].

#### 4. Predicting the risk of coastal water eutrophication

Coastal phytoplanktonic production is controlled by the interplay of several factors including its morphology and hydrological properties as well as nutrient delivery from land base sources or direct discharge. An unbalanced availability of nutrients with nitrogen and phosphorus in excess over silica can lead to the preferential development of non-siliceous, often harmful, algal species rather than of diatoms which initiate linear trophic chains leading to fish [Billen and Garnier, 2007, Garnier et al., 2010, 2021]. Such *new* production of non-siliceous algae, normally restricted to *regenerated* production, is the common characteristics of coastal eutrophication which can



**Figure 2.** N flows (in ktonN/yr) through the current agro-food systems (a) and in a prospective agro-ecological scenario at the 2050 horizon (b) of the Seine watershed and the Great West (Brittany, Basse Normandie, Pays de la Loire) region. Both regions have a very close surface area. The former is characterized by a much higher population and agriculture oriented toward cereal production for export. The latter is more oriented toward intensive livestock farming, with animal products exported to the rest of France. (All flux values are extracted from the study at European scale by Billen et al., 2022.)

results in very diverse manifestations such as the accumulation of mucilaginous material, toxin production, or oxygen deficits caused by degradation of accumulated unpalatable algal biomass.

The ICEP has been proposed by Billen and Garnier [2007] as an indicator of the eutrophication potential of nutrient delivery by river watersheds. It is defined as the excess of N or P over Si with respect to the requirements of balanced diatoms growth (according to Redfield ratios) and is expressed in kgC/day/km<sup>2</sup> (i.e. an algal biomass potentially produced per km<sup>2</sup>

of watershed) for a direct comparison of N-ICEP and P-ICEP among river systems:

$$N-ICEP \text{ (kgC/d/km}^2\text{)} = [\text{FlxN (ktonN/yr)/14} * 16 - \text{FlxSi (ktonSi/yr)/28} * 20] * 106 * 12/365/\text{Wa (km}^2\text{)}$$

$$P-ICEP \text{ (kgC/d/km}^2\text{)} = [\text{FlxP (ktonP/yr)/31} - \text{FlxSi (ktonSi/yr)/28} * 20] * 106 * 12/365/\text{Wa (km}^2\text{)}$$

where FlxN, FlxP and FlxSi (kton/yr) are the average annual river loads of nitrogen, phosphorus and silica



respectively, and  $Wa$  ( $\text{km}^2$ ) the watershed area.

N- or P-ICEP is preferred according to the most limiting element in the marine environment.

The ICEP permits to assess the potential for non-siliceous algal growth to be sustained by N and P in excess over silica in the nutrient load of a given river and can easily be calculated by the Riverstrahler model, at the outlet of large rivers discharging into coastal seas (Figure 3a).

The ICEP indicator has recently been included as an indicator for marine eutrophication in the UN Sustainable Development Goals (SDG14: <https://sdg.iisd.org/news/ioc-unesco-provides-update-on-sdg-14-indicator-development/>).

The ICEP indicator does not take into account any characteristics of the receiving coastal zone. Yet, according to their morphological and hydrological peculiarities, the pelagic ecosystem of different bays can react quite differently to a given loading of nutrients. For this reason, a derived indicator, the B\_ICEP (for Bay integrated indicator of coastal eutrophication potential) was further developed, combining ICEP with some characteristics of the specific receiving Bay [Garnier et al., 2021]. It is defined as the ratio of the riverine flux of nutrient in excess over silica to the volume of the receiving bay multiplied by its flushing rate.

$$B\_ICEP \text{ (mgC/L)} = [ICEP \text{ (kgC/d/km}^2) * Wa \text{ (km}^2) * 1000] / [Volb \text{ (m}^3) \cdot fr \text{ (d}^{-1})]$$

where  $Wa$  is the watershed area ( $\text{km}^2$ ),  $Volb$  ( $\text{m}^3$ ) is the volume of the receiving bay, and  $fr$  the flushing rate of the bay by both marine currents from the surrounding sea water bodies and by the river flow. The flushing rate ( $fr$ ) of a given bay can be empirically estimated based on its average salinity  $SalB$ (‰) compared with that of the surrounding seawater bodies.

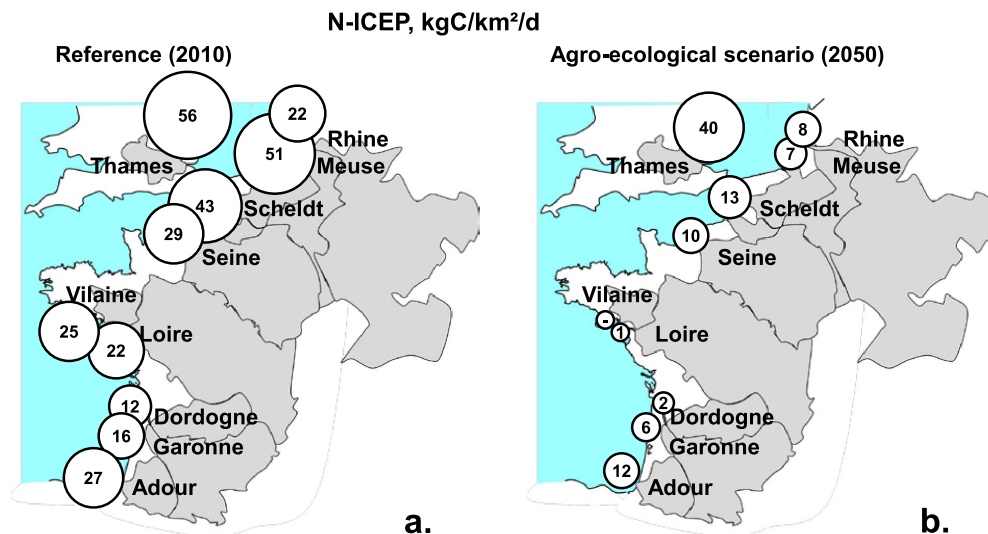
So defined, the B-ICEP represents the maximum concentration of non-siliceous algae which can be developed based on excess riverine N and P over Si in a particular marine bay. The pertinence of the B\_ICEP approach, i.e., the ability of this indicator to predict the order of magnitude of undesirable algal blooms in particular bays, has been demonstrated by using an idealized model of marine algal development [Garnier et al., 2021].

## 5. Integrated agro-ecological scenarios of future water-agro-food systems

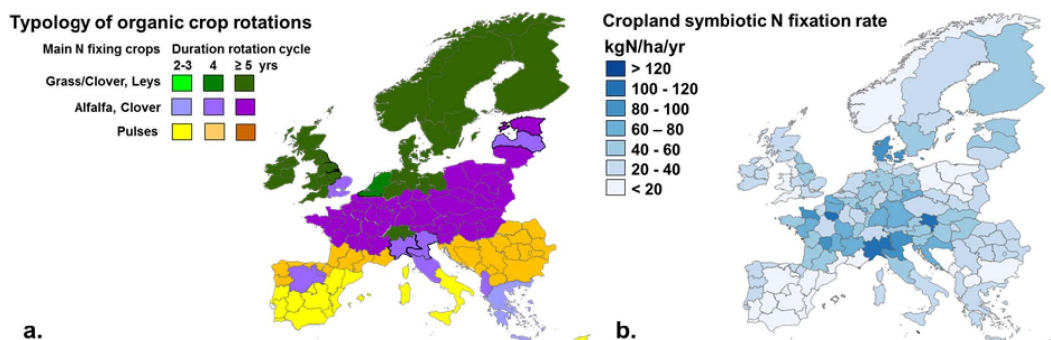
The modelling chain GRAFS—Riverstrahler—B\_ICEP has the potential for an integrated assessment of water quality over the whole continuum from headwaters to coastal sea resulting from a given scenario of the agro-food system. Here, as an example of the approach, we will explore some results from an extreme scenario which makes the assumption of a generalization of agro-ecological practices in agriculture. The purpose is to assess the capacity of such a scenario, which would consider organic farming without recourse to synthetic fertilizer, to feed the human population, as well as its effect on water quality and coastal eutrophication, and its ability to reduce greenhouse gas emission.

Three major levers have been combined in conceiving the scenario. The first one concerns human diet. WHO [2019] considers a protein intake of 3.4 kgN-protein per capita per year, as the vital minimum requirement. Values twice as high are observed in western countries, and have been associated to important health problems [Willet et al., 2019]. Importantly, the share of animal-based protein consumption varies between less than 10% in some African countries to as much as 70% in rich western countries in Europe and the USA. The EAT-Lancet Commission recently recommended a Reference Healthy Diet which, compared to the current European diet, implies a strong reduction not only of total protein intake but also of the share of animal products Willet et al. [2019] The assumption of the agro-ecological scenario presented here is very close to the EAT-Lancet one (Table 1).

The second lever of the agro-ecological scenario is the adoption of organic farming practices, strictly banning the use of synthetic inputs (fertilizers and pesticides). There is large diversity of such practices world-wide, often adjusted by the experience of farmers to territorial climatic peculiarities [Altieri, 2002, Altieri and Nicholls, 2017, Compagnone et al., 2018, Garnier et al., 2016, Billen et al., 2021]. All of them use long and diversified crop rotations leaving ample room to grain and forage legumes, the symbiotic  $N_2$  fixation of which constitutes the main source of new fertilizing nitrogen input to the whole rotation. An extensive review of organic systems effectively in use in Europe led to a typology of these crop



**Figure 3.** Values, linked to different sizes of circles for ICEP (Indicator of Coastal Eutrophication Potential, in kgC/km<sup>2</sup>/day) at the outlet of the main rivers discharging along the North Atlantic façade. (a) Current situation (2010); (b) in the agro-ecological scenario. (The data are extracted from the Emosem program, Desmit et al., 2018.)



**Figure 4.** (a) Typology of organic crop rotation in Europe [Billen et al., 2021] and (b) associated symbiotic N<sub>2</sub> fixation capacity [Billen et al., 2022].

rotations, and to an estimate of their N<sub>2</sub> fixation capacity (Figure 4).

The third lever to be operated in such a scenario concerns the reconnection of crop and livestock farming. Animals have not only the function of providing meat and milk to human: they are also the agent able to convey nutrients from legumes to other crops and from grassland to cropland. In our scenario, livestock density in each region is determined by the available resources of grass and forage crop production, with no recourse to feed imports from

outside.

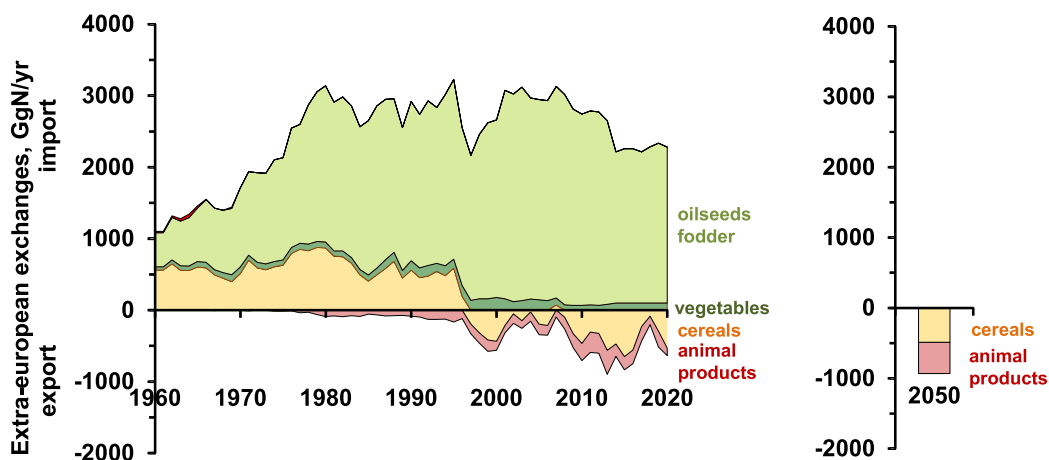
The results of the application of this scenario to the two contrasting regions taken as examples above, the Seine basin and the Great West, are illustrated in Figure 3b. The scenario would lead to reintroduction of livestock in the center of the Paris basin, while its density would be strongly reduced in Brittany. In the scenario, both regions would feed their human and livestock populations without import of feed and would still export ample food surpluses. Similarly, at the scale of France, this scenario predicts that

**Table 1.** Current average European diet per day (d) or per year (yr) in terms of apparent consumption [Westhoek et al., 2015] and prospective diet considered in the agro-ecological scenario exposed in this paper at the horizon 2050, close to the EAT-Lancet reference healthy diet

			Average current European diet				Prospective healthy diet 2050			
	kcal/g	%N	g/d	kcal/d	kgN/cap/yr	%	g/d	kcal/d	kgN/cap/yr	%
Vegetal products					2.35	41			3.5	70
Cereals	3.5	2	240	840	1.75		285	998	2.08	
Grain legumes	1	3.5	15	15	0.19		59	59	0.75	
Roots and tubers	0.8	0.25	100	80	0.09		80	64	0.07	
Fresh vegetables	0.3	0.3	150	45	0.16		300	90	0.33	
Fresh fruits	0.3	0.15	140	42	0.08		200	60	0.11	
Nuts	6.5	1	20	130	0.07		50	325	0.18	
Animal products (excluding fish and seafood)					3.15	54			1.25	25
Milk products	0.8	0.7	565	452	1.44		224	179	0.57	
Meat	1.5	3.25	125	188	1.48		50	74	0.59	
Eggs	1.4	2	30	42	0.22		12	17	0.09	
Fish & seafood	1	2.9	27	27	0.29	5	23.5	24	0.25	5
Products without N					[kgNeq/cap/yr]*				[kgNeq/cap/yr]*	
Added sugar	3	0	50	150	0.24*		20	60	0.09*	
Oil	7	0	40	280	1.10*		40	280	1.10*	
Total			1502	2291	5.78	100	1343	2229	5.0	100

Europe refers here to the EU28 plus UK, non EU ex-Yugoslavian countries, Norway and Switzerland.

\*Nequivalent in corresponding harvested product before extraction of oil or sugar (0.075 equN/kg oil and 0.013 kgN/kg sugar).



**Figure 5.** Extra-European exchanges of cereals, animal products, oilseed fodder and vegetables from 1961 to 2020 [Billen et al., 2021 and FAOstat data] and in the agro-ecological scenario for 2050 [Billen et al., 2022].

French exportation could still amount 220 ktN/yr as vegetable products (40% of the current value) and 3 ktonN/yr as animal products (10% of current ex-

ports), without importing any soybean based feed [Billen et al., 2018a]. For Europe the conclusions are similar: the agro-ecological scenario would be able

to feed the population without synthetic N fertilizers while still exporting cereals and animal products at rates similar to current levels to the rest of the world (Figure 5).

At the global scale, the possibility of feeding a population of 10 billion people by agro-ecological agriculture, with no deforestation, and less recourse to international trade (hence more food sovereignty) was demonstrated by several authors [Billen et al., 2015, Lassaletta et al., 2016, Erb et al., 2016], provided that a diet not exceeding 30–40% of animal products in the total protein ingestion is adopted everywhere. This work thus demonstrate that, contrarily to what is sometimes reported [Connor, 2013, 2018, Barbieri et al., 2021], agro-ecological practices can feed the world provided they are combined with dietary changes and structural reorganization of the relationships between crop and livestock farming, allowing to close the nitrogen cycle.

The effect of this scenario on water quality can be assessed by comparing the leaching water nitrate concentration calculated by the GRAFS model between the current and the scenarized situation (Figure 6). Further, coupling GRAFS and Riverstrahler models shows how nitrate concentration would be reduced in the drainage network of the different rivers of the Western Atlantic façade and English Channel [Desmit et al., 2018, Garnier et al., 2019b, 2022b], reducing the risks of coastal eutrophication (Figure 4b).

Greenhouse gas emissions by agriculture are also predicted to be drastically reduced in this agro-ecological scenario, owing to reduced livestock numbers (hence reduction of methane emissions) and suppression of synthetic N fertilizers (hence reduction of CO<sub>2</sub> emissions for fertilizer synthesis and N<sub>2</sub>O emissions from cropland (Figure 7).

For the Seine Basin, the Great West and France as a whole, the contribution of agriculture to total greenhouse gas emissions in terms of CO<sub>2</sub> equivalent is estimated to have reduced by 33%, 74% and 50% respectively [Garnier et al., 2019a, 2022a].

## 6. Conclusions

The water-agro-food continuum concept, gradually developed in the past 30 years, was a guide for dealing with the tremendous complexity of the intertwined challenges of food production, international

trade, environmental quality and climate changes. It stresses the central role that agriculture and food supply play in a globalized world. It has helped conceiving a vision of our future where these challenges would be met altogether. These challenges are global in nature as they ultimately concern the whole planet. However, the solutions, just like the research on which they are based, needs to be rooted in territories. The Seine watershed, from that perspective was a very fruitful case study, because it is a paradigmatic example of a territory organized around a huge metropolis of global significance, as well as a highly specialized agricultural area with extremely intensive farming practices. This is likely to be one of the main reasons for the exceptional longevity and productivity of the PIREN-Seine program.

## Conflicts of interest

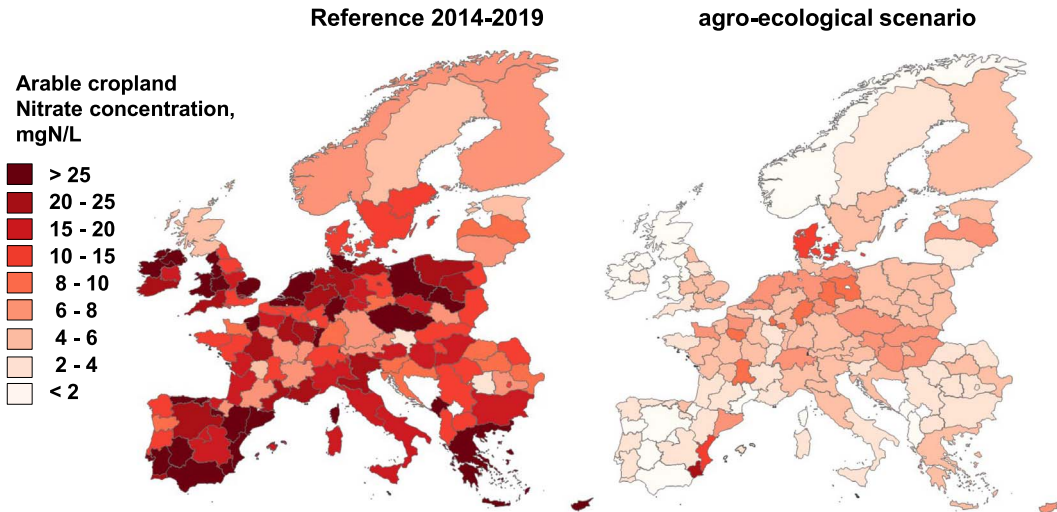
Authors have no conflict of interest to declare.

## Acknowledgments

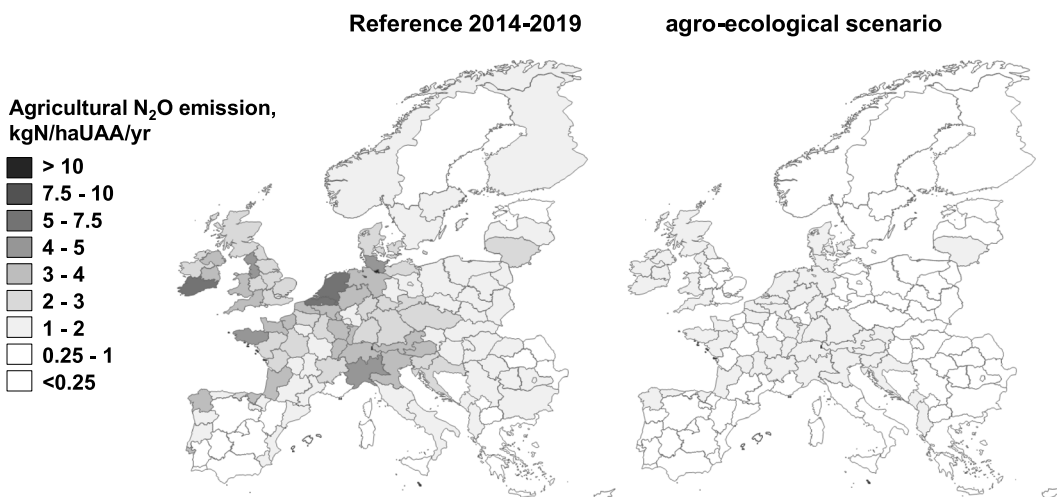
The authors wish to express their deep gratitude to all their students and collaborators who, since three decades, were involved in the gradual elaboration of the approaches exposed in this paper. Ghislain de Marsily welcomed both of us in his lab where he ensured the best scientific environment for developing creative and interdisciplinary research. We are extremely grateful to him for his openness and attention. His catchphrase was: *“I let the researchers do the research they want, because that is what they do best!”*.

## References

- Altieri, M. A. (2002). Agroecology: the science of natural resource management for poor farmers in marginal environments. *Agric. Ecosyst. Environ.*, 93, 1–24.
- Altieri, M. A. and Nicholls, C. I. (2017). The adaptation and mitigation potential of traditional agriculture in a changing climate. *Climatic Change*, 140, 33–45.
- Anglade, J., Billen, G., Makridis, T., Garnier, J., Puech, T., and Tittel, C. (2015). Nitrogen soil surface balance of organic vs conventional cash crop



**Figure 6.** Cropland leaching water N concentration (in mgN/l) calculated from N surplus and runoff in the different European regions in 2014–2019 and in the agro-ecological scenario [data from Billen et al., 2022].



**Figure 7.** Total agricultural N<sub>2</sub>O emission in the different European regions in 2014–2019 and in the agro-ecological scenario [data from Billen et al., 2022].

farming in the Seine watershed. *Agric. Syst.*, 139, 82–92.

Barbieri, P., Pellerin, S., Seufert, V., Smith, L., Ramankutty, R., and Nesme, T. (2021). Global option space for organic agriculture is delimited by nitrogen availability. *Nat. Food*, 2, 363–372.

Barles, S., Buclet, N., and Billen, G. (2011). L'écologie territoriale : du métabolisme des sociétés à la gou-

vernance des flux d'énergie et de matières. In *CIST2011—Fonder les sciences du territoire*, pages 16–22. Collège international des sciences du territoire (CIST), Paris, France.

Beaudoin, N., Venet, E., Maucorps, J., Vandenberghe, C., Pugeaux, N., Viennot, P., Gourcy, L., Brayer, C., Machet, J. M., Couturier, A., Billy, C., Vigour, N., Hulin, G., Dorel, G., and Mary, B. (2021). Long

- term response of water and nitrogen fluxes to good agricultural practices at field and catchment scales. *Sci. Total Environ.*, 776, article no. 145954.
- Billen, G. (2001). Le PIREN-Seine : un programme de recherche né du dialogue entre scientifiques et gestionnaires. *Revue pour l'Histoire du CNRS*, 4, 86–91.
- Billen, G., Aguilera, E., Einarsson, R., Garnier, J., Gingrich, S., Grizzetti, B., Lassaletta, L., Le Noë, L., and Sanz-Cobena, A. (2021). Reshaping the European agro-food system and closing its nitrogen cycle: the potential of combining dietary change, agroecology, and circularity. *One Earth*, 4, 839–850.
- Billen, G., Aguilera, E., Einarsson, R., Garnier, J., Gingrich, S., Lassaletta, L., Le Noë, J., and Sanz-Cobena, A. (2022). European GreenDeal Scenario Project 2. Final Report. EU Joint Resarch Center, Ispra.
- Billen, G. and Garnier, J. (2007). River basin nutrient delivery to the coastal sea: assessing its potential to sustain new production of non siliceous algae. *Mar. Chem.*, 106, 148–160.
- Billen, G. and Garnier, J. (2021). Nitrogen biogeochemistry of water-agro-food systems: the example of the Seine land-to-sea continuum. *Biogeochemistry*, 154, 307–321.
- Billen, G., Garnier, J., Grossel, A., Thieu, V., Théry, S., and Hénault, C. (2020a). Modeling indirect N<sub>2</sub>O emissions along the N cascade from cropland soils to rivers. *Biogeochemistry*, 148, 207–221.
- Billen, G., Garnier, J., and Hanset, P. (1994). Modelling phytoplankton development in whole drainage networks: The RIVERSTRAHLER model applied to the Seine river system. *Hydrobiologia*, 289, 119–137.
- Billen, G., Garnier, J., and Lassaletta, L. (2013). The nitrogen cascade from agricultural soils to the sea: modelling N transfers at regional watershed and global scales. *Phil. Trans. R. Soc. B*, 368, article no. 20130123.
- Billen, G., Garnier, J., Le Noë, J., Viennot, P., Gallois, N., Puech, T., Schott, C., Anglade, J., Mary, B., Beaudoin, N., Léonard, J., Mignolet, C., Théry, S., Thieu, V., Silvestre, M., and Passy, P. (2020b). The Seine watershed water-agro-food system: Long-term trajectories of C, N and P metabolism. In Flipo, N., Labadie, P., and Lestel, L., editors, *The Seine River Basin*, The Handbook of Environmental Chemistry. Springer, Cham.
- Billen, G., Garnier, J., Mouchel, J.-M., and Silvestre, M. (2007). The Seine System: introduction to a multidisciplinary approach of the functioning of a regional river system. *Sci. Total Environ.*, 375, 1–12.
- Billen, G., Garnier, J., and Rousseau, V. (2005). Nutrient fluxes and water quality in the drainage network of the Scheldt basin over the last 50 years. *Hydrobiologia*, 540, 47–67.
- Billen, G., Lassaletta, L., and Garnier, J. (2015). A vast range of opportunities for feeding the world in 2050: trade-off between diet, N contamination and international trade. *Environ. Res. Lett.*, 10, article no. 025001.
- Billen, G., Le Noë, J., and Garnier, J. (2018a). Two contrasted future scenarios for the French agro-food system. *Sci. Total Environ.*, 637–638, 695–705.
- Billen, G., Ramarson, A., Thieu, V., Théry, S., Silvestre, M., Pasquier, C., Hénault, C., and Garnier, J. (2018b). Nitrate retention at the river–watershed interface: a new conceptual modeling approach. *Biogeochemistry*, 139, 31–51.
- Brion, N. and Billen, G. (2000). Wastewater as a source of nitrifying bacteria in river systems: the case of River Seine downstream from Paris. *Water Res.*, 34, 3213–3221.
- Brion, N., Billen, G., Guézennec, L., and Ficht, A. (2000). Distribution of nitrifying activity in the Seine River (France) from Paris to the estuary. *Estuaries*, 23, 669–682.
- Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, D., Bertuzzi, P., Burger, P., Bussière, F., Cabidoche, Y. M., Cellier, P., Debaeke, P., Gaudillère, J. P., Hénault, C., Maraux, F., Seguin, B., and Sinoquet, H. (2003). An overview of the crop model STICS. *Eur. J. Agron.*, 18, 309–32.
- Buclet, N. (2015). *Essai d'Ecologie Territoriale: l'exemple d'Aussois en Savoie*. CNRS editions, Paris.
- Chestérikoff, A., Garban, B., Billen, G., and Poulin, M. (1992). Inorganic nitrogen dynamics in the river Seine downstream from Paris (France). *Biogeochemistry*, 17, 147–164.
- Compagnone, C., Lamine, C., and Dupré, L. (2018). La production et la circulation des connaissances en agriculture interrogées par l'agro-écologie. *Revue d'anthropologie des connaissances*, 12, 111–138.
- Connor, D. J. (2013). Organically grown crops do not a cropping system make and nor can organic agriculture nearly feed the world. *Field Crops Res.*

- 144, 145–147.
- Connor, D. J. (2018). Organic agriculture and food security: A decade of unreason finally implodes. *Field Crops Res.*, 225, 128–129.
- Cugier, P., Billen, G., Guillaud, J. F., Garnier, J., and Ménesguen, A. (2005). Modelling eutrophication of the Seine Bight under present, historical and future Seine river nutrient loads. *J. Hydrol.*, 304, 381–396.
- Desmit, X., Thieu, V., Dulière, V., Ménesguen, A., Campuzano, F., Lassaletta, L., Sobrinho, J. L., Silvestre, M., Garnier, J., Neves, R., Billen, G., and Lacroix, G. (2018). Reducing marine eutrophication may require a paradigmatic change. *Sci. Total Environ.*, 635, 1444–1466.
- Erb, K.-H., Lauk, C., Kastner, T., Mayer, A., Theurl, M. C., and Haberl, H. (2016). Exploring the biophysical option space for feeding the world without deforestation. *Nat. Commun.*, 7, article no. 11382.
- Even, S., Poulin, M., Garnier, J., Billen, G., Servais, P., Chestérikoff, A., and Coste, M. (1998). River ecosystem modelling: Application of the PROSE model to the Seine river (France). *Hydrobiologia*, 373/374, 27–45.
- Garnier, J., Anglade, J., Benoit, M., Billen, G., Puech, T., Ramarson, A., Passy, P., Silvestre, M., Lassaletta, L., Trommenschlager, J.-M., Schott, C., and Tallec, G. (2016). Reconnecting crop and cattle farming to reduce nitrogen losses in river water of an intensive agricultural catchment (Seine basin, France). *Environ. Sci. Policy*, 63, 76–90.
- Garnier, J., Beusen, A., Thieu, V., Billen, G., and Bouwman, L. (2010). N:P:Si nutrient export ratios and ecological consequences in coastal seas evaluated by the ICEP approach. *Global Biogeochem. Cycles*, 24, article no. GB0A05. Special issue: Past and Future Trends in Nutrient Export from Global Watersheds and Impacts on Water Quality and Eutrophication.
- Garnier, J., Billen, G., and Coste, M. (1995). Seasonal succession of diatoms and chlorophyceae in the drainage network of the River Seine: Observations and modelling. *Limnol. Oceanogr.*, 40, 750–765.
- Garnier, J., Billen, G., and d Cébron, A. (2007). Modelling nitrogen transformations in the lower Seine river and estuary (France): impact of waste water release on oxygenation and N<sub>2</sub>O emission. *Hydrobiologia*, 588, 291–302.
- Garnier, J., Billen, G., Hannon, E., Fonbonne, S., Videna, Y., and Soulie, M. (2002). Modeling transfer and retention of nutrients in the drainage network of the Danube River. *Estuar. Coast. Shelf Sci.*, 54, 285–308.
- Garnier, J., Billen, G., Lassaletta, L., Vigiak, O., Nikolaidis, N. P., and Grizzetti, B. (2021). Hydromorphology of coastal zone and structure of watershed agro-food system are main determinants of coastal eutrophication. *Environ. Res. Lett.*, 16, article no. 023005.
- Garnier, J., Billen, G., Sanchez, N., and Leporcq, B. (2000). Ecological functioning of a large reservoir in the upstream basin of the river Seine (Marne reservoir, France). *Regul. Rivers Res. Manage.*, 16, 51–71.
- Garnier, J., Billen, G., Tournebize, J., Barré, P., Mary, B., and Baudin, F. (2022a). Storage or loss of soil active carbon in cropland soils: The effect of agricultural practices and hydrology. *Geoderma*, 407, article no. 115538.
- Garnier, J., Lassaletta, L., Billen, G., Romero, E., Grizzetti, B., Némery, J., Le, Q. L. P., Pistocchi, C., Aissa-Grouz, N., Luu, M. T. N., Vilmin, L., and Dorioz, J.-M. (2015). Phosphorus budget in the water-agro-food system at nested scales in two contrasted regions of the world (ASEAN-8 and EU-27). *Global Biogeochem. Cycles*, 29, 1348–1368.
- Garnier, J., Le Noë, J., Marescaux, A., Sanz-Cobena, A., Lassaletta, L., Silvestre, M., Thieu, V., and Billen, G. (2019a). Long-term changes in greenhouse gas emissions from French agriculture and livestock (1852–2014): from traditional agriculture to conventional intensive systems. *Sci. Total Environ.*, 660, 1486–1501.
- Garnier, J., Meybeck, M., Ayrault, S., Billen, G., Blanchoud, H., Carré, C., Flipo, N., Gasperi, J., Lestel, L., de Marsily, G., Mouchel, J. M., Servais, P., and Tales, E. (2022b). Chapter 5.2—Continental Atlantic rivers: the Seine Basin. In Tockner, K., Zarfl, C., and Robinson, C. T., editors, *Rivers of Europe*, pages 291–330. Elsevier, 2nd edition. ISBN 9780081026120.
- Garnier, J., Némery, J., Billen, G., and Théry, S. (2005). Nutrient dynamics and control of eutrophication in the Marne River system: modelling the role of exchangeable phosphorus. *J. Hydrol.*, 304, 397–412.
- Garnier, J., Ramarson, A., Billen, G., Théry, S., Thiéry, D., Thieu, V., Minaudo, C., and Moatar, F. (2018a).

- Nutrient inputs and hydrology together determine biogeochemical status of the Loire River (France): current situation and possible future scenarios. *Sci. Total Environ.*, 637–638, 609–624.
- Garnier, J., Ramarson, A., Thieu, V., Némery, J., Théry, S., Billen, G., and Coynel, A. (2018b). How can water quality be improved when the urban waste water directive has been fulfilled? A case study of the Lot river (France). *Environ. Sci. Pollut. Res.*, 25, 11924–11939.
- Garnier, J., Riou, P., Le Gendre, R., Ramarson, A., Billen, G., Cugier, P., Schapira, M., Théry, S., Thieu, V., and Ménesguen, A. (2019b). Managing the agro-food system of watersheds to combat coastal eutrophication: A land-to-sea modelling approach to the French coastal English channel. *Geosciences*, 9, article no. 441.
- Gérardin, A. (1875). Altération de la Seine aux abords de Paris depuis novembre 1874 jusqu'à mai 1875. *C. R. Acad. Sci. Paris*, 80, 1326–1328.
- Goyette, J.-O., Bennett, E. M., Howarth, R. W., and Maranger, R. (2016). Changes in anthropogenic nitrogen and phosphorus inputs to the St. Lawrence sub-basin over 110 years and impacts on riverine export. *Glob. Biogeochem. Cycles*, 30, 1000–1014.
- Lassaletta, L., Billen, G., Garnier, J., Bouwman, L., Velazquez, E., Mueller, N. D., and Gerber, J. S. (2016). Nitrogen use in the global food system: Past trends and future trajectories of agronomic performance, pollution, trade, and dietary demand. *Environ. Res. Lett.*, 11, article no. 095007.
- Lassaletta, L., Billen, G., Grizzetti, B., Garnier, J., Leach, A. M., and Galloway, J. N. (2014). Food and feed trade as a driver in the global nitrogen cycle: 50-year trends. *Biogeochemistry*, 118, 225–241.
- Le, T. P. Q., Billen, G., Garnier, J., and Chau Van, M. (2015). Long-term evolution of the biogeochemical functioning of the Red River (Vietnam): past and present situations. *Reg. Environ. Change*, 15, 329–339.
- Le Noë, J., Billen, G., Esculier, F., and Garnier, J. (2018). Long-term socioecological trajectories of agro-food systems revealed by N and P flows in French regions from 1852 to 2014. *Agr. Ecosyst. Env.*, 265, 132–143.
- Le Noë, J., Billen, G., and Garnier, J. (2017). How the structure of agro-food systems shapes nitrogen, phosphorus, and carbon fluxes: the generalized representation of agro-food system applied at the regional scale in France. *Sci. Total Environ.*, 586, 42–55.
- Le Noë, J., Billen, G., and Garnier, J. (2019a). Carbon dioxide emission and soil sequestration for the French agro-food system: present and prospective scenarios. *Front. Sustain. Food Syst.*, 3, article no. 19.
- Le Noë, J., Billen, G., Mary, B., and Garnier, J. (2019b). Drivers of long-term carbon dynamics in cropland: a bio-political history (France, 1852–2014). *Environ. Sci. Policy*, 93, 53–65.
- Ledoux, E., Gomez, E., Monget, J.-M., Viavattene, C., Viennot, P., Ducharme, A., Benoit, M., Mignolet, C., Schott, C., and Mary, B. (2007). Agriculture and groundwater nitrate contamination in the Seine Basin. The STICS-MODCOU modelling chain. *Sci. Total Environ.*, 375(1–3), 33–47.
- Lesouëf, A. and André, A. (1982). Mise au point d'un modèle de qualité de la Seine de Montereau à Poses. In *XVIIèmes Journées de l'Hydraulique. Nantes sept 1982. L'assainissement de demain, hydraulique des eaux pluviales et usées*. Société Hydrotechnique de France.
- Marescaux, A., Thieu, V., Gypens, N., Silvestre, M., and Garnier, J. (2020). Modeling the inorganic carbon dynamics in the Seine River continuum in France. *Hydrol. Earth Syst. Sci.*, 24, 1–20.
- Meybeck, M., de Marsily, G., and Fustec, E. (1998). *La Seine en son bassin : fonctionnement écologique d'un système fluvial anthropisé*. Elsevier, Masson.
- Minaudo, C., Curie, F., Jullian, Y., Gassama, N., and Moatar, F. (2018). QUAL-NET, a high temporal-resolution eutrophication model for large hydrographic networks. *Biogeosciences*, 15, 2251–2269.
- Neitsch, S. L., Arnold, J. G., Kiniry, J. R., and Williams, J. R. (2001). *Soil and Water Assessment Tool-Use Manual Version 2000*. Blackland Research Center, Agricultural Research Service, Texas, USA.
- Raimonet, M., Thieu, V., Silvestre, M., Oudin, L., Rabouille, C., Vautard, R., and Garnier, J. (2018). Coastal eutrophication potential under future climate change: a landward perspective. *Front. Mar. Sci. Sec. Mar. Biogeochem.*, 5, article no. 136.
- Rockström, J., Steffen, W., Noone, K., Persson, A., Chapin, F. S., Lambin, E. F., Lenton, T. M., Scheffer, M., Folke, C., Schellnhuber, H. J., Nykvist, B., de Wit, C. A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P. K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R. W., Fabry,



- V. J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, and Foley, J. A. (2009). A safe operating space for humanity. *Nature*, 461, 472–475.
- Sanz-Cobena, A., Lassaletta, L., Estelles, F., Del Prado, A., Gurdia, G., Abalos, D., Aguilera, E., Pardo, G., Vallejo, A., Sutton, M. A., Garnier, J., and Billen, G. (2014). Yield scaled mitigation of ammonia emission from N fertilization: the Spanish Case. *Environ. Res. Lett.*, 9, article no. 1250005.
- Sferratore, A., Billen, G., Garnier, J., Smedberg, E., Humborg, C., and Rahm, L. (2008). Modelling nutrient fluxes from sub-arctic basins: comparison of pristine versus dammed rivers. *J. Mar. Syst.*, 73, 236–249.
- Streeter, H. W. and Phelps, E. B. (1925). *Studies of the pollution and natural purification of the Ohio River, Part III, Factors concerned in the phenomena of oxidation and reareation*. Public Health bulletin No. 146. U.S. Public Health Service, Washington, DC.
- Thieu, V., Billen, G., and Garnier, J. (2009). Nutrient transfer in three contrasting NW European watersheds: The Seine, Somme, and Scheldt Rivers. A comparative application of the Seneque/Riverstrahler model. *Water Res.*, 43, 1740–1748.
- Westhoek, H. et al. (2015). Nitrogen on the table: the influence of food choices on nitrogen emissions, greenhouse gas emissions and land use in Europe. ENA special report on Nitrogen and Food. CEH, UK.
- WHO (2019). Sustainable healthy diets—Guiding principles. <http://www.fao.org/3/ca6640en/ca6640en.pdf>. Rome.
- Willet, W., Rockström, J., Loken, B., Springmann, M., Lang, T., Vermeulen, S., et al. (2019). Food in the Anthropocene: the EAT–Lancet Commission on healthy diets from sustainable food systems. *Lancet Commissions*, 393(10170), 447–492.
- Williams, J. R., Jones, C. A., Kiniry, J. R., and Spanel, D. A. (1989). The EPIC crop growth model. *Trans. ASAE*, 32(2), 497–511.





Research article

## Geo-hydrological Data & Models

# Rivers help us to quantify the socio-ecological functioning of their basin at the Anthropocene: the Seine example (1850–2020)

Michel Meybeck <sup>\*,a</sup>, Gabrielle Bouleau <sup>b</sup>, Catherine Carré <sup>c</sup>, Josette Garnier <sup>a</sup> and Laurence Lestel <sup>\*,a</sup>

<sup>a</sup> Sorbonne Université, CNRS EPHE, UMR Metis 7619, 4 place Jussieu, 75005, Paris, France

<sup>b</sup> UGE, CNRS, INRAE, LISIS, Marne-la-Vallée, France

<sup>c</sup> Université Paris 1 Panthéon-Sorbonne, LADYSS, Paris, France

*E-mails:* michel.meybeck@upmc.fr (M. Meybeck), gabrielle.bouleau@inrae.fr (G. Bouleau), catherine.carre@univ-paris1.fr (C. Carré), josette.garnier@upmc.fr (J. Garnier), laurence.lestel@sorbonne-universite.fr, laurence.lestel@upmc.fr (L. Lestel)

**Abstract.** River basins were identified very early on as a key component of chemical fluxes from continents to oceans, driven by weathering and biogeochemical cycles in their basin. Fifty years ago, important riverine changes attributed to human impacts started to be studied at the global scale, an evolution which has been the foundation for Anthropocene studies in hydrosystems since 2000. However, the co-evolution of river basin systems and sociosystems remains poorly addressed, particularly over long periods (>100 y). Medium-sized river basins make it possible to perform such detailed analyses, which combines historical river fluxes, material flows, river ecology, environmental history and political ecology.

Such an interdisciplinary approach at the scale of the Seine basin, undertaken by the PIREN-Seine research program, started in 1989, has shown how scientific and technological knowledge, environmental awareness, environmental regulations and policies, and political decisions have played a role on water quality during that period.

**Keywords.** Seine river, Socio-hydrosystem, Anthropocene, Environmental trajectories, River basin metabolism, Water quality.

*Manuscript received 2 May 2022, revised and accepted 30 June 2022.*

## 1. Introduction

### 1.1. PIREN-Seine creation and overall status

When the PIREN-Seine program was established in 1989 by Ghislain de Marsily, professor in hydrogeol-

ogy at Sorbonne Université, to enhance the knowledge on water resources and on river quality, together with many actors of the basin, none of us would have thought that more than three decades later the program would be more active than ever. At the beginning of PIREN Seine, interdisciplinarity meant for scientists and stakeholders the study of hydrogeology, biogeochemistry and ecology at the scale of the Seine River and its basin, with caveats from man-

\* Corresponding authors.

agers to integrate other disciplines such e.g., as sociology, climate science... Water managers faced a painful stalemate with Prime Minister Rocard refusing to expand the gigantic wastewater treatment plant located downstream of Paris. They welcomed the PIREN-Seine as a scientific forum for reallocating investments for pollution reduction [Bouleau, 2014]. Researchers and managers represented two separate communities, the former, academic, financially supporting the latter, technical. To this effect, G. de Marsily's main action was to set up and lead a coordination committee to pool knowledge between the various academic and operational institutions and thus initiated a societal and applied research for river hydrographic network, which was developed by French CNRS as the river component of the Programme Interdisciplinaire de Recherche en Environnement (PIREN), such as the PIREN-Rhône [Amoros *et al.*, 1987, Bravard *et al.*, 1986].

Few years later, the program opened to agricultural and social sciences, developed new research themes and tools at the scale of the entire Seine basin, by initiating concepts such as the man-land-sea continuum and water agro-food system [Billen and Garnier, 2021, Garnier *et al.*, 2021], and developed a long-term approach to carry out a prospective analysis. The program, accredited by the CNRS (the French fundamental research institution), already at the scale of the whole basin, from headwaters to the coast, expanded its topics and temporal scope to (i) to diverse environmental issues as greenhouse gases (GHG), carbon budgets, climate change impact, agricultural practices, consumer habits, emerging contaminants and (ii) to a temporal window up to two hundred years. We did not expect that this project would generate its own community over time, gathering within the PIREN-Seine annual forum scientists from various disciplines with managers and engineers of a dozen key institutions and companies, which have supported the program since its beginning, and that it would generate hundreds of environmental specialists. Our syntheses in the late 1990s [Garnier and Mouchel, 1999, Meybeck *et al.*, 1998], in 2007 [Billen *et al.*, 2007a] and in 2021 [Flipo *et al.*, 2021b, Garnier *et al.*, 2022] reflect this evolution.

## 1.2. *PIREN-Seine, Earth system science and the Anthropocene*

River basins are natural territorial scales to address sustainable water resources management and water governance [Pahl-Wostl *et al.*, 2012]. They are also major components of Earth System science [Kabat *et al.*, 2004], integrating lateral fluxes to oceans over the whole watershed. As such they can be used to quantify the changes occurring on the continental surface [Vörösmarty *et al.*, 2004]. When, in 2000, Paul Crutzen coined the term Anthropocene era, to qualify the global change that was observed on the Earth System due to human activities in the industrial period, the river and coastal ocean scientific communities were one of the very first to use that new concept [Luciano, 2022] and the Seine basin and its hydrosystem controlled by human activities was chosen to illustrate it [Meybeck, 2002, 2003, Meybeck and Vörösmarty, 2005, Vörösmarty and Meybeck, 2004].

## 1.3. *Evolution of the PIREN-Seine vision, from a hydrosystem to a socio-hydrosystem*

As early as 1989, we considered the Seine River basin as an hydrosystem, i.e. a well-bounded spatial entity, linking streams, river network, aquifers and reservoirs to the estuary, on which water balance and budgets of riverborne materials could be measured and modelled taking into account the inputs to surface water. We then gradually integrated how much it was impacted by multiple and interacting social and economic activities (17 million people, 76,000 km<sup>2</sup>), and regulated in many ways, generally by dedicated institutions, some of them established more than one hundred years ago. We also found that our analysis of present conditions had to consider the basin history over the "longue durée" (the long term, >100 y) [Billen *et al.*, 2007b, Le Noé *et al.*, 2018, Lestel and Carré, 2017]. Finally, we pointed out that this system was no longer a closed one, as in the classical hydrological approach, and that external inputs and exchanges (water, materials, information, technology, norms, aquatic species, etc.), with other territories had to be considered [e.g., Barles, 2010, Belliard *et al.*, 2018, Billen *et al.*, 2012, Bouleau *et al.*, 2017, Meybeck and Lestel, 2017]. We thus

have considered the technical, economic, and cultural components of the basin, along with the water-dedicated institutions, water users, operating rules, power relations, water laws, norms and regulations [Bouleau and Pont, 2015, Bouleau *et al.*, 2009, 2017, Carré *et al.*, 2017, 2018]. Each of these subsystems operates on specific time and space scales, and are all interacting between them, as a coevolutionary coupled human-nature system [Savenije *et al.*, 2014, Sivapalan *et al.*, 2012, Werther *et al.*, 2021, Wesselink *et al.*, 2017].

#### 1.4. *The Seine River basin is well suited for such a socio-hydrosystem analysis*

Too few river systems have shown rich information on their evolution over the *longue durée*, particularly for river management, water chemistry and water quality. The Seine, a medium-size basin, is one of those, together with emblematic rivers, first studied for river water resources management as the Tennessee River [Knop, 1976], then for river ecology and restoration, e.g. the Rhône River [Amoros and Petts, 1993], the Rhine River [Wantzen *et al.*, 2022], the Mersey River [Burton *et al.*, 2003] or the Chesapeake basin [Cooper and Brush, 1991]. The PIREN-Seine has allowed a lot of data to be unearthed and analyzed, providing an exceptional set of quantitative data and qualitative descriptions, started 150 years ago by Belgrand [1872] and the Montsouris observatory (1876–1937) [Meybeck *et al.*, 2017], completed by the Seine River authority engineers [MDBSN and AFBSN, 1976], then by the PIREN-Seine: historical population census, accurate and high-resolution archival maps, pioneer chemical analysis, historical records of fish assemblage, river engineers reports, agricultural census, hydrological monitoring, followed then by biogeochemical surveys in the 1970's, etc. [Belliard *et al.*, 2018, Carré *et al.*, 2021, Flipo *et al.*, 2021b, Romero *et al.*, 2016, and references herein].

The program took up a major challenge: to understand major pressures and their interactions with the components of the hydrosystem of this river, and endeavoured to share between researchers and operational staff a common knowledge to act on human practices (domestic, industrial, agricultural ones) in order to improve the quality of the water and the environment. We are presenting in this paper a selection of the Seine River basin socio-hydrosystem

analysis and propose a perspective of its evolution in the 21st century: (i) a nested spatial analysis of the Seine hydrosystem, (ii) the PIREN-Seine modelling approach for water and material transfers, (iii) the trajectories of some major water quality issues of the Lower Seine over the *longue durée*.

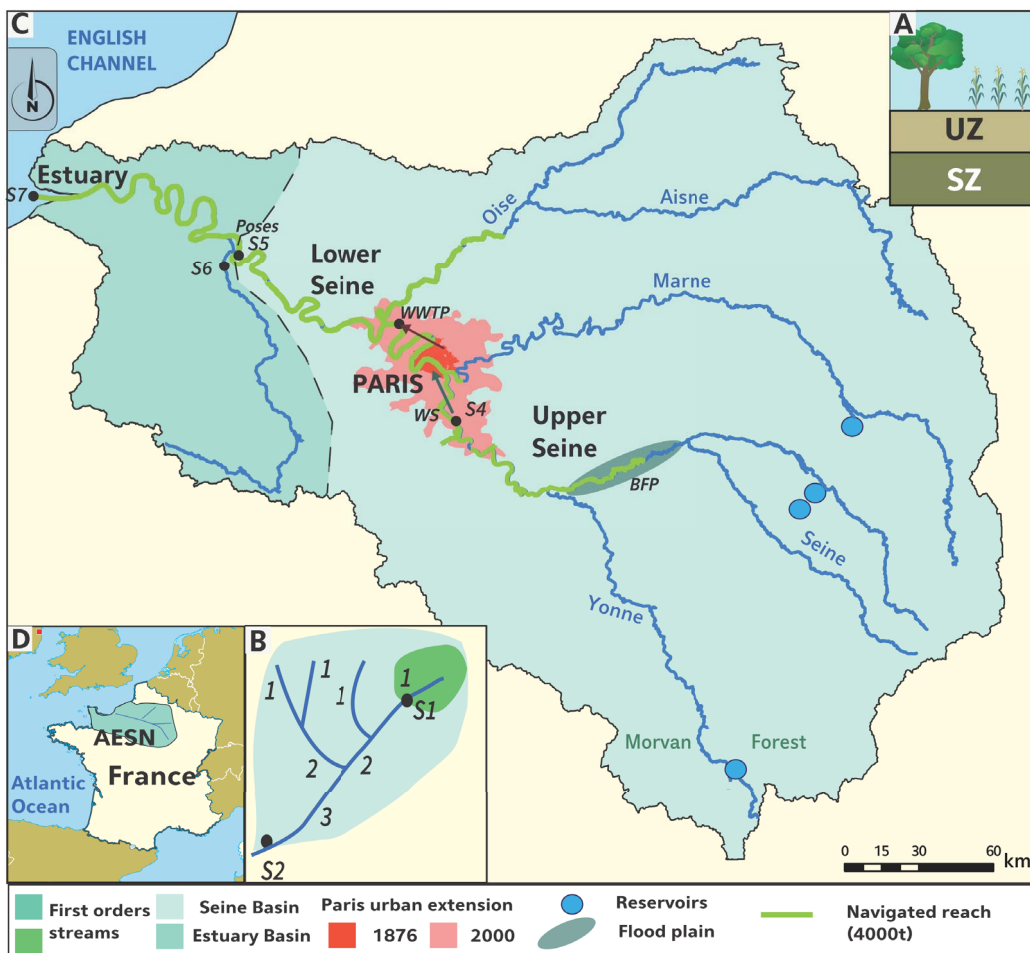
## 2. The Seine hydrosystem from headwaters to the coast, under multiple constraints and controls

One of the main characteristics of the Seine River is its degree of “artificialization” and human control over it. Paris megacity, now administratively termed *Metropole du Grand Paris*, is central to the Seine Basin and dominates the whole socio-hydrosystem, estuary included [Lestel *et al.*, 2022]. Paris' administrative area has increased in the last two hundred years, from 75 km<sup>2</sup> in 1800 to 17,175 km<sup>2</sup> in 2010, of which 2850 km<sup>2</sup> is urbanized (Figure 1C).

Hundreds of sand-pits have been excavated after 1950 to supply this Parisian construction and its radial infrastructure, as far as in the Bassée, once a natural floodplain of 320 km<sup>2</sup> some 100–150 km upstream of Paris (BFP, Figure 1). Paris growth has also generated a well-developed agricultural hinterland. The influence of Paris is maximum on the Lower Seine, downstream of Paris (fluvial transport, locks and dredged channel), but it also extends to upstream parts of the hydrosystem, on the Upper Seine and its Yonne and Marne tributaries, through the control of water levels by dam reservoirs (Figure 1). These structures have generated permanent alterations of the original river course, islands, riparian forests and floodplains [Lestel *et al.*, 2021, 2022].

Few essential key players are in charge of managing river hydrology, water quality and aquatic biodiversity. Some are historical institutions; others were created after the 1964 Water Law. In the following list, PIREN-Seine historical partners are indicated with an asterisk (\*).

- (1) The Seine-Normandie basin institution (AESN\*, established by the 1964 water law) for its pollution control missions raises water taxes to fund the wastewater treatment plants (WWTPs) and is in charge of the river quality monitoring and assessments since 1971 [MDBSN-AFBSN, 1976]. AESN has been



**Figure 1.** The nested spatial vision of the Seine socio-hydrosystem. (A) Plot scale hydrological and biochemical processes. UZ: unsaturated zone; SZ: saturated zone. (B) Low-order streams (1 to 3) draining forest and cropland. (C) River basin, major tributaries and their regulation reservoirs, specific estuarine basin and extension of Paris metropolis from 1876 to 2000. (D) Seine-Normandie water management authority (AESN), France, European Union and Atlantic coast. S1 to S7: schematic position of river observing stations. WS: river water supply for Paris; WWTP: treated urban sewage released to the Seine River. BFP: Bassée floodplain. In green: navigated river network regulated by locks.

and remains the main funder of the program since 1989.

- (2) The Seine reservoirs institution (Seine Grands Lacs\*) was created in 1969 to regulate the discharge of the river within Paris region, through a set of reservoirs totaling 99 km<sup>2</sup> and 800 Mm<sup>3</sup>, built between the 1949 and 1990, some 300 km upstream of Paris (Figure 1C), thus reducing the flood stage in the Paris region and sustaining summer

low flows by the release of additional water up to 75 m<sup>3</sup>·s<sup>-1</sup>, to maintain a sufficient quality for water intakes and drinking water production.

- (3) The urban sanitation syndicate [SIAAP\*, created in 1971, Rocher and Azimi, 2017], a public entity, gathers the majority of the waste waters of Paris megacity through a gigantic network of sewers (440 km). The initial sewer network, conceived in 1860 within historical

Paris, combined both domestic/industrial wastes and storm runoff, but two separate suburban networks were constructed then, mostly after 1950. The total discharge of treated water released in the Central Parisian Seine sector is  $30 \text{ m}^3 \cdot \text{s}^{-1}$ , i.e., equivalent to the lowest natural summer flow, showing the indirect control of this institution on the chemical water quality of Central Seine and Lower Seine.

- (4) The institutions of fluvial navigation and harbors (VNF\* and HAROPA) control the water levels and velocity of the Lower Seine and parts of the Upper Seine, Marne and Oise tributaries (Figure 1C). Since the end of the 19th century, almost the entire course of the Seine River has been channelized and 400 km long network, totaling 20 million tons of freight per year, operating through 29 weirs and locks in which the natural low flows depth has been gradually increased by dredging (1.6 then 2 m in the 1860s, 3.2 m in the 1880s, 4.5 m in the 1930s for the Seine River between Rouen and Paris). The Seine navigated waterways are connected through canals with the Loire River (in 1632), the Rhône River (1832) and the Scheldt River in Belgium (1939), which has favored the invasion of aquatic species [Belliard *et al.*, 2018].
- (5) Two public water supply organizations, the SEDIF\* (142 suburban municipalities totaling 4 M people) and Eau de Paris (Paris and some other municipalities, totaling 3 M people), divert some  $19 \text{ m}^3 \cdot \text{s}^{-1}$  from the Upper Seine, Marne and Oise rivers and from aquifers; Eau de Paris also controls a 1700 km network of pipes for cleaning streets, sewers and watering parks.

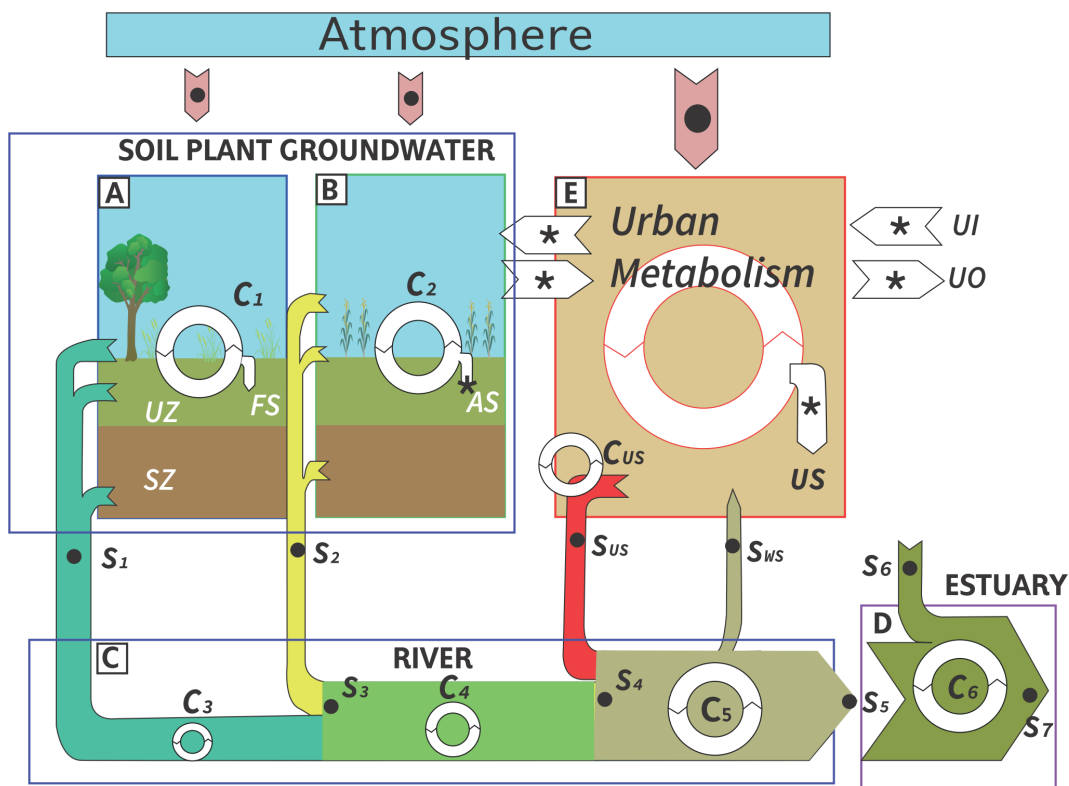
Today, the Seine River is set to remain a highly modified river, regulated to provide the economical functionalities expected. With the help of their exchanges with their partners, the researchers have gradually acquired a better understanding of the visible and invisible forms of degradation of water quality and aquatic environments due to anthropization. This has been made possible by the improvement of measurements, long term monitoring and the development of models.

### 3. The river basin metabolism, its representation and modelling

Modelling has been central to the construction of a shared understanding of the river basin system, between managers and researchers, and among researchers themselves. In the last three decades, numerical dynamic models were developed, taking into account various resolutions and major pressures, including point and non-point pollution, which were initially poorly monitored.

#### 3.1. The PIREN-Seine spatial observation strategy

The program has combined multiple and nested spatial field observations (Figure 1A to 1D), which integrate the hydrological network (from Strahler stream orders 1 to 7), land cover, land use and agricultural practices, human pressure in chosen river sectors (Figure 2): (i) the hydrological and biogeochemical processes and cycles (Figure 2,  $C_1$ – $C_6$ ) are studied at the scale of major land use types (croplands, grassland, and forest), integrating at specific representative study sites and their exchanges with the root zone, unsaturated zone and aquifers, (ii) the background levels of nutrients and heavy metals are measured on forested headwaters, stream orders 1 and 2 (Figure 1,  $S_1$ ,  $S_2$ ), with minimal population density ( $<5 \text{ people}/\text{km}^2$ ), (iii) from stream orders 1 to 7, i.e., up to the basin outlet station of Poses ( $250 \text{ people}/\text{km}^2$ ), the transfer of water and biogeochemical elements are determined at stations representative of each stream order, in small watershed and along longitudinal profiles in major river axis, schematically positioned at  $S_1$ – $S_6$ ; (iv) the general water quality degradation by human activities is studied by documenting point and diffuse sources on the stream order set, particularly on the Seine River upstream and downstream of the Paris megacity ( $S_4$  and  $S_5$ ); (v) some peri-urban (Grand-Morin River) and urban rivers (Bièvre, Orge), where population density reaches  $1000 \text{ people}/\text{km}^2$ , are also considered. These observations are then processed to determine the temporal and spatial distribution of hydrological and biogeochemical fluxes in terms of in/-out budgets, material flux metabolisms, and modelling.



**Figure 2.** Schematic conception of the Seine basin for whole basin modeling, river network budgeting by stream orders, simulation of river water quality and ecological functions. The analysis of territorial metabolism of the basin socio-hydrosystem includes all components (A) to (D): (A) low impacted headwaters, (B) intensive agricultural cropland, (C) river network and its transfers, (E) urban territory, mainly Paris megacity, (D) specific estuarine section. UZ, SZ: unsaturated and saturated zones.  $S_1$  to  $S_7$ , position of main budgeting stations in the land-ocean continuum (see also Figure 1).  $S_{US}$  and  $S_{WS}$ , observing stations for urban sewage release and river water supply. UI and UO, urban input and output. (\*: position of territorial metabolism quantification). FS, AS and US: calculated forest, agricultural soils storage and urban storage.  $C_1$  and  $C_2$ , soil-plant hydrological and biogeochemical cycles;  $C_3$  to  $C_6$ , biogeochemical cycles in aquatic system (cycles in river reservoirs are not shown).  $C_{US}$ , biogeochemical cycle in wastewater treatment plant.

### 3.2. The river basin models

The Seine basin modelling has been a continuous and central activity of the PIREN-Seine since its beginning. Models encapsulate a large part of the generated knowledge, help checking the coherence of the results, and simulate multiple functions of the hydrosystem and basic variables of its water quality.

PROSE [Even *et al.*, 1998, 2007a,b] simulates Seine River hydrodynamics in 2D at a fine scale every 15 mn and 500 m. The whole-basin hydro-biogeochemical model RIVERSTRAHLER allows the calculation of

flows and water quality variables at any point of the river network with a spatial resolution of 1 km [Billen *et al.*, 1994]. The hydrogeological model CAWAQS, spatially distributed into elementary units, 11 km<sup>2</sup> on average [Baratelli *et al.*, 2018, Flipo *et al.*, 2007, 2021a], has been gradually developed, coupled with STICS, a water-plant agronomical model [Ledoux *et al.*, 2007]. These models have been linked to simulate the future evolution of the hydrosystem in response to climate change scenarios and/or to socio-economic changes [e.g. consumption habits, organic agriculture, new river hydrostructures, Passy *et al.*,



2018]. The biogeochemical RIVE model (<https://www.federation-fire.cnrs.fr/rive/>) is common to surface water models (PROSE, RIVERSTRAHLER). Diffuse sources, such as groundwater nitrogen contamination, are determined from STICS and/or GRAFS (an accounting model for soil organic C, N and P balance), at the department (NUTS3) or region (NUTS2) scales, depending on available agricultural census data [Billen *et al.*, 2014]. The estuarine river basin has a dedicated model linked to the coastal zone [ECOMARS-3D: Romero *et al.*, 2019, Garnier *et al.*, 2019].

These water-quality models are used by several partners of the PIREN-Seine. The basin authority (AESN) is using them to complement its regulatory monitoring when reporting for the European Water Framework Directive [EU-WFD, 2000], to assess the good ecological state of unmonitored water bodies [Passy *et al.*, 2018, Carré *et al.*, 2017]. The sewage treatment institution (SIAAP) is using PROSE (including RIVE), which can be considered as a decision support system. It has been used to simulate the water quality of the Lower Seine at high frequency [Vilmin *et al.*, 2018], instead of the monthly regulatory monitoring recommended by the EU-WFD, and to test the impacts of the new SIAAP facilities for phosphorus and nitrogen treatments. It is also used to simulate the impacts of combined sewer overflows during storm events, a critical point for determining when and where future temporary sites, considered for bathing and water sports in the Seine River within Paris, will have to be closed.

These models have also been used to simulate the past water quality in the absence of monitoring before the start of regulatory survey in the early 1970s, etc. [e.g., Billen *et al.*, 2007b] or possible scenarios for the future of the socio-hydrosystem, taking into account different climatic, demographic, agricultural, consumerism scenarios and their combinations [Ducharme *et al.*, 2007, Garnier *et al.*, 2019].

Models have also generated indicators of ecological functioning and status of the land-sea aquatic continuum: (i) the Production/Respiration ratio (P/R) [Billen and Garnier, 2007], (ii) the Indicator of Coastal Eutrophication Potential (ICEP), now included as a Sustainable Development Goal indicator [Billen and Garnier, 2007, Garnier *et al.*, 2021], and tested at the global scale [Garnier *et al.*, 2010, Howarth *et al.*, 2011].

### 3.3. *The territorial metabolism of the Seine socio-hydrosystem*

Earth System scientists seek a deeper understanding of the physical, chemical, biological and human interactions that determine the past, current and future states of the Earth. The Earth System has indeed changed or been regulated by human activities, particularly since the beginning of the industrial era and across the land-to-ocean continuum. The Seine River basin, in which natural fluxes are low, offers an excellent opportunity to quantify these changes (Table 1).

The river basin scale is particularly well-suited to study nutrients and carbon metabolism, their circulation in the anthroposphere as food, organic wastes, chemical fertilizers, N- and P-containing products (detergents), and the effect of the agricultural revolution. During 1950s, after World War II, in France and many other countries, the agricultural sector made a technical revolution by the conversion of pasture to cropland and the systematic use of agrochemicals [Le Noé *et al.*, 2018]. This intensification of agriculture, which generated diffuse excess nitrogen fertilizer and pesticides to the Seine basin streams, has been particularly studied by the PIREN-Seine since 1990 Billen *et al.* [2020], Blanchoud *et al.* [2021]. Other metabolism studies have been established for heavy metals [Meybeck *et al.*, 2007, Thévenot *et al.*, 2007], for PAHs, drug residues and emerging pollutants [Gateuille *et al.*, 2021, Labadie *et al.*, 2021, Tamtam *et al.*, 2008], for specific micropollutants as PCBs, PAHs, phthalates, PBDE, pesticides, microplastics, in rural, and peri-urban areas [Blanchoud *et al.*, 2013, 2021, Dris *et al.*, 2015, Gateuille *et al.*, 2014, Teil *et al.*, 2014]. These studies show the extent to which the material is imported, transported, transformed, recycled, and exported within the anthroposphere.

Other metabolism studies have been established at the infra-basin scale, for the Paris megacity [Billen *et al.*, 2009, Esculier *et al.*, 2019, Lestel, 2012], and at the supra-basin scale, for the whole country, and eventually for Europe [Billen *et al.*, 2021] (Figure 1D).

The circulation of materials (e.g., N, P, metals, potassium) within the Seine socio-hydrosystem is much greater than the natural fluxes (generated by leaching/erosion, weathering, N-fixation, primary production in non-agricultural land, etc.)

**Table 1.** Indicators of impact of the Seine basin anthroposphere on the receiving river for particulate metals, nitrate and phosphate

	Cd	Hg	Pb	Zn	N (NO <sub>3</sub> <sup>-</sup> )	P (PO <sub>4</sub> <sup>3-</sup> )
(A) Pre-industrial references	μg·g <sup>-1</sup>	μg·g <sup>-1</sup>	μg·g <sup>-1</sup>	μg·g <sup>-1</sup>	mg·L <sup>-1</sup>	mg·L <sup>-1</sup>
Seine River natural background (1)	0.22	0.03	20	60	0.2	0.004
Seine River regulatory reference (2)	1.2	0.5–1	20–50	100–200	5.3	0.11
(B) French norms before EU WFD	μg·g <sup>-1</sup>	μg·g <sup>-1</sup>	μg·g <sup>-1</sup>	μg·g <sup>-1</sup>	mg·L <sup>-1</sup>	mg·L <sup>-1</sup>
Dredged sediments (1990s) (3)	2	1	100	300		
WWTP sludge reuse (1986) (4)	26	16	750	2500		
Water quality (green-yellow threshold) (5)					0.5	0.03
(C) Maximum contamination level	μg·g <sup>-1</sup>	μg·g <sup>-1</sup>	μg·g <sup>-1</sup>	μg·g <sup>-1</sup>	mg·L <sup>-1</sup>	mg·L <sup>-1</sup>
Seine R. (downstream Paris) (date)	35 (1960)	26 (1973)	558 (1945)	1317 (1967)	6.36 (2006)	1.14 (1976)
(D) Maximum enrichment factor EF						
Seine R. (downstream Paris)	160	1000	29	22	32	285
(E) Per capita excess fluxes	g·cap <sup>-1</sup> ·y <sup>-1</sup>	g·cap <sup>-1</sup> ·y <sup>-1</sup>	g·cap <sup>-1</sup> ·y <sup>-1</sup>	g·cap <sup>-1</sup> ·y <sup>-1</sup>	kg·cap <sup>-1</sup> ·y <sup>-1</sup>	kg·cap <sup>-1</sup> ·y <sup>-1</sup>
Seine R. maximum (1955–60 for metals; 1990 for N and P)	3.4	0.6	50	160	5.9	1.08
Seine R. 2000	0.07	0.06	5	21	4.8	0.09
(F) Leakage ratio (percent)						
Seine R. 1960s	11.6%	ND	0.94%	1.9%	25%	5%
Seine R. 2000s	0.4%	~10%	0.08%	0.25%	25%	-22%

Per capita excess loads (g·cap<sup>-1</sup>·y<sup>-1</sup>), leakage rates (%), enrichment factors (dimensionless,  $EF > 1.0$ ). (1) historical sediments 4000 BP and forested streams (PIREN-Seine); (2) “natural levels”, AESN, 1979; (3) criteria for safe dredged material reuse (VNF-1); (4) criteria for safe WWTP sludge re-use in 1986; (5) green/yellow limit, SEQ-Eau [Oudin and Maupas, 2003]. From Lestel and Carré [2017], Meybeck [2013], Meybeck and Lestel [2017], Meybeck *et al.* [2007]. Basin population in 1955: 10.5 M people, 1990: 14.5 M people.

which are, in the Seine, generated by medium runoff (10 L·s<sup>-1</sup>·km<sup>-2</sup>) and very-low sediment transport (10 t·km<sup>-2</sup>·y<sup>-1</sup>), i.e., a factor from 10 to 100 times for some elements (metals, nutrients). Specifically, the “longue durée” budget of phosphorus in agricultural soils has shown a bell-shape behavior with a long storage, very rarely quantified, due to over fertilization by chemical P fertilizers (1950–1990), followed by de-storage, while the P balance still remains positive [Billen *et al.*, 2020, Garnier *et al.*, 2015]. Also, the circulation of metals and metal-containing products is completely different from the erosion and weathering processes which generate their natural river fluxes [Meybeck *et al.*, 2007, Thévenot *et al.*, 2007].

The determination of natural references is a key point: (i) they are the starting point of the calculation of the enrichment factor  $EF$  ( $EF = C_{\text{observed}}/C_{\text{natural}}$ ),

as defined by geochemists [Meybeck *et al.*, 2007], (ii) they are used in water quality management, to differentiate the “very good” quality, without any impacts (blue color coded in the WFD) from the “good” quality (green coded) in which impacts are considered as negligible. Reference levels have been established by the PIREN-Seine from pre-industrial alluvial sediments or in low order forested streams (Figures 1B and 2A). They can be very different from the regulatory figures. At one stage of the water quality management, the blue level of the national scale for metals (SEQ-Eau) [Oudin and Maupas, 2003] was simply considered as 1/10 of the green–yellow regulatory limit, often much below the observed pre-industrial values. Examples of French norms (SEQ-Eau and others) before the EU WFD are presented in Table 1. Water quality norms may greatly

evolve with time as they result from scientific knowledge, environmental awareness, analytical facilities, policy objectives and lobbying [Bouleau *et al.*, 2017, Meybeck *et al.*, 2018] as for the nitrate pollution [Carré *et al.*, 2018]. After 2006 the French norms were replaced by those of the EU-WFD. Moreover, the regulatory criteria to qualify the dredged sediments and the recycling of treated sewage sludges on cropland have greatly evolved in the past 40 years [Meybeck *et al.*, 2007].

The leakage rate LR is a dimensionless indicator which can be used for intercomparison between riverborne materials and/or between rivers [Meybeck, 2013]. First, excess loads ( $t$  per year) generated by the socio-hydrosystem are determined as the difference between observed fluxes and the pre-industrial reference flux at the system outlet ( $S_5$ , Figure 1). Then, excess loads are compared to the inputs to the basin (Figure 2,  $IN_{agr}$  and  $IN_{urb}$ ) to determine the leakage rate (LR) of the material. LR reflects the efficiency of the socio-hydrosystem with regard to the fluvial environment: for heavy metals, it was between 1 and 10% after World War II, and has now decreased by an order of magnitude, except for mercury, which is still around 10% [Meybeck *et al.*, 2007].

The per capita excess load is the excess load relative to the population of the socio-hydrosystem at a given period. It is very low or cannot be measured in river basins where the population lives in full equilibrium with its resources. The early river budgets of major ions [Meybeck, 1979] have found that per capita excess fluxes of  $Cl^-$ ,  $SO_4^{2-}$ ,  $Na^+$ ,  $K^+$  were actually proportional to their energy consumption, an indicator of their industrial development. The study of the evolution of this indicator in the Seine basin for metals and for phosphorus adds a new dimension: for these elements, in post-industrial stage and under higher environmental constraints, the per capita excess load can be decreased without negatively affecting economic development.

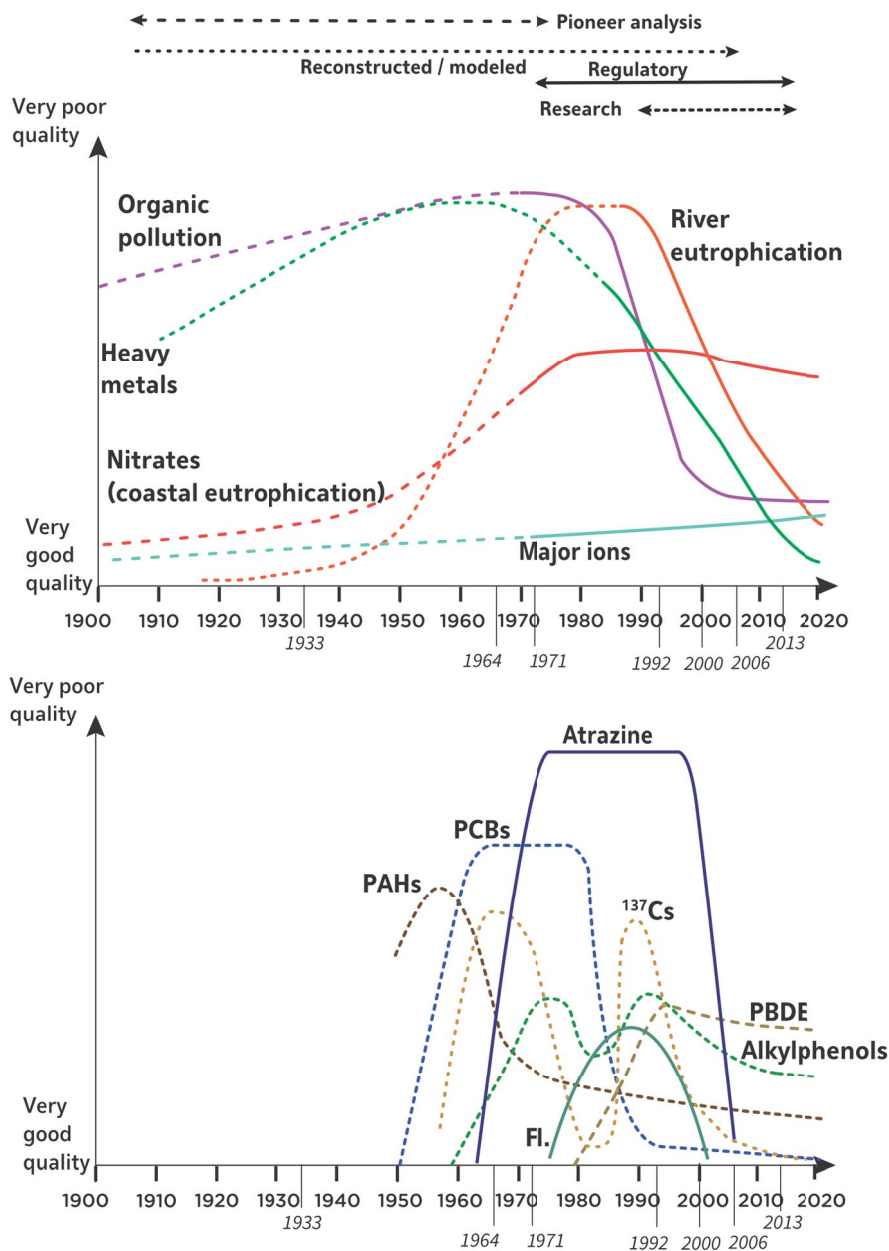
Per capita excess load and leakage rate are specific for the socio-hydrosystem at a given stage of its development and environmental concern. If they can be linked to available socio-economic data (e.g., energy consumption, agricultural production), they can be used to make first-order estimates of human impacts from regional scales to global scale for poorly monitored riverborne materials.

#### 4. Longue durée trajectories of indicators of the state of the Lower Seine quality

Patterns of impacts change over long periods. For most water quality degradations, identifying these impacts, understanding their causes and mitigating the pollution or impact take more than twenty years, often much more and some impacts are yet to happen. In the Seine basin, the longue durée river state has been reconstructed over more than 150 years for many issues [Meybeck *et al.*, 2018]. This is first based on historical archives, maps, fish observations, and water chemistry records generated by the Montsouris observatory, then by other institutions. They allow assessment of the evolution of river course morphology [Lestel *et al.*, 2021], hydrological (river flow) and hydraulic (water level and velocity) regimes, fish assemblage [Belliard *et al.*, 2018], concentrations of major ions, nutrients, dissolved oxygen [Cun *et al.*, 1997, Meybeck and Lestel, 2017], and fecal coliforms [Mouchel *et al.*, 2021]. A second set is generated by biogeochemical models applied to past basin conditions, as for eutrophication [Billen *et al.*, 2007b]. The third set has been generated by the chemical analyses of alluvial cores (sediment archives) after 2006 [Ayrault *et al.*, 2021, Meybeck *et al.*, 2007], which allowed reconstruction of the past contamination trends (from 1920 or 1950 to 2006) for dozens of micropollutants, metallic and organic, which had either not been monitored or insufficiently monitored before 2006.

These state indicator series have then been transformed into quality levels which refer to a single metric, based on a reference water quality (WQ) scale. For the Seine basin, we have chosen those developed by the EU-WFD [Meybeck *et al.*, 2018]. The water quality level (from very good to very poor) allows the intercomparison of issue severity through time in graphical representations (Figure 3). They do not represent how the issue was perceived or declared by the society in the past (actually not documented for most of the time) but correspond to our present analysis of the river environmental trajectory, based on our current scientific knowledge and a unique set of criteria.

Each water quality issue, its development, extension, amplitude and resolution, is specific (Figure 3). In addition, the trajectories may be quite different within each type of issue (e.g. among nu-



**Figure 3.** Longue durée patterns of selected river quality issues in the Lower Seine (at station  $S_5$ ). Relevant dissolved and particulate concentrations observed in the river are converted into issue severity levels according to 2006 scales. Upper graph: classical water quality issues. Lower graph: toxic micropollutants and emerging issues (Fl: flumequine antibiotics). Upper temporal horizontal lines bars: pioneer and historical analyses, reconstruction from sedimentary archives and from models (PIREN-Seine), regulatory analyses performed by AESN, analyses generated by PIREN-Seine. Key dates of water quality construction: 1933, first wastewater treatment plant (WWTP) of Paris sewage, 1964: Water Law, 1971: National Pollution Inventory, 1992: Aquatic systems Law, 2000: EU water framework directive (WFD), 2006: LEMA law (Loi sur l'eau et les milieux aquatiques), 2013: end of implementation of nitrogen treatment at Paris WWTP (Seine-Aval).

trients, heavy metals, pesticides, endocrine disruptors) and specific trajectories must be established (e.g., cadmium, atrazine, flumequine). Also, for a given riverborne material the trajectory analysis depends on viewpoint and on the relevant water quality criteria: the nitrate issue, analyzed with the drinking water criteria ( $50 \text{ mgNO}_3^- \cdot \text{L}^{-1}$ ) used in the basin during several decades, is very different from the one considering the coastal eutrophication guideline ( $10 \text{ mgNO}_3^- \cdot \text{L}^{-1}$ ), recommended by scientists.

The issues that have been solved are characterized by a bell shape which could suggest a virtuous and linear cause-effect relationships with early warning by whistle blowers, early monitoring and assessment, inscription on the political agenda, research of consensus on solutions, and finally, adequate responses. Such ideal succession has only been observed for a few water bodies, such as some alpine lakes, Annecy and Léman (Geneva Lake), [Meybeck, 2002]. In many cases improvements have only been noticeable since the 1970s (organic pollution, DDT, PAHs, PCBs, heavy metals, fecal coliforms), or even later (last 20y for eutrophication) (Figure 3). Some issues have reached a plateau or are slightly improving, such as nitrate pollution, while physical changes of the river course (not represented here) should be considered as permanent [Lestel *et al.*, 2022]. Although sediment records evidence other types of pollution, they are still publicly qualified as emergent, i.e. without regular monitoring nor related WQ criteria. Some WQ issues which have been continuously surveyed since the beginning of monitoring (salinization, acidification) have never shown any problem in the Seine basin (green level on quality scale).

For understanding the evolution of the river quality, a classical analysis is the DPSIR (Drivers of activities, Pressures on the aquatic system, State of the river quality, economic, health or cultural Impacts and policy Responses) recommended by the Europe Environmental Agency and by the United Nation program GEMS-Water. The DPSIR has the hidden hypothesis of a linear, rapid and virtuous process. However, when a detailed and specific analysis is made on the multidecadal, often centennial, time scale which represents the start of an issue, its peak and its resolution, a completely different situation is found, as for the metal pollution, the PCBs contamination, and many others in the Seine basin [Ayrault *et al.*, 2021, Meybeck *et al.*, 2007, 2018]. This example re-

vealed that (i) river quality should be studied over several human generations, (ii) linear cause-effect relation between the DPSIR elements may not exist, (iii) past impacts are generally unknown, (iv) the environmental improvement may occur much before any voluntary response, (v) scientific knowledge may not trigger any action, etc. Whereas DPSIR only captures how problems are understood today, the longue durée water quality trajectory is necessary to understand how social and physical processes have interacted. It incorporates to DPSIR other critical elements as scientific and technical knowledge, all monitoring elements, political agenda and its debates, norms, which allows the intercomparison of issues within a given river basin and between basins, for a given type of issues [Lestel and Carré, 2017].

Detailing the temporality of the multiple elements of a given issue in terms of the first analysis in the hydrosystem, first awareness, start of monitoring, maximum impact, duration of impact, start of response [Meybeck, 2002, Meybeck *et al.*, 2018] as presented in Table 2 for selected issues makes it possible to show the time lags between knowledge and action, knowledge not leading to action and, conversely, action that can be taken without local knowledge of pollution. Actually, water quality is a continuous, spiraling, social construction, informed by past issues and inherited modes of monitoring [Bouleau *et al.*, 2017].

Impacts are generally the least documented components of DPSIR. In the Seine a large data set has been assembled concerning (i) basin drivers, as population, economical statistics, agricultural practices, (ii) basin pressures as sewer inputs, use of fertilizer, chemicals, metals, and (iii) basin responses as environmental laws and regulations, waste treatment, etc. However, impacts of river quality degradation on people's health, economic sectors, on the aquatic environment's health (aquatic ecology and biodiversity), and on related degradation of coastal waters, as well as on economic sectors (tourism, fishing, ...), are still insufficiently assessed in the Seine basin. The national statistics were not conceived to collect such information [Lestel *et al.*, 2007]: hydrological watersheds do not coincide with administrative limits, economical loss remains barely known, past and present environmental impacts on human health in the basin remain poorly known.

The bottlenecks are multiple and refer to different

**Table 2.** Elements of the triplets knowledge, (1)–(3), monitoring, (4)–(6) and responses, (7)–(9) to selected types of pollution of the Lower Seine

(1) Pollution type	(2) Impact start	(3) First awareness	(4) Monitoring start	(5) Objectives of survey	(6) Maximum impact	(7) First regulatory assessment	(8) Regulatory response and quality standards	(9) Technical response
Dissolved oxygen	<1874	1874	1880s	AB	1874–1990	1960s 1976	1971 grid 5 mg·L <sup>-1</sup> SEQ-eau 8 mg·L <sup>-1</sup>	BOD <sub>5</sub> treatment in WWTP
Ammonia	1864	1864	1880s	AB	1970s	1960s 1976	1971 grid 0.5 mg·L <sup>-1</sup> SEQ-eau	NH <sub>4</sub> <sup>+</sup> treatment WWTP 2007
Nitrate DW	1950s	1960s	1874	DW	2000	1971	50 mg·L <sup>-1</sup>	Nitrate treatment WWTP 2007
Nitrate CR		1980s	1874	CR; RT	(2000)		10 mg·L <sup>-1</sup>	Expected
Chloride & Salts	<1874	None	1874	DW; I	1950–2000	1971	WHO-DW 200 mgCl <sup>-</sup> ·L <sup>-1</sup>	Not needed
Phosphate (river eutroph- ication)	1960	1976	1985 (chloro a)	DW; AB; RT	1970–1990		1992 P use restriction	P treatment in WWTP (2000)
PCBs	1952	1970	1990s	FC	1960–1975	1990s	Use ban	PCBs de- struction
Cadmium	<1930	1972	1985 (river bed sediment)	DW	1960	None	Use restriction then ban (1985)	Recycling industrial wastes
Atrazine	(1970)	1980	1985	DW	1970–2000		Ban (2003)	

Objectives of survey: AB aquatic biota degradation; FC food consumption; DW drinking water/human health; RT recreation tourism; I irrigation; CR coastal resources. WHO World Health Organization.

registers of action (see Table 3). Certain levels of river pollution, which dominant actors and experts consider inevitable, is reinforced by ignorance of the effects of industrial products, but also by lobbies, prohibiting any political debate on certain types of pollution. This prevents or limits the implementation of adequate financial or regulatory means. This also explains why the knowledge acquired by monitoring networks and researchers' modeling is not mobilized for action.

## 5. Conclusions and perspectives

PIREN-Seine has provided precise knowledge of the pressures and their dynamics throughout the basin. The models have made it possible to consider point and diffuse inputs, discharges, from the headwaters of rivers to estuaries and coastal waters. The Seine River basin has already been used to test new approaches on a well-documented territory, as the Anthropocene Hypothesis [Meybeck and Les- tel, 2017, Rockström *et al.*, 2014, Vörösmarty *et al.*, 2015], urban socio-ecological trajectories [Barles, 2015], land to sea continuum of nitrogen cycle [Billen

**Table 3.** Typology of the critical points encountered in the construction of the water quality in the Seine River basin, limiting (in no order of priority) the resolution of the issue (with some examples and dates)

Critical points	Examples
State experts accepting pollution as a necessary evil for modernity	Industrial and municipal discharge into rivers until the 1970s
Lack of political concern	All toxics before 1971
Lack of scientific knowledge	Heavy metals before 1970s
Un-appropriate monitoring	Most pesticides before 2006; metals on coarse bed sediments (1985)
Lobbying	Debate on nitrate origins (1980s); debate on river eutrophication cause (1985–95)
Unnecessary frequent monitoring	Major ions; conductivity (since 1971)
Laboratory issues	Contaminated analyses (dissolved metals); high detection limits for many trace components; high analytical cost. (until 2000s)
Wrong WQ target	Choice of drinking criteria for nitrate management (since 1971)
Inconclusive or limited WQ assessment	Metal issue (since 1971); pesticide issue (1985–2006)
Inadequate technical solution	Sewage farms (1870s–1990s)
Lack of financial support	No pollution tax resources before 1968 and lack of political support for financing de-pollution with existing taxes
Lack of administrative tool	No dedicated institution for WQ management and for wastewater management before 1968
Lack of water quality scale	Many examples; emerging pollutants etc.
Underestimated issues	Most toxic contaminants before 2006

and Garnier, 2021].

This knowledge has been used by the public players, depending on the opportunities available, or ignored when the political situation did not allow it. However, the exploration of agricultural and climatic scenarios is now questioning production and consumption patterns and their necessary transformation of the agro-food system, beyond technical solutions alone, with the now recognized effects of climate change.

The PIREN-Seine hydrological models, taking into account the global climate change (GCM) scenarios, predict less summer rainfall, although with localized very intense rainfalls, which may greatly change the hydrosystem functioning, and the availability of water resources, and generate new water-related hazards, as the exceptional flood which occurred in June 2016 [Flipo *et al.*, 2018]. Scenarios of hydrological changes under climate changes can be com-

bined with the ones already explored for deep structural changes in the agrofood system. This knowledge should now be articulated with better understanding of how economic sectors and local territories might be impacted and could improve and adapt. Some processes can be studied independently from actors' strategies; others directly depend on resources, paradigms and routines of social organizations. Therefore credible and useful scenarios must be elaborated with stakeholders willing to address these issues, in order to set robust decisions based on what are the most limiting factors for population resilience.

Interactions between scientists and water managers (dealing with water, but also with agriculture, urban planning, transportation, and health), must be given priority because it is actually on this scale that (i) innovations are made, under social and political pressure, and that (ii) the action value of scientific in-

formation is collectively built.

While the 6th planetary limit concerning the freshwater cycle was crossed in April 2022, the social demand for waterfront reclamation and access to urban nature is increasing and social groups experiencing new practices may foster new concerns for the environment. Opportunities for new perceptions of river systems may arise. The future challenge for research is to find which actors can remove the identified bottlenecks in the management of water quality problems. For this, more detailed studies on the diversity of actors in each sector and their levers of action are crucial. Many events in the last few years have profoundly shaken the ways in which actors project themselves into the future and questioned certainties. As the beginning of the history of PIREN-Seine program shows, only when lock-ins are suddenly released, may the different sources of knowledge accumulated in the program be used for innovative decisions.

### Conflicts of interest

Authors have no conflict of interest to declare.

### Acknowledgments

The authors wish to acknowledge the numerous Seine basin partners for having continuously supported the PIREN Seine program since 1989. This long-term program belongs to the Zone Atelier Seine, an environment observatory instrument created by the CNRS in 2000, part of the international Long Term Socio-Ecological Research (LTSER) network.

### References

- Amoros, C. and Petts, G. E. (1993). *Hydrosystèmes fluviaux*, volume 300. Masson, Paris.
- Amoros, C., Roux, A. L., Reygrobellet, J. L., Bravard, J. P., and Pautou, G. (1987). A method for applied ecological studies of fluvial hydrosystems. *Regul. Rivers Res. Manage.*, 1(1), 17–36.
- Ayrault, S., Meybeck, M., Mouchel, J. M., Gasperi, J., Lestel, L., Lorgeoux, C., and Boust, D. (2021). Simentary archives reveal the concealed history of micropollutant contamination in the Seine River basin. In Flipo, N., Labadie, P., and Lestel, L., editors, *The Seine River Basin*, Handb Environ Chem, 90, pages 269–300. Springer, Cham.
- Baratelli, F., Flipo, N., Rivière, A., and Biancamaria, S. (2018). Retrieving river baseflow from SWOT spaceborne mission. *Remote Sens. Environ.*, 218, 44–54.
- Barles, S. (2010). Society, energy and materials: the contribution of urban metabolism studies to sustainable urban development issues. *J. Environ. Plann. Manage.*, 53(4), 439–455.
- Barles, S. (2015). The main characteristics of urban socio-ecological trajectories: Paris (France) from the 18th to the 20th century. *Ecol. Econ.*, 118, 177–185.
- Belgrand, E. (1872). *La Seine: études hydrologiques*. Dunod, Paris.
- Belliard, J., Beslagic, S., Delaigue, O., and Tales, E. (2018). Reconstructing long-term trajectories of fish assemblages using historical data: The Seine River basin (France) during the last two centuries. *Environ. Sci. Pollut. Res.*, 25, 23430–23450.
- Billen, G., Aguilera, E., Einarsson, R., Garnier, J., Gingrich, S., Grizzetti, B., Lassaletta, L., Le Noë, L., and Sanz-Cobena, A. (2021). Reshaping the European agro-food system and closing its nitrogen cycle: the potential of combining dietary change, agroecology, and circularity. *One Earth*, 4, 839–850.
- Billen, G., Barles, S., Chatzimpiros, P., and Garnier, J. (2012). Grain, meat and vegetables to feed Paris: where did and do they come from? Localising Paris food supply areas from the eighteenth to the twenty-first century. *Reg. Environ. Change*, 12, 325–335.
- Billen, G., Barles, S., Garnier, J., Rouillard, J., and Benoit, P. (2009). The food-print of Paris: long-term reconstruction of the nitrogen flows imported into the city from its rural hinterland. *Reg. Environ. Change*, 9(1), 13–24.
- Billen, G. and Garnier, J. (2007). River basin nutrient delivery to the coastal sea: assessing its potential to sustain new production of non siliceous algae. *Mar. Chem.*, 106, 148–160.
- Billen, G. and Garnier, J. (2021). Nitrogen biogeochemistry of water-agro-food systems: the example of the Seine land-to-sea continuum. *Biogeochemistry*, 154(2), 307–321.
- Billen, G., Garnier, J., and Hanset, Ph. (1994). Modelling phytoplankton development in whole drainage networks: the RIVERSTRAHLER model applied to the Seine river system. *Hydrobiologia*, 289, 119–137.



- Billen, G., Garnier, J., Le Noë, J., Viennot, P., Gallois, N., Puech, T., Schott, C., Anglade, J., Mary, B., Beaudoin, N., Léonard, J., Mignolet, C., Théry, S., Thieu, V., Silvestre, M., and Passy, P. (2020). The Seine watershed water-agro-food system: long-term trajectories of C, N, P metabolism. In Flipo, N., Labadie, P., and Lestel, L., editors, *The Seine River Basin*, Handb Environ Chem, 90. Springer, Cham.
- Billen, G., Garnier, J., Mouchel, J. M., and Silvestre, M. (2007a). The Seine system: introduction to a multidisciplinary approach of the functioning of a regional river system. *Sci. Total Environ.*, 375, 1–12.
- Billen, G., Garnier, J., Némery, J., Sebilo, M., Sferatore, A., Barles, S., Benoit, P., and Benoit, M. (2007b). A long term view of nutrient transfers through the Seine river continuum: mechanisms and long term trends. *Sci. Total Environ.*, 375, 80–97.
- Billen, G., Lassaletta, L., and Garnier, J. (2014). A biogeochemical view of the global agro-food system: Nitrogen flows associated with protein production, consumption and trade. *Global Food Secur.*, 3–4, 209–219.
- Blanchoud, H., Barriuso, E., Nicola, L., Schott, C., Roose-Amsaleg, C., and Tournebize, J. (2013). La contamination de l'Orgeval par les pesticides. In Loumagne, C. and Tallec, G., editors, *L'observation long terme en environnement, exemple du bassin versant de l'Orgeval*, pages 159–174. Eds Quae, Versailles.
- Blanchoud, H., Schott, C., Tallec, G., Queyrel, W., Gallois, N., Habets, F., Viennot, P., Ansart, P., Desportes, A., Tournebize, J., and Puech, T. (2021). How should agricultural practices be integrated to understand and simulate long-term pesticide contamination in the Seine river basin? In Flipo, N., Labadie, P., and Lestel, L., editors, *The Seine River Basin*, Handb Environ Chem, 90, pages 141–162. Springer, Cham.
- Bouleau, G. (2014). The co-production of science and waterscapes: the case of the Seine and the Rhône Rivers, France. *Geoforum*, 57, 248–257.
- Bouleau, G., Argillier, C., Souchon, Y., Barthélémy, C., and Babut, M. (2009). How ecological indicators construction reveals social changes—the case of lakes and rivers in France. *Ecol. Indic.*, 9(6), 1198–1205.
- Bouleau, G., Marchal, P. L., Meybeck, M., and Lestel, L. (2017). La construction politique de la commune mesure de la qualité des eaux superficielles en France: de l'équivalent-habitant au bon état (1959–2013). *Développement durable et territoires. Économie, Géographie, Politique, Droit, Sociologie*, 8(1), 1–19.
- Bouleau, G. and Pont, D. (2015). Did you say reference conditions? Ecological and socio-economic perspectives on the European water framework directive. *Environ. Sci. Policy*, 47, 32–41.
- Bravard, J. P., Amoros, C., and Pautou, G. (1986). Impact of civil engineering works on the successions of communities in a fluvial system: a methodological and predictive approach applied to a section of the Upper Rhône River, France. *Oikos*, 47(1), 92–111.
- Burton, L. R., Howard, A., and Goodall, B. (2003). Construction of a historical water pollution index for the Mersey Basin. *Area*, 35(4), 438–448.
- Carré, C., Boccarossa, A., Hellier, E., and Dupont, N. (2018). Quality standards and monitoring as water management tools: the case of nitrogen in rivers in France and Brittany. *Water Policy*, 20(5), 1059–1074.
- Carré, C., Meybeck, M., and Esculier, F. (2017). The water framework directive's "percentage of surface water bodies at good status": unveiling the hidden side of a "hyperindicator". *Ecol. Indic.*, 78, 371–380.
- Carré, C., Meybeck, M., Garnier, J., Chong, N., Deroubaix, J. E., Flipo, N., Goutte, A., Le Pichon, C., Seguin, L., and Tournebize, J. (2021). River basin visions: Tools and approaches from yesterday to tomorrow. In Flipo, N., Labadie, P., and Lestel, L., editors, *The Seine River Basin*, Handb Environ Chem, 90, pages 381–414. Springer, Cham.
- Cooper, S. R. and Brush, G. S. (1991). Long-term history of Chesapeake Bay anoxia. *Science*, 254(5034), 992–996.
- Cun, C., Bousquet, G., and Vilaginès, R. (1997). A 90-year record of water quality data of Paris Seine and Marne rivers. *J. Water Supply Res. Technol. AQUA*, 46, 150–164.
- Dris, R., Gasperi, J., Rocher, V., Saad, M., Renault, N., and Tassin, B. (2015). Microplastic contamination in an urban area: a case study in Greater Paris. *Environ. Chem.*, 12(5), 592–599.
- Ducharne, A., Baubion, C., Beaudoin, N., Benoit, M., Billen, G., Brisson, N., Garnier, J., Kieken, H., Lebonvallet, S., Ledoux, E., Mary, B., Mignolet, C., Poux, X., Sauboua, E., Schott, C., Théry, S., and Vi-

- ennot, P. (2007). Long term prospective of the Seine River system: Confronting climatic and direct anthropogenic changes. *Sci. Total Environ.*, 375(1–3), 292–311.
- Esculier, F., Le Noë, J., Barles, S., Billen, G., Créno, B., Garnier, J., Lesavre, J., Petit, L., and Tabuchi, J. P. (2019). The biogeochemical imprint of human metabolism in Paris megacity: A regionalized analysis of a water-agro-food system. *J. Hydrol.*, 573, 1028–1045.
- EU-WFD, Water Framework Directive (2000). *Official Journal of the European Communities*, 327(43), 10–72.
- Even, S., Billen, G., Bacq, N., Théry, S., Ruelland, D., Garnier, J., Cugier, P., Poulin, M., Blanc, S., Lamy, F., and Paffoni, C. (2007a). New tools for modelling water quality of hydrosystems: An application in the Seine River basin in the frame of the water framework directive. *Sci. Total Environ.*, 375(1–3), 274–291.
- Even, S., Mouchel, J. M., Servais, P., Flipo, N., Poulin, M., Blanc, S., Chabanel, M., and Paffoni, C. (2007b). Modelling the impacts of combined sewer overflows on the river Seine water quality. *Sci. Total Environ.*, 375(1–3), 140–151.
- Even, S., Poulin, M., Garnier, J., Billen, G., Servais, P., Chesterikoff, A., and Coste, M. (1998). River ecosystem modelling: Application of the PROSE model to the Seine river. *Hydrobiologia*, 373–374, 27–45.
- Flipo, N., Even, S., Poulin, M., Théry, S., and Ledoux, E. (2007). Modeling nitrate fluxes at the catchment scale using the integrated tool CAWAQS. *Sci. Total Environ.*, 375(1–3), 69–79.
- Flipo, N., Gallois, N., Labarthe, B., Baratelli, F., Viennot, P., Schuite, J., Rivière, A., Bonnet, R., and Boé, J. (2021a). Pluri-annual water budget on the Seine basin: past, current and future trends. In Flipo, N., Labadie, P., and Lestel, L., editors, *The Seine River Basin*, Handb Environ Chem, 90, pages 59–89. Springer, Cham.
- Flipo, N., Labadie, P., and Lestel, L., editors (2021b). *The Seine River Basin*. Handb Environ Chem, 90. Springer, Cham. 420 pp.
- Flipo, N., Mouchel, J. M., Fisson, C., Wang, S., and Le Gall, M. (2018). Les effets de la crue de Juin 2016 sur la qualité de l'eau du bassin de la Seine. Fascicule #17. PIREN-Seine. ARCEAU-IdF, 70 pages.
- Garnier, J., Beusen, A., Thieu, V., Billen, G., and Bouwman, L. (2010). N:P:Si nutrient export ratios and ecological consequences in coastal seas evaluated by the ICEP approach. *Glob. Biogeochem. Cycles*, 24(4), article no. GB0A05.
- Garnier, J., Billen, G., Lassaletta, L., Vigiak, O., Nikolaidis, N. P., and Grizzetti, B. (2021). Hydromorphology of coastal zone and structure of watershed agro-food system are main determinants of coastal eutrophication. *Environ. Res. Lett.*, 16, article no. 023005.
- Garnier, J., Billen, G., Legendre, R., Riou, Ph., Cugier, Ph., Schapira, M., Théry, S., Thieu, V., and Menesguen, A. (2019). Managing the agri-food system of watersheds to combat coastal eutrophication: A land-to-sea modelling approach to the French Coastal English channel. *Geosciences*, 9, article no. 441.
- Garnier, J., Lassaletta, L., Billen, G., Romero, E., Grizzetti, B., Némery, J., Le, Q. L. P., Pistocchi, C., Aissa-Grouz, N., Luu, M. T. N., Vilmin, L., and Dorioz, J.-M. (2015). Phosphorus budget in the water-agro-food system at nested scales in two contrasted regions of the world, ASEAN-8 and EU-27. *Global Biogeochem. Cycles*, 29(9), 1348–1368.
- Garnier, J., Meybeck, M., Ayrault, S., Billen, G., Blanchoud, H., Carré, C., Flipo, N., Gasperi, J., Lestel, L., de Marsily, G., Mouchel, J. M., Servais, P., and Tales, E. (2022). Continental Atlantic rivers: the Seine basin. In Tockner, K., Zarfl, C., and Robinson, C. T., editors, *Rivers of Europe*, pages 291–330. Elsevier, Amsterdam, 2nd edition.
- Garnier, J. and Mouchel, J. M. (1999). Man and river system. The functioning of river systems at the basin scale. *Hydrobiologia*, 410, ix–xii.
- Gateuille, D., Evrard, O., Lefevre, I., Moreau-Guigon, E., Alliot, F., Chevreuil, M., and Mouchel, J. M. (2014). Mass balance and decontamination times of Polycyclic Aromatic Hydrocarbons in rural nested catchments of an early industrialized region (Seine River basin, France). *Sci. Total Environ.*, 470, 608–617.
- Gateuille, D., Gasperi, J., Briand, C., Guigon, E., Alliot, F., Blanchard, M., Teil, M. J., Chevreuil, M., Rocher, V., Azimi, S., Thévenot, D., Moilleron, R., Brignon, J. M., Meybeck, M., and Mouchel, J. M. (2021). Mass balance of PAHs at the scale of the Seine River basin. In Flipo, N., Labadie, P., and Lestel, L., editors, *The Seine River Basin*, Handb Environ Chem, 90, pages 163–187. Springer, Cham.

- Howarth, R., Chan, F., Conley, D. J., Garnier, J., Doney, S. C., Marino, R., and Billen, G. (2011). Coupled biogeochemical cycles: eutrophication and hypoxia in temperate estuaries and coastal marine ecosystems. *Front. Ecol. Environ.*, 9(1), 18–26.
- Kabat, P., Claussen, M., Dirmeyer, P. A., Gash, J. H., de Guenni, L. B., Meybeck, M., Pielke, R. A., Vörösmarty, C. J., Hutjes, R. W. A., and Lütkeemeier, S., editors (2004). *Vegetation, Water, Humans and the Climate: A New Perspective on an Interactive System*. Springer, Berlin, Heidelberg. 566 pp.
- Knop, H., editor (1976). *The Tennessee Valley Authority Experience*, volume 1 of *Proceedings of the First Conference on Case Studies of Large Scale Planning Projects*. IIASA, Laxenburg. 752 pp.
- Labadie, P., Alligant, S., Berthe, T., Budzinski, H., Bigot-Clivot, A., Collard, E., Dris, R., Gasperi, J., Guigon, E., Petit, E., Rocher, V., Tassin, B., Tramoy, R., and Treilles, R. (2021). Contaminants of emerging concern in the Seine River basin: overview of recent research. In Flipo, N., Labadie, P., and Lestel, L., editors, *The Seine River Basin*, *Handb Environ Chem*, 90, pages 355–380. Springer, Cham.
- Le Noé, J., Billen, G., Esculier, E., and Garnier, J. (2018). Long term socio-ecological trajectories of agro-food systems revealed by N and P flows: the case of French regions from 1852 to 2014. *Agric. Ecosyst. Environ.*, 265, 132–143.
- Ledoux, E., Gomez, E., Monget, J. M., Viavattene, C., Viennot, P., Ducharme, A., Benoit, M., Mignolet, C., Schott, C., and Mary, B. (2007). Agriculture and groundwater nitrate contamination in the Seine basin. The STICS–MODCOU modelling chain. *Sci. Total Environ.*, 375(1–3), 33–47.
- Lestel, L. (2012). Non-ferrous metals (Pb, Cu, Zn) needs and city development: the Paris example (1815–2009). *Reg. Environ. Change*, 12(2), 311–323.
- Lestel, L. and Carré, C., editors (2017). *Les rivières urbaines et leur pollution*. Eds Quae, Paris. 282 pp.
- Lestel, L., Eschbach, D., Meybeck, M., and Gob, F. (2021). The evolution of the Seine basin water bodies through historical maps. In Flipo, N., Labadie, P., and Lestel, L., editors, *The Seine River Basin*, *Handb Environ Chem*, 90, pages 29–57. Springer, Cham.
- Lestel, L., Meybeck, M., Carré, C., and Belliard, J. (2022). The Seine, the river dedicated to Paris. In Wantzen, K. M., editor, *River Cultures: Life as a Dance to the Rhythm of the Waters*. Unesco Press, Paris. Ch. 27, in press.
- Lestel, L., Meybeck, M., and Thévenot, D. R. (2007). Metal contamination budget at the river basin scale: an original flux-flow analysis (F2A) for the Seine River. *Hydrol. Earth Syst. Sci.*, 11(6), 1771–1781.
- Luciano, E. (2022). *The Anthropocene Hypothesis: birth and epistemology*. PhD thesis, Ludwig Maximilian University of Munich, Germany.
- MDBSN-AFBSN (1976). Les bassins de la Seine et des cours d'eau normands. Mission Déléguée de bassin Seine-Normandie, Agence Financière de bassin Seine-Normandie. Nanterre.
- Meybeck, M. (1979). Concentrations des eaux fluviales en éléments majeurs et apports en solution aux océans. *Rev. Geol. Dyn. Geogr. Phys.*, 21, 215–246.
- Meybeck, M. (2002). Riverine quality at the Anthropocene: Propositions for global space and time analysis, illustrated by the Seine River. *Aquat. Sci.*, 64(4), 376–393.
- Meybeck, M. (2003). Global analysis of river systems: from Earth system controls to Anthropocene syndromes. *Phil. Trans. R. Soc. Lond. Ser. B*, 358(1440), 1935–1955.
- Meybeck, M. (2013). Heavy metal contamination in rivers across the globe: an indicator of complex interactions between societies and catchments. In Arheimer, B., editor, *Understanding Freshwater Quality Problems in a Changing World*, volume 361 of *Int. Ass. Hydrol. Sci. Publ.*, pages 3–16. Copernicus, Göttingen.
- Meybeck, M., Fustec, E., and de Marsily, G. (1998). *La Seine en son bassin: fonctionnement écologique d'un système fluvial anthropisé*. Elsevier, Masson. 736 pp.
- Meybeck, M. and Lestel, L. (2017). A western European river at the Anthropocene. In Kelly, J. M., Scarpino, P., Berry, H., Syvitski, J., and Meybeck, M., editors, *Rivers of the Anthropocene*, pages 84–100. University of California Press, Berkeley, CA. Ch. 7.
- Meybeck, M., Lestel, L., Bonté, P., Moilleron, R., Colin, J. L., Rousselot, O., Hervé, D., de Pontevès, C., Grosbois, C., and Thévenot, D. R. (2007). Historical perspective of heavy metals contamination (Cd, Cr, Cu, Hg, Pb, Zn) in the Seine River basin (France) following a DPSIR approach (1950–2005). *Sci. Total Environ.*, 375(1–3), 204–231.
- Meybeck, M., Lestel, L., and Briand, C. (2017). La

- Seine sous surveillance: Les analyses des impacts de l'agglomération parisienne par l'Observatoire de Montsouris de 1876 à 1937. In Lestel, L. and Carré, C., editors, *Les rivières urbaines et leur pollution*, pages 32–43. Quae, Versailles.
- Meybeck, M., Lestel, L., Carré, C., Bouleau, G., Garnier, J., and Mouchel, J. M. (2018). Trajectories of river chemical quality issues over the Longue Durée: the Seine River (1900S–2010). *Environ. Sci. Pollut. Res.*, 25(24), 23468–23484.
- Meybeck, M. and Vörösmarty, C. (2005). Fluvial filtering of land-to-ocean fluxes: from natural Holocene variations to Anthropocene. *C. R. Geosci.*, 337(1–2), 107–123.
- Mouchel, J. M., Lucas, F. S., Moulin, L., Wurtzer, S., Euzen, A., Haghe, J. P., Rocer, V., Azimi, S., and Servais, P. (2021). Bathing activities and microbiological river water quality in the Paris area: A long-term perspective. In Flipo, N., Labadie, P., and Lestel, L., editors, *The Seine River Basin*, volume 90 of *Handb Environ Chem*, pages 323–353. Springer, Cham.
- Oudin, L. C. and Maupas, D. (2003). *Système d'évaluation de la qualité de l'eau des cours d'eau (SEQ-Eau)*. Ministère de l'écologie et du développement durable-Agences de l'eau, Paris. 106 pp.
- Pahl-Wostl, C., Lebel, L., Knieper, C., and Nikitina, E. (2012). From applying panaceas to mastering complexity: toward adaptive water governance in river basins. *Environ. Sci. Policy*, 23, 24–34.
- Passy, P., Viennot, P., Gallois, N., Billen, G., Garnier, J., Silvestre, M., Thieu, V., and Flipo, N. (2018). Modélisation des apports diffus d'azote et de phosphore aux masses d'eau de surface du bassin Seine-Normandie. Technical Report 1074634 (1) 2017, AESN FIRE/Armines. Convention d'aide financière. 60 pp.
- Rocher, V. and Azimi, S. (2017). *Evolution de la qualité de la Seine en lien avec les progrès de l'assainissement de 1970 à 2015*. Johanet, Paris. 76 pp.
- Rockström, J., Falkenmark, M., Allan, T., Folke, C., Gordon, L., Jägerskog, A., Kummu, M., Lannerstad, M., Meybeck, M., Molden, D., Postel, S., Savenije, H. H. G., Svedin, U., Turton, A., and Varis, O. (2014). The unfolding water drama in the Anthropocene: towards a resilience-based perspective on water for global sustainability. *Ecohydrol*, 7(5), 1249–1261.
- Romero, E., Garnier, J., Billen, G., Ramarson, A., Riou, Ph., and Le Gendre, R. (2019). The biogeochemical functioning of the Seine estuary and the nearby coastal zone: export, retention and transformations. A modelling approach. *Limnol. Oceanogr.*, 64, 895–912.
- Romero, E., Le Gendre, R., Garnier, J., Billen, G., Fisson, C., Silvestre, M., and Riou, Ph. (2016). Long-term water quality in the lower Seine: lessons learned over 4 decades of monitoring. *Environ. Sci. Policy*, 58, 141–154.
- Savenije, H. H., Hoekstra, A. Y., and van der Zaag, P. (2014). Evolving water science in the Anthropocene. *Hydrol. Earth Syst. Sci.*, 18(1), 319–332.
- Sivapalan, M., Savenije, H. H., and Blöschl, G. (2012). Socio-hydrology: A new science of people and water. *Hydrol. Process.*, 26(8), 1270–1276.
- Tamtam, F., Mercier, F., Le Bot, B., Eurin, J., Dinh, Q. T., Clément, M., and Chevreuil, M. (2008). Occurrence and fate of antibiotics in the Seine River in various hydrological conditions. *Sci. Total Environ.*, 393(1), 84–95.
- Teil, M. J., Tlili, K., Blanchard, M., Labadie, P., Alliot, F., and Chevreuil, M. (2014). Polychlorinated biphenyls, polybrominated diphenyl ethers, and phthalates in roach from the Seine River basin (France): impact of densely urbanized areas. *Arch. Environ. Contam. Toxicol.*, 66(1), 41–57.
- Thévenot, D. R., Moilleron, R., Lestel, L., Gromaire, M. C., Rocher, V., Cambier, P., Bonté, P., Colin, J. L., de Pontevès, C., and Meybeck, M. (2007). Critical budget of metal sources and pathways in the Seine River basin (1994–2003) for Cd, Cr, Cu, Hg, Ni, Pb and Zn. *Sci. Total Environ.*, 375(1–3), 180–203.
- Vilmin, L., Flipo, N., Escoffier, N., and Groleau, A. (2018). Estimation of the water quality of a large urbanized river as defined by the European WFD: what is the optimal sampling frequency? *Environ. Sci. Pollut. Res.*, 25, 23485–23501.
- Vörösmarty, C., Lettenmaier, D., Levêque, C., Meybeck, M., Pahl-Wostl, C., Alcamo, J., Cosgrove, W., Grassl, H., Hoff, H., Kabat, P., Lansigan, F., Lawford, R., and Naiman, R. (2004). Humans transforming the global water system. *Eos Trans. Am. Geophys. Union*, 85(48), 509–514.
- Vörösmarty, C. J. and Meybeck, M. (2004). Responses of continental aquatic systems at the global scale: new paradigms, new methods. In Kabat, P., Claussen, M., Dirmeyer, P. A., Gash, J. H., de Guenni, L. B., Meybeck, M., Pielke, R. A., Vörös-

- marty, C. J., Hutjes, R. W. A., and Lütkeemeier, S., editors, *Vegetation, Water, Humans and the Climate*, pages 375–413. Springer, Berlin, Heidelberg.
- Vörösmarty, C. J., Meybeck, M., and Pastore, C. L. (2015). Impair-then-repair: a brief history & global-scale hypothesis regarding human-water interactions in the anthropocene. *Dædalus*, 144(3), 94–109.
- Wantzen, K. M., Uehlinger, U. F., Van der Velde, G., Leuven, R. S., Schmitt, L., and Beisel, N. (2022). The Rhine River basin. In Tockner, K., Zarfl, C., and Robinson, C. T., editors, *Rivers of Europe*, pages 331–389. Elsevier, Amsterdam, 2nd edition.
- Werther, L., Mehler, N., Schenk, G. J., and Zielhofer, C. (2021). On the way to the fluvial Anthroposphere—current limitations and perspectives of multidisciplinary research. *Water*, 13(16), article no. 2188.
- Wesselink, A., Kooy, M., and Warner, J. (2017). Socio-hydrology and hydrosocial analysis: Toward dialogues across disciplines. *Water*, 4(2), article no. e1196.



# **Modelling Philosophy & Law**







---

Research article

Geo-hydrological Data & Models

# On the (im)possible validation of hydrogeological models

Vazken Andréassian <sup>a</sup>

<sup>a</sup> Université Paris-Saclay, INRAE, HYCAR Research Unit, Antony, France

*E-mail:* [vazken.andreassian@inrae.fr](mailto:vazken.andreassian@inrae.fr)

**Abstract.** This paper revisits the controversy on the validation of hydrogeological models, 30 years after it broke out with the publications by Konikow and Bredehoeft [1992a] and de Marsily et al. [1992]. In that debate, Konikow and Bredehoeft [1992a] argued that the word “valid” was misleading to the public and should not be used with respect to models. de Marsily et al. [1992] answered that while the bases of hydrogeological models (conservation of mass and Darcy’s law) were uncontested and unconditionally valid, specific validation exercises were dearly needed to evaluate the parameters and the geometry of these models (confronting the models with data they had not seen during the calibration phase). By updating and extending the literature review, we reanalyze this debate and the arguments presented and conclude by proposing an extension of de Marsily’s position, which underlines the necessity to look at validation from two distinct viewpoints, i.e. the point of view of the model’s explanatory power (theoretical content) and the point of view of its predictive power. The explanatory and predictive dimensions of model validation are to be considered separately.

**Keywords.** Hydrogeological model, Model validation, Corroboration, Falsifiability, Ghislain de Marsily.

*Manuscript received 24 March 2022, accepted 11 July 2022.*

## 1. Introduction

While validation exercises cannot ensure perfection, they help hydrogeologists achieve their level best by increasing their confidence in the model used: This is how one could summarize the position of de Marsily et al. [1992] in the model validation debate that set them in opposition to Konikow and Bredehoeft [1992a]. The argumentation of de Marsily et al. [1992] was quite straightforward: They insisted that model validation was an essential exercise for hydrogeology, and that it was excessive to call upon the Popperian vision of falsifiability [Popper, 1959] to renounce testing exhaustively hydrogeological models: “Groundwater flow models rely essentially on two concepts: (i) mass balance, (ii) Darcy’s law. The former is a principle, not a theory. No one is going to seriously argue that the mass conservation principle may one day be invalidated. [...] Darcy’s law is not a theory; it is an empirical observation, which is applied in a huge number of cases (although it can be in error in a few very special cases, and even so, the departure from the linear Darcy law will be of little significance in most applications)” [de Marsily et al., 1992, p. 367].

Beyond the principles that they considered pointless to contest, de Marsily et al. [1992] insisted that both the parameters and the geometry (the structure) of a hydrogeological model remain uncertain and this is precisely why the so-called *validation* exercises were needed: to either refine the model progressively, or to confirm the robustness of past parametric and structural choices. They argued that the validation exercises were meaningful and that they necessitated using the model in a predictive mode and confronting it with data it had not seen during the calibration phase. This process “increases the confidence” in the model in question, and even if certainty and perfection remain out of reach, this is already a worthy result.

In their response, Konikow and Bredehoeft [1992b] wrote that “using the word ‘valid’ with respect to models misleads the public” and makes hydrogeologists “look foolish to our scientific colleagues”. However, they agreed that the exercise they called “postaudit” (and which consists in revisiting past predictions after a few years) was useful.

As in all controversies, the vocabulary used is not always well defined. “Valid” comes from the Latin

“validus”, meaning “strong, healthy”. The concept of validity has a precise definition in logic, that of a univocal link between the premises and the conclusion of an argument (i.e., if the premises are true, then the conclusion has to be true). It is a well-established concept in law, where a norm is valid if conditions of form (the procedure is respected) and substance (the superior rules of law are respected) are ensured. It is not precisely defined in hydrogeology, where neither Konikow and Bredehoeft [1992a] nor de Marsily et al. [1992] provided a clear definition (we would not consider Konikow and Bredehoeft’s definition of validation as “a process that can guarantee that a model is a correct representation of the physical world” to be precise, because the term “correct” is as vague as the term “valid” was in the first place).

Thirty years have passed since the publication of the articles by de Marsily et al. [1992] and Konikow and Bredehoeft [1992a], and we posit that it is time to propose a critical appraisal of this debate, in the light of more recent contributions on the model validation issue.

## 2. Further contributions from hydrogeologists

Several hydrogeologists brought a further contribution to the debate: Carrera et al. [1993] started by stressing that to them an accurate characterization of geological media was “absurdly utopic”, adding that due to the numerous unknowns and uncertainties in both physical processes and underground media properties, validation was a “rather elusive concept, probably more controlled by the modeler’s background and views of reality than by actual facts”. They insisted on the fact that the qualitative nature of many observations necessarily results in a somewhat subjective conceptualization by the modeler, resulting in several equally likely alternative models. At that point, it was essential to agree on an objective model selection process, and the authors proposed a selection process involving (i) an analysis of model residuals, (ii) an analysis of model parameters (with the aim of having “reasonable” values), and (iii) the computation of theoretical measures of model validity. Acknowledging the difficulty of linking model parameter values with field measurements (because of the scale issue), they insisted on aiming at parameter stability and added that parsimony was a

good means to obtain robust parameters. In conclusion, they underlined that different people perceive the validation process differently, and suggested that models be seen as simple theories about the behavior of the natural systems, to reduce the “drama and controversy often associated with the concept of validation”.

Gorokhovski and Nute [1996] also contributed to the debate: considering the “Popperian” validation of hydrogeological models impossible, they proposed to focus on improving the evaluation of modelling uncertainties using full models and surrogate models, in what they name a “two-level modelling approach”.

The vision of Doherty [2011] is also worth of mention in the frame of this debate: in an editorial of *Groundwater*, he discussed the relative merits of complex (“picture-perfect”) and simple (“abstract”) models, which should both have a role to play for the sake of extracting as much information as possible from historical data. He added that the abstract models are too-often discarded “just because the model does not ‘look like’ what we imagine reality to look like”, while “a model deserves criticism only when it fails to achieve the only thing that it has a right to claim—quantification of uncertainty and maximum reduction of uncertainty through optimal processing of environmental data”.

### 3. The model validation debate on the other side of the hydrological fence (among the “surface” hydrologists)

We all agree that there is only one unique water cycle, and that the border between *hydrogeology* and *surface hydrology* is only cultural, mostly an inheritance of the too-narrow disciplinary teaching of the 20th century. There are however different traditions in surface hydrology (where models focus on reproducing the precipitation-streamflow relationship without mentioning groundwater levels most of the time), and hydrogeology (where the reproduction of piezometric levels is of primary importance and the surface processes are only considered under a “recharge” perspective). Let us now see how the issue of model validation has been dealt with by the “surface” hydrologists.

The loudest voice on this topic has unarguably been that of Vit Klemeš, former president of the International Association of Hydrological Sciences. He entered the debating arena with a paper published few years before the article by de Marsily et al. [1992]. Klemeš [1986] defended the generalization of what he called *split sample tests*<sup>1</sup> (SSTs), and proposed a progressive four-level calibration–validation testing scheme to assess hydrological models. Klemeš’s SST focuses on model transposability in time and space, with increasing difficulties presented to the model: (i) the elementary SST is based on calibrating and validating the model on two independent periods, (ii) the proxy-basin SST is based on transferring parameters between neighboring catchments, (iii) the differential SST is based on calibrating and validating the model on two independent and contrasting (dry/wet or cold/warm) periods, and (iv) the proxy-basin differential SST is based on transferring parameters between neighboring catchments on contrasting periods. Klemeš’s hope was that a wider adoption of SST practices could lead to reducing “the most glaring abuses of simulation models” and in promoting realistic assessments among modelers by avoiding “exaggerated claims regarding model capabilities.”<sup>2</sup> All this is quite similar to de Marsily’s objective: “increasing confidence”.

A few years after the paper by de Marsily et al. [1992], Refsgaard and Knudsen [1996] published a

<sup>1</sup>Note that Klemeš never claimed to have invented the concept [see, e.g., Larson, 1931, Mosteller and Tukey, 1988]: he wrote that the SST “contains no new and original ideas; it is merely an attempt to present an organized methodology based on standard techniques, a methodology that can be viewed as a generalization of the routine split sample test”. But hydrologists still refer very often to his article, which is by far the most cited of his papers (over 750 citations as of December 2021), and SST during the last decade has seen a resurgence of interest [see, e.g., Coron et al., 2012, Seifert et al., 2012, Teutschbein and Seibert, 2013, Thirel et al., 2015, Dakhlaoui et al., 2019, Nicolle et al., 2021].

<sup>2</sup>Many years after publishing his famous paper, Klemeš (personal communication) wrote to us that he had in fact always been skeptical about the capacity of hydrologists to validate rigorously their model. He wrote that he knew in advance that the tests he had suggested would be “avoided under whatever excuses available because modelers, especially those who want to ‘market’ their products, know only too well that they would not pass it.” He concluded: “I had no illusions in this regard when I wrote my paper, but the logic of modelling led me to develop the ‘testing principle’ to its, let’s say, ‘theoretical limit.’”

paper entitled “Operational validation and intercomparison of different types of hydrological models.” They applied Klemeš’s four-level SST scheme to three models of increasing complexity, as their aim was to study the comparative robustness of different models. In this paper, they provide their own definition of “model validation”, which appears to be a nice synthesis of the opinions of de Marsily et al. [1992] and Konikow and Bredehoeft [1992a]: “Model validation is here defined as the process of demonstrating that a given site-specific model is capable of making accurate predictions for periods outside a calibration period. A model is said to be validated if its accuracy and predictive capability in the validation period have been proven to lie within acceptable limits or errors. It is important to notice that the term model validation refers to a site-specific validation of a model. This must not be confused with a more general validation of a generalized modelling system which, in principle, will never be possible” (p. 2190). The same group of authors developed their vision on the subject in subsequent papers [Refsgaard and Henriksen, 2004, Henriksen et al., 2003].

Over the past three decades, Professor Keith Beven has actively discussed the model validation issue. He, however, advocates and promotes a *rejectionist* approach, where “the question is not really validation but rather on what basis should a model run survive invalidation” [Beven, personal communication]. In a recent synthesis [Beven, 2019a], he defends the idea that “a simulation model should be shown to be fit-for-purpose, corroborated against some kind of observation or judgment, even if there are few rules about precisely what constitutes ‘fit’ and ‘purpose’, such that its use can be justified.” He proposes for model evaluation an approach called “limits of acceptability”, considering that there will be “a gradation of acceptability from the ‘best’ models that can be found, to those that are clearly not acceptable as simulators of the system of interest: in this context, the equifinality concept is intrinsically linked to model calibration and validation. The equifinality thesis suggests that there will be no single model representation of an environmental system, but rather an evolving ensemble of models that are considered acceptable in the sense of being useful in prediction as new information becomes available” [Beven, 2019b].

Let us mention here also our own past contribu-

tion to this debate [Andréassian et al., 2009]: While avoiding the terminological debate on the possible or impossible validation of hydrological models, we did argue that it was important to test models as exhaustively and vigorously as possible, with truly demanding tests that we proposed to call *crash tests*: Since the car industry can learn by destroying on purpose an exemplary of their production, we hydrologists should not be ashamed of taking our models to their limits and even a little beyond. We also underlined that the validation of a given model structure would require that tests be conducted on large sets of catchments, as large and varied as possible [see also on this topic Andréassian et al., 2006, Gupta et al., 2014]. A few years later, Biondi et al. [2012] proposed two “code of practices”, one for the validation of the performances of hydrological models, and another for what they call the “scientific validation” of the model. They insist on discussing model limitations “with the same detail that is dedicated to model strengths”, taking the example (Table 1) of the well-known SWOT analysis [on this issue of valuing the evaluation of model failures, see also our discussion in Andréassian et al., 2010].

#### 4. Other relevant contributions from the fields of science history, ecology, and statistics

Science historian Naomi Oreskes made several relevant contributions to the debate, with some explicit references to the dialogue between de Marsily et al. [1992] and Konikow and Bredehoeft [1992a]. In an initial paper, Oreskes et al. [1994] argued that models can only be evaluated in relative terms (i.e., a model should not be declared “good” but only “better” than an alternative one). They underlined that “the term validation does not necessarily denote an establishment of truth. Rather, it denotes the establishment of legitimacy typically given in terms of contracts, arguments and methods. A valid contract is one that has not been nullified by action or inaction. A valid argument is one that does not contain obvious errors of logic. By analogy, a model that does not contain known or detectable flaws and is internally consistent can be said to be valid.” Oreskes et al. [1994] explicitly referred to the position of de Marsily et al., which they commended as *honest* (but not easily marketable...), considering that it fell under the

**Table 1.** Schematic representation of a SWOT analysis for models [modified from Biondi et al., 2012]

		Factors related to the model's predictive power	
		Strengths	Weaknesses
Factors related to the model's explanatory power	Opportunities	Highlight model strengths and related opportunities	Highlight model weaknesses and how they can be mitigated
	Risks	Highlight how model strengths allow avoiding risks	Highlight which risks are caused by model weaknesses

van Fraassen school of thought, i.e., *constructive empiricism*, where the goal of a scientific theory cannot be truth (unobtainable) but rather what van Fraassen names *empirical adequacy*.

In a second paper, Oreskes [1998] returned to the topic of validation in order to address issues related to models used to evaluate/support public policies: There, the semantic debate becomes overwhelming and Oreskes argued that “rather than talking about strategies for validation, we should be talking about means of evaluation”. A very interesting point in Oreske's [1998] paper is a remark on the surprising reluctance of most scientists toward evaluation tests: “Most scientists are aware of the limitations of their models, yet this private understanding contrasts the public use of affirmative language to describe model results.”

In a third paper, Oreskes and Belitz [2001] first expressed semantic regrets—“the term ‘validation’ is an unfortunate one”—then underlined that the main problem lies with the extrapolation capacity of models: “Models may match available observations, yet still be conceptually flawed. Such models may work in the short run, but later fail. [...] Rather than think of models as something to accept or reject [...] it may be more useful to think of models as tools to be modified in response to knowledge gained through continued observation of the natural systems being represented.”

For the ecological sciences, Caswell [1976] discussed the model validation issue and argued that validation should be looked at differently depending on the purpose of the model: He considered it essential to distinguish between predictive models and theoretical models (i.e., models aimed at providing insight into how the system operates). Caswell deemed that theoretical models should be examined according to the Popperian sequence of “conjectures and refutations”, and proposed reserving the term

“validation” for predictive models *only* (and to use the *Popperian* term of corroboration for theoretical models). He explained that the same model can be judged on both grounds, and can eventually be simultaneously declared predictively validated and theoretically refuted.

Two decades later, Power [1993] suggested a two-step approach to validation that would first check that candidate models are able to reproduce the statistical properties of the observations, in order to eliminate models with poor statistical properties. Only in a second phase would the models predictive properties be evaluated. Rykiel [1996] published an exhaustive review of model testing and validation practices in the field of ecological modeling, and his review shows that ecologists do not agree on the semantics or on the practices: In this way, they do not differ from the hydrogeologists! From an ecological research perspective, Rykiel [1996] considered that “the validation problem reflects ambiguity about how to certify the operational capability of a model versus how to test its theoretical content. The crux of the matter is deciding (1) if the model is acceptable for its intended use, i.e., whether the model mimics the real world well enough for its stated purpose, and, (2) how much confidence to place in inferences about the real system that are based on model results. The former is validation, the latter is scientific hypothesis testing. [...] Models can indeed be validated as acceptable for pragmatic purposes, whereas theoretical validity is always provisional.” In conclusion, the author insisted that “validation is not a procedure for testing scientific theory or for certifying the ‘truth’ of current scientific understanding, nor is it a required activity of every modelling project. Validation means that a model is acceptable for its intended use because it meets specified performance requirements.”

More recently, the statistician Shmueli [2010] pub-

lished a synthesis paper entitled “To explain or to predict?” where he discussed in much detail the distinction between explanatory and predictive models. This distinction seems to be central in the model validation debate; indeed, an explanatory model is to be validated qualitatively (and not necessarily quantitatively), while a predictive model is to be validated quantitatively (and could possibly be a “black-box” model, without any explicit explanatory capacity): “Predictive models are advantageous in terms of negative empiricism: a model either predicts accurately or it does not, and this can be observed. In contrast, explanatory models can never be confirmed and are harder to contradict.” Shmueli [2010] argued that misunderstandings arise from the frequent conflation between explanatory power and predictive power in science: “While explanatory power provides information about the strength of an underlying causal relationship, it does not imply its predictive power.” To conclude, the author suggested considering explanatory and predictive abilities as two dimensions: “explanatory power and predictive accuracy are different qualities and a model will possess some level of each.”

## 5. Discussion

### 5.1. *Validation from a model uncertainty perspective*

Over the past 30 years, uncertainty assessments have progressively become an inseparable part of modeling practice. The estimation of predictive uncertainty is seen as a kind of “quality insurance” [Refsgaard et al., 2005] and is as such considered good practice for any environmental modeling activity [Refsgaard et al., 2007]. In groundwater modeling, the uncertainty topic has obviously been discussed for years [de Marsily, 1978, Delhomme, 1979] but no general agreement has yet been reached on how to adequately quantify it; see, for example Barnett et al. [2012] and Guillaume et al. [2016] for a review. Notwithstanding the present popularity of uncertainty assessment exercises, which are now becoming part of the common modeling evaluation practice, it is important to stress here that they can only be seen as a necessary but not sufficient means for model validation, because they only refer to the predictive dimension of models (cf. the aforementioned discussion of the 2010 Shmueli paper). And

one can find in the history of science models that were “right but for the wrong reason” [e.g., the Ptolemaic planetary model and its famous epicycles, Klemesš, 1986].

### 5.2. *Validation from a sensitivity analysis perspective*

Sensitivity analysis (SA) is as old as model construction, but the last three decades have seen a renewed interest in the use of SA techniques. Keeping a model slim is not enough to make it a good model, but it can definitely contribute to turn the model validation process more efficient. According to Saltelli et al. [2000], SA can help investigate “whether a model resembles the system or processes under study; the factors that most contribute to the output variability and that require additional research to strengthen the knowledge base; the model parameters (or parts of the model itself) that are insignificant, and that can be eliminated from the final model; if there is some region in the space of input factors for which the model variation is maximum; the optimal regions within the space of the factors for use in a subsequent calibration study; if and which (group of) factors interact with each other.”

### 5.3. *Validation from a data availability perspective*

Over the past 30 years, the type and amount of data available for model validation has evolved, and this has had an impact on the “feasibility” of validation exercises. On the positive side, distributed data from satellites are now available, sometimes at high frequency. New measurements have appeared, allowing evaluating models at a regional scale rather than at a point scale: one can mention here NASA’s Gravity Recovery and Climate Experiment (GRACE), which provides since 2002 a quantitative measurement of terrestrial water storage changes, allowing the estimation of groundwater storage changes [Tapley et al., 2004]. Other satellite products offer information on actual evaporation and snow extent, and while the quality of satellite precipitation estimates remains rather modest, it has improved too. Water quality and water temperature sensors are also increasingly available, so that in many regions of the world

the possibilities for quantitative validation of hydrogeological predictions have increased. Of course, there is another side to every coin... and one should also mention that in many areas of the world, the density of ground stations (measuring either streamflow, piezometric level or precipitation) has actually decreased...

#### 5.4. *Validation or evaluation?*

Among the criticisms made to the 1992 model validation debate, one is full of good sense: since there is so much controversy around the word “validation”, let us choose another softer one and give it a precise definition. This is the point of view developed by Oreskes [1998]: “rather than talking about strategies for validation, we should be talking about means of evaluation. That is not to say that language alone will solve our problems or that the problems of model evaluation are primarily linguistic. The uncertainties inherent in large, complex models will not go away simply because we change the way we talk about them. But this is precisely the point: calling a model validated does not make it valid.” This is certainly right, but on the other side we must acknowledge that it is extremely complicated to fight language habits! For example, the French language is full of undesirable anglicisms, which the Académie Française is fighting against... with limited success. We have been able to introduce “ordinateur” to replace “computer”, but we keep using the English language “sport” instead of its old French equivalent “desport”. If we decide to wait for our colleagues to accept and adopt our naming conventions... we may need a lot of patience.

## 6. Conclusion

Thirty years after the publication of the article by de Marsily et al. [1992], our literature review has allowed us to shed some new light on the model validation debate. For de Marsily et al. [1992], model validation exercises were meant to *increase the confidence* that a hydrogeologist would have in his/her model. This notion of *confidence* was multifactorial, as a model was to hold both explanatory power (cf. the reference to Darcy’s law and the principle of mass conservation) and predictive power (cf. the reference to success obtained in tests on an independent period).

This distinction between the predictive and explanatory dimensions of validation [underlined among others by Caswell, 1976, Beven, 2001 and Shmueli, 2010] is essential: With regard to our model validation debate, it implies that model validation can have two dimensions (hence the possibility of misunderstandings for those who did not realize it in the first place). It also implies the possibility of searching for compromises between these two dimensions: A “strong” predictive model could be preferred to a “weak” explanatory model, and vice versa. Obviously, validation becomes a multi-objective endeavor, and as such, it will require hydrogeologists to look for compromises (which may remain a matter of debate among them).

To conclude this conclusion, we would like to propose our own definition of model validation by extending that of de Marsily; and while this new definition is unavoidably shifted towards the way surface hydrologists look at models, we do believe that it retains enough generality to be of common interest to the hydrological and hydrogeological sciences:

- (1) The validation of models is possible and necessary;
- (2) When judging the validity of a model, one needs to keep in mind that a model remains an abstraction and a simplification;
- (3) Judging the validity of a hydrological model requires one to consider the model’s objectives as well as its space and time scale;
- (4) Validity can be considered from the point of view of the model’s explanatory power (theoretical content) and/or from the point of view of its predictive power. The explanatory and predictive dimensions of model validation must be considered separately: A model can eventually be simultaneously declared predictively validated and theoretically refuted;
- (5) When validity cannot be assessed in an absolute way, the value of a model can be examined from a comparative perspective;
- (6) When judging a model’s predictive power, the quantitative predictions are at least to be judged based on measurements that have not been used for model calibration, and possibly on measurements requiring a higher extrapolation capacity;

- (7) An assessment of the model's predictive uncertainty can be helpful with the validation process.

### Conflicts of interest

The author has no conflict of interest to declare.

### Acknowledgements

The author acknowledges the review of two anonymous referees, which helped him improve significantly his manuscript.

### References

- Andréassian, V., Hall, A., Chahinian, N., and Schaake, J. (2006). Introduction and synthesis: Why should hydrologists work on a large number of basin data sets? In *Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment—MOPEX*, volume 307, pages 1–5. IAHS Publ.
- Andréassian, V., Perrin, C., Berthet, L., Le Moine, N., Lerat, J., Loumagne, C., Oudin, L., Mathevet, T., Ramos, M. H., and Valéry, A. (2009). Crash tests for a standardized evaluation of hydrological models. *Hydrol. Earth Syst. Sci.*, 13, 1757–1764.
- Andréassian, V., Perrin, C., Parent, E., and Bardossy, A. (2010). Editorial—the court of miracles of hydrology: can failure stories contribute to hydrological science? *Hydrol. Sci. J.*, 55(6), 849–856.
- Barnett, B., Townley, R., Post, V., Evans, R., Hunt, R. J., Peeters, L., Richardson, S., Werner, A. D., Knapp, A., and Boronkay, A. (2012). Australian groundwater modelling guidelines. Report no 82. National Water Commission, Canberra.
- Beven, K. (2001). On explanatory depth and predictive power. *Hydrol. Process.*, 15, 3069–3072.
- Beven, K. (2019a). Invalidation of models and fitness-for-purpose: A rejectionist approach. In Beisbart, C. and Saam, N. J., editors, *Computer Simulation Validation*, pages 145–171. Springer, Cham.
- Beven, K. (2019b). Validation and equifinality. In Beisbart, C. and Saam, N. J., editors, *Computer Simulation Validation*, pages 791–809. Springer, Cham.
- Biondi, D., Freni, G., Iacobellis, V., Mascaro, G., and Montanari, A. (2012). Validation of hydrological models: Conceptual basis, methodological approaches and a proposal for a code of practice. *Phys. Chem. Earth*, 42–44, 70–76.
- Carrera, J., Mousavi, S. F., Usunoff, E. J., Sánchez-Vila, X., and Galarza, G. (1993). A discussion on validation of hydrogeological models. *Reliab. Eng. Syst. Saf.*, 42, 201–216.
- Caswell, H. (1976). The validation problem. In Patten, B., editor, *Systems Analysis and Simulation in Ecology*, volume IV, pages 313–325. Academic Press, New York, NY.
- Coron, L., Andréassian, V., Perrin, C., Lerat, J., Vaze, J., Bourqui, M., and Hendrickx, F. (2012). Crash testing hydrological models in contrasted climate conditions: An experiment on 216 Australian catchments. *Water Resour. Res.*, 48, article no. W05552.
- Dakhlaoui, H., Ruelland, D., and Trambly, Y. (2019). A bootstrap-based differential split-sample test to assess the transferability of conceptual rainfall-runoff models under past and future climate variability. *J. Hydrol.*, 575, 470–486.
- de Marsily, G. (1978). *De l'identification des systèmes hydro-géologiques*. Doctorat d'état thesis, Université Pierre et Marie Curie, Paris.
- de Marsily, G., Combes, P., and Goblet, P. (1992). Comment on 'Ground-water models cannot be validated', by L.F. Konikow and J.D. Bredehoeft. *Adv. Water Resour.*, 15, 367–369.
- Delhomme, J. P. (1979). Spatial variability and uncertainty in groundwater flow parameters: a geostatistical approach. *Water Resour. Res.*, 15, 269–280.
- Doherty, J. (2011). Modeling: Picture perfect or abstract art? *Groundwater*, 49(4), 455.
- Gorokhovski, V. and Nute, D. (1996). Validation of hydrogeological models is impossible: what's next? In *Calibration and Reliability in Groundwater Modelling*, volume 237, pages 417–424. IAHS Red Book.
- Guillaume, J. H. A., Hunt, R. J., Comunian, A., Blakers, R. S., and Fu, B. (2016). Methods for exploring uncertainty in groundwater management predictions. In Jakeman, A. J., Barreteau, O., Hunt, R. J., Rinaudo, J.-D., and Ross, A., editors, *Integrated Groundwater Management: Concepts, Approaches and Challenges*, pages 711–737. Springer International Publishing, Cham.
- Gupta, H. V., Perrin, C., Kumar, R., Blöschl, G., Clark, M., Montanari, A., and Andréassian, V. (2014).



- Large-sample hydrology: a need to balance depth with breadth. *Hydrol. Earth Syst. Sci.*, 18, 463–477.
- Henriksen, H. J., Trolldborg, L., Nyegaard, P., Sonnenborg, T., Refsgaard, J. C., and Madsen, B. (2003). Methodology for construction, calibration and validation of a national hydrological model for Denmark. *J. Hydrol.*, 280, 52–71.
- Klemeš, V. (1986). Operational testing of hydrological simulation models. *Hydrol. Sci. J.*, 31, 13–24.
- Konikow, L. F. and Bredehoeft, J. D. (1992a). Groundwater models cannot be validated. *Adv. Water Resour.*, 15, 75–83.
- Konikow, L. F. and Bredehoeft, J. D. (1992b). Reply to comment. *Adv. Water Resour.*, 15, 371–372.
- Larson, S. C. (1931). The shrinkage of the coefficient of multiple correlation. *J. Educ. Psychol.*, 22, 45–55.
- Mosteller, F. and Tukey, J. W. (1988). Data analysis, including statistics. In *The Collected Works of John W. Tukey: Graphics 1965-1985*, volume 5. CRC Press, Boca Raton.
- Nicolle, P., Andréassian, V., Royer-Gaspard, P., Perrin, C., Thirel, G., Coron, L., and Santos, L. (2021). Technical note: RAT – a robustness assessment test for calibrated and uncalibrated hydrological models. *Hydrol. Earth Syst. Sci.*, 25, 5013–5027.
- Oreskes, N. (1998). Evaluation (not validation) of quantitative models. *Environ. Health Perspect.*, 106, 1453–1460.
- Oreskes, N. and Belitz, K. (2001). Philosophical issues in model assessment. In Anderson, M. G. and Bates, P. D., editors, *Model Validation: Perspectives in Hydrological Science*, pages 23–41. John Wiley and Sons, Ltd, London.
- Oreskes, N., Shrader-Frechette, K., and Belitz, K. (1994). Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, 263, 641–646.
- Popper, K. (1959). *The Logic of Scientific Discovery*. Routledge, London.
- Power, M. (1993). The predictive validation of ecological and environmental models. *Ecol. Model.*, 68, 33–50.
- Refsgaard, J. C. and Henriksen, H. J. (2004). Modelling guidelines—terminology and guiding principles. *Adv. Water Resour.*, 27, 71–82.
- Refsgaard, J. C., Henriksen, H. J., Harrar, W. G., Scholten, H., and Kassahun, A. (2005). Quality assurance in model based water management—review of existing practice and outline of new approaches. *Environ. Model. Softw.*, 20, 1201–1215.
- Refsgaard, J. C. and Knudsen, J. (1996). Operational validation and intercomparison of different types of hydrological models. *Water Resour. Res.*, 32, 2189–2202.
- Refsgaard, J. C., van der Sluijs, J. P., Hojberg, A. L., and Vanrolleghem, P. A. (2007). Uncertainty in the environmental modelling process—A framework and guidance. *Environ. Model. Softw.*, 22, 1543–1556.
- Rykiel, E. J. (1996). Testing ecological models: the meaning of validation. *Ecol. Model.*, 90, 229–244.
- Saltelli, A., Chan, K., and Scott, E. M. (2000). *Sensitivity Analysis*. John Wiley, Hoboken, NJ.
- Seifert, D., Sonnenborg, T. O., Refsgaard, J. C., Højberg, A. L., and Trolldborg, L. (2012). Assessment of hydrological model predictive ability given multiple conceptual geological models. *Water Resour. Res.*, 48, article no. W06503.
- Shmueli, G. (2010). To explain or to predict? *Stat. Sci.*, 25, 289–310.
- Tapley, B. D., Bettadpur, S., Watkins, M. M., and Reigber, C. (2004). The gravity recovery and climate experiment; mission overview and early results. *Geophys. Res. Lett.*, 31(9), article no. L09607.
- Teutschbein, C. and Seibert, J. (2013). Is bias correction of regional climate model (RCM) simulations possible for nonstationary conditions? *Hydrol. Earth Syst. Sci.*, 17, 5061–5077.
- Thirel, G., Andréassian, V., Perrin, C., Audouy, J.-N., Berthet, L., Edwards, P., Folton, N., Furusho, C., Kuentz, A., Lerat, J., Lindström, G., Martin, E., Mathevet, T., Merz, R., Parajka, J., Ruelland, D., and Vaze, J. (2015). Hydrology under change: an evaluation protocol to investigate how hydrological models deal with changing catchments. *Hydrol. Sci. J.*, 60, 1184–1199.





Research article

Geo-hydrological Data & Models

# High-level radioactive waste repository: How Geology combined with societal principles can lead to public acceptance—the German experiment

Maria-Theresia Schafmeister<sup>©, a, b</sup>

<sup>a</sup> Applied Geology/Hydrogeology Department, Greifswald University, Friedrich-Ludwig-Jahn-Str. 17a, D-17487 Greifswald, Germany

<sup>b</sup> Member of the National Citizens' Oversight Committee (Assisting the Search for a High-Level Radioactive Waste Disposal Site)

E-mail: [schaf@uni-greifswald.de](mailto:schaf@uni-greifswald.de)

**Abstract.** Using nuclear power for civil energy generation was a promising idea in the second half of the 20th century. Compared to fossil fuels, nuclear power results in almost no CO<sub>2</sub> bearing emissions, which are known to have triggered and enhanced the global climate change. However, nuclear power bears other risks, e.g. environmental and health damage caused by radiation, unwanted release of nuclides being transported by water, accidents of nuclear power plants, and not to mention the possibility of non-intended military use. In any case, generating power from nuclear fission will always produce radiating waste, albeit in small amounts. Due to its longevity, this waste is a significant challenge for scientists, politicians and for the international social community to find a repository, which can be considered safe for millions of years and is accepted by the public. As the general consensus is that the rocks of earth crust's are the best option for a controlled storage, scientists including geologists, geophysicists and hydrogeologists are looking worldwide for safe underground repository sites. Given different political preconditions, the nations, especially in the European Union, have developed different strategies. The following article intends to describe the current strategy in Germany, which is based on a specific "Act on the search for and selection of a site for a final repository for highly radioactive waste" (in short: Site Selection Act) and attempts to learn from mistakes in the seventies. It is our strong conviction that fundamental geological aspects are as important as essential societal principles in order to gain maximum public acceptance.

**Keywords.** High-level radioactive waste, Site selection, Geoscience, Social science, Germany.

*Manuscript received 30 March 2022, revised and accepted 25 October 2022.*

## 1. Introduction

"It is not a sin to take excessive safety precautions; it should be the rule. After all, who blamed Che-Huang-ti, Chinese emperor of the 3rd century B.C., for building the Great Wall (probably one of the greatest works of man to this day) even if it was never needed?" [de Marsily, 1982]. It is with these words that Ghislain de Marsily ended his foreword to the

proceeding volume of the 26th Congress of IAMG (International Association for Mathematical Geosciences) held within the International Geological Congress (IGC), Paris in 1980. The small but distinguished IAMG had introduced mathematical and statistical prediction methods in order to estimate mineral resources. Here, their focus was on prediction of potential risks associated with nuclear power and its waste management.

A few lines earlier, he stated: “However, they (the volume’s papers) also make it clear that geologists are a long way from the ultimate goal which is meaningful risk analysis. Until this goal is reached, it is the responsibility of the world geological community (i.e. each of us) to ensure that no decisions are made, which take the stability of the present world for granted, ignoring basic geodynamic processes, which may increase drastically the risk associated with the decision. Whereas individuals are free to accept various risks, society is not, and the risk to which our own species may be exposed in the future should be reduced to a minimum, which is the law of nature in any ecosystem.”

The following paper aims at demonstrating how Germany, i.e. parliament and government, society and scientists, is honestly struggling to live up to this principle of calculating, assessing and minimizing all risks associated with the underground storage of high-level radioactive waste. This is currently necessary after all previous plans to select an appropriate high-level waste (HLW) repository in Germany had failed.

After WWII, many nations have profited from civil use of nuclear power. Also, Germany became enthusiastic about an apparently endless energy source which seemingly causes no environmental damages compared to the consumption of fossil fuels. Interestingly, in the early 60s when West Germany enhanced the mining of vast amounts of hard coal, some energy supply companies were skeptical to bear the economic risk of nuclear energy production. This attitude changed soon, when the so-called Energy-Gap was envisioned. Both Germanies started to rely on nuclear energy production [End-Kom, 2016]. The first atomic power plant in West Germany was put into operation in Bavaria in 1966, the former German Democratic Republic (East Germany) followed with the nuclear power plant Bruno-Leuschner in 1974 in Lubmin-Greifswald. Its latest block operated only a month, in November 1989—the year of the “Wende”—before the plant was finally shut down. The former power plant’s site is now used as one of the sixteen interim storage facilities in Germany. It also provides technical knowledge and services in dismantling and disposal of nuclear energy facilities [EWN, 2022].

Nuclear power plants produce the most intensively radiating waste. Whereas HLW [IAEA, 2009]

only encompasses ten percent of the total volume of radioactive waste, it covers more than 99% of the total radioactivity. By the end of 2022, which is the date when all German nuclear power plants were planned to be shut down, approx. 10,500 tons of high-level radioactive waste, which corresponds to a volume of 27,000 m<sup>3</sup>, are estimated. According to the German understanding, this volume also includes heat-emitting intermediate-level radioactive waste, since the heat generation may have a major impact on the configuration of the subsurface repository and its host rock. In total, Germany will face 600,000 m<sup>3</sup> of radioactive waste, most of which are classified as low- and intermediate-level waste (LILW). The safe disposal of this radiating waste is not yet ensured. This paper focuses on the deposition of HLW; however, it should be noted that LILW has been disposed in subsurface structures in the past, e.g. Asse/Lower Saxony and Morsleben/Saxony-Anhalt, and that during the selection process of a potential repository site for HLW, the concurrent storage of LILW needs to be explored [StandAG, 2017].

After seventy years of operating nuclear power plants in Germany, there is still no final repository for HLW in sight, although many attempts have been made and some underground sites had been suggested. For example, the salt mine Asse in Lower Saxony has been used as repository; however, it is now considered a contaminated site, since groundwater inrush became evident in 2013. The deposited waste now needs to be recovered, which is a high technical challenge for the operating federal company (BGE, Federal Company for Radioactive Waste Disposal, “Bundesgesellschaft für Endlagerung”). Three other sites, all located in the North German Basin have been under debate: a former iron-ore mine Konrad/Lower-Saxony, and two salt-domes Morsleben and Gorleben. It may not be a coincidence of history that both the latter sites are located close to the former German–German border opposite each other.

While the storage of nuclear waste is a worldwide problem, different nations follow different selection strategies. Germany additionally faces the problem of two separate developments regarding the use of nuclear energy, its threats, and the final disposal of nuclear waste, which have emerged over forty years of different political systems. In West Germany, a broad resistance had evolved already in the seven-

ties. The Anti-Atomic-Power movement drew attention on environmental and health hazards by numerous demonstrations, which culminated in the illicit occupation of a 12 ha area at Gorleben in 1980. Here, a deep exploration drill hole was planned. The so-called “Republik Freies Wendland” was cleared only one month after its founding by police and federal border guards [Bundeszentrale für Politische Bildung, 2020]. While anti-nuclear movements in West Germany were causing public uproar, in the GDR the advantages of nuclear power were stressed. Uranium mining in the Erzgebirge/Saxony, the close science collaboration with the former Soviet Union, as well as the construction and operating of nuclear power plants provided employment and moderate economic livelihood, benefits that were hardly publicly questioned [Stude, 2019]. Potential risks and nuclear accidents, e.g. the Super MCA (maximum credible accident) in Chernobyl/Ukraine (former Soviet Union) and its fall-out in 1986 were systematically concealed [Pflugbeil, 2006]. Only after this event anti-atomic movements evolved and became more and more visible also in the GDR.

In retrospect, 40 years of unfortunate political approaches and technical misjudgments in connection with the final disposal of HLW is reason to develop new search strategies in order to achieve the greatest possible social confidence in the future repository. This paper intends to present the recent strategies in Germany to select a safe repository for Germany's HLW. A brief introduction into the German geological situation with respect to the disposal of heat-developing HLW will be given. Thereafter, the geological and societal decision criteria will be highlighted. The actual activities and their various technical, societal and legislative parties will be summarized, now that the first step along the new German selection strategy was finalized in 2020.

## 2. The novel concept of site selection

Some legal boundary conditions and specific events must be mentioned prior to explaining the active search of an HLW repository in Germany. In 2011, after the Fukushima event, Germany decided on an irreversible withdrawal from nuclear power generation. The amendment of the Atomic Act [AtG, 2011] rules that all nuclear power plants shall be shut down by the end of 2022. For reason of actuality, it must be

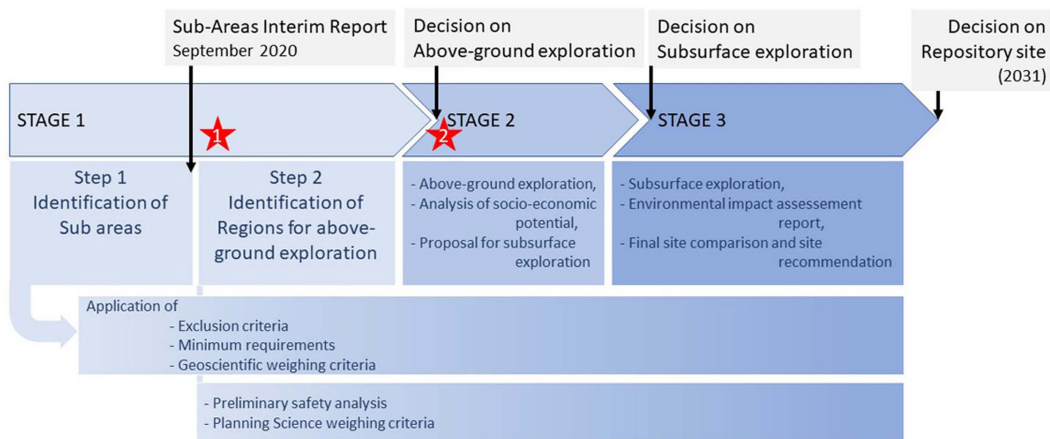
mentioned that the unforeseen Ukraine crisis and its consequences to the European energy economy have led to intensive political debates in Germany as to whether or not the last three operating nuclear power plants should continue to do so. The current governing coalition in Germany has now agreed on an extension until mid April 2023. Existing fuel rods will continue to be used.

A first version of the Site Selection Act was adopted in 2013 and further specified after two years of intensive work in the Commission for the Storage of High-Level Waste [EndKom, 2016] in 2017 [StandAG, 2017]. The crucial paradigm shift was not to explore a site with respect to its appropriateness for a safe storage but to consider that other sites might be relatively more favorable. While all exploration activities so far focused on the Gorleben salt dome, the entire territory of the FRG is now to be considered as a potential repository site until technical arguments lead to the exclusion of regions. Thus, it is expected to select the most appropriate location by comparison [Däuper et al., 2013, EndKom, 2016]. Since nuclear waste management is considered a national task, the selection procedure is organized at the federal level, i.e., the state administrations cannot interfere, thus individual economic or social interests of the states are to be prevented.

Another lesson learnt from the past is that the process of selection needs to meet the following five attributes: citizen participative, science-based, transparent, self-questioning and learning. In order to ensure these requirements, the Site Selection Act [StandAG, 2017] defines various operative and supervising parties and several stages along the process, when decisions need to be debated and finally made. The declared aim is to give the citizens of Germany the greatest possible confidence in the process and the selected site.

### 2.1. Organization of selection process as ruled by the site selection act

The stated objective of the selection process is to identify an underground structure with the best characteristics to ensure safe containment of HLW for one million years. A secondary condition is that 500 years after the disposal of the radioactive waste and the closure of the site, recovery should be possible.

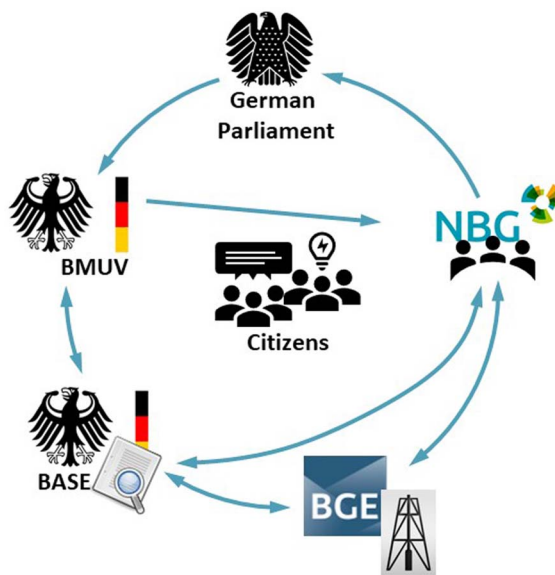


**Figure 1.** Process of Site selection. Stars indicate the times of public participation: 1—Sub-area conference, 2—Regional conferences. Modified after BGE [Reiche et al., 2021].

The selection process works along specific criteria which are subject to increasingly stringent testing, a process that can be easily illustrated with the help of the well-known grain size analysis, in which a series of sieves with decreasing mesh size is ultimately used to separate the smallest grain fraction from all the coarser ones. While in the beginning, the scientific, i.e. geological and radiation physical aspects dominate, societal and planning science criteria gain weight as the closer potential regions are identified (Figure 1).

Various parties and political institutions are involved in the process (Figure 2): the operating federal company (BGE, Federal Company for Radioactive Waste Disposal), which actively performs the site selection, repository construction and nuclear waste storage. These activities are monitored by a federal authority (BASE, Federal Office for the Safety of Nuclear Waste Management), which also has to ensure public participation. Finally, the German Federal Ministry for the Environment, Nature Conservation, Nuclear Safety and Consumer Protection (BMUV) and the parliament, which makes the final decision. The Site Selection Act also stipulates that the citizens of Germany must be involved. This is achieved by establishing technical conferences, where the above listed parties, scientists and citizens can discuss the preliminary results of specific steps in the selection process. They may also propose their own requests [StandAG, 2017].

Finally, the Site Selection Act rules that an inde-



**Figure 2.** Parties involved in the site-selection process. Arrows indicate direction of interaction. Modified after [BGE, 2020a].

pendent board is installed to oversee all phases of the search process, reviewing all documents, and thus monitoring compliance with the five fundamental principles (NBG, National Citizens' Oversight Committee). This panel is pluralistically composed by twelve recognized personalities from science and society, who are appointed by the parliament, and six citizens including young people. The citizen repre-

sentatives are elected out of a pool of randomly addressed and willing persons. All NBG members work on an honorary basis. The panel takes insight in all documents from all involved parties and issues recommendations to stakeholders and the parliament. It is further authorized to commission expert opinions, especially on the basis of economic geological data, which may not be published due to property rights. This right is derived from a law passed specifically on the occasion of the search for a final repository site, which regulates the publication of geological data [GeolDG, 2020]. The NBG is part of the lessons learnt from earlier processes and therefore plays a crucial role.

These different parties interact continually during the selection process. While the process scheme looks straightforward, it is actually an iterative procedure which at any time can respond to mistakenly made decisions. Although there is a definite timeline (Figure 1), it is a guiding principle that diligence takes precedence over time pressure.

It goes without saying that the process will be closely supported and driven by independent research at universities and research facilities, e.g. DAEF (German Working Group on Repository Research).

Germany's State Geological Surveys are involved in the selection process by providing geological data and introducing regional geological expertise. They are, however, not an active part of the site selection.

## 2.2. Brief geological background of Germany

Following the strategy of screening out the more suitable sites and excluding the less suitable ones, the selection started in 2017 with the so-called "White Map of Germany". Naturally, this map should increasingly take on color according to the geological characteristics with regard to a possible HLW disposal site. Three rock types are unanimously considered as potential host rocks for HLW: rocksalt, claystone and crystalline rock, all three of which occur in Germany. Therefore, a short overview of the geology in Germany seems advisable. A detailed description can be found in [Meschede and Warr, 2019].

Germany can roughly be subdivided into three different geological zones. North Germany, as part of the Central European Depression, is covered by up to 200 m thick unconsolidated remnants from the

Pleistocene glaciation period. Locally, sub-glacial channels incise more than 500 m deep into the pre-Pleistocene layers. Among the deep pre-Quaternary sediments, the Permian evaporite formations are of special interest. They occur as flat-bedded salt pillows or as steep salt domes, which have uplifted and partly penetrated the overlying Mesozoic strata. The Gorleben and Morsleben salt domes are prominent examples. Unconsolidated Tertiary and consolidated pre-Tertiary clays are also encountered at relevant depths.

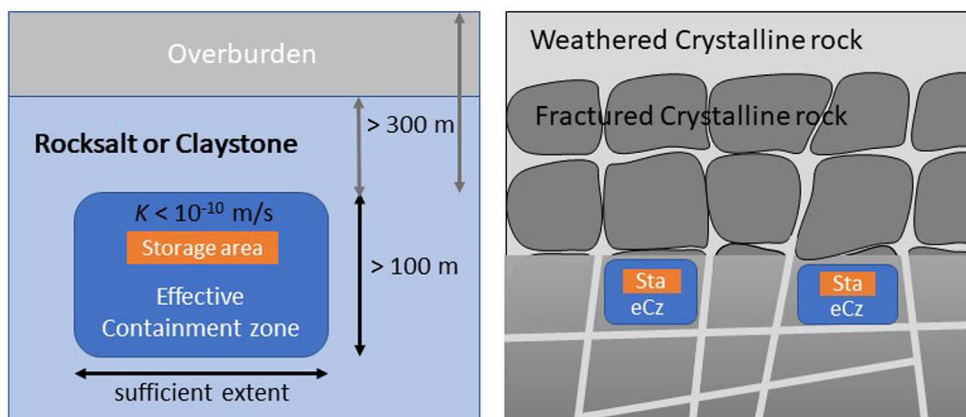
The Low Mountain range, south of the greatest extension of the Pleistocene glaciers, i.e. a line from the Westfalen basin–Hannover–Leipzig, consists of Mesozoic sediments, which are locally penetrated by Tertiary volcanic rocks. The Variscan orogeny is represented by intensively folded metamorphic and igneous rocks (e.g. Black Forest, Bavarian Forest and Ore Mountains). Potential sites in crystalline host rocks are looked up here. The NNE–SSW trending Tertiary Upper-Rhein-Graben is a prominent tectonic structure.

The Alps, which build the southern high mountain border of Germany, accumulated the Tertiary Molasse north of the advancing fold-and-thrust belt. These up to several km thick clastic sediments cover Mesozoic (Jurassic) limestone and clay layers, the latter being potentially interesting for HLW disposal.

## 2.3. Decision criteria according to the Site Selection Act

In the first step of stage 1 of the selection process (Figure 1) the "White Map" of Germany is restricted to subareas where potential host rocks occur. Areas where one or more out of six exclusion criteria (EC) apply are not considered. These EC address uplift rates, seismic, volcanic, tectonic, and former mining activities. Areas where young groundwater has been encountered are also excluded.

If one or more of the potential host rocks are identified, five minimum requirements (MRs) need to be met by the effective containment zone (eCz), which is defined as that part of a rock mass which, in the case of repository systems based essentially on geological barriers, in interaction with the engineered and geotechnical seals, ensures the safe confinement of the radioactive waste in a repository



**Figure 3.** Configuration of host rocks and effective Containment zone (eCz) and Minimum requirements (MR). Left: rocksalt and claystone, right: Crystalline rocks. Modified after [BGR Federal Institute for Geosciences and Natural Resources, 2022].

[StandAG, 2017]. The MRs account for depth, thickness, and lateral extent of the eCz, as well as permeability and preservation of its barrier effect (Figure 3). The ranges of these parameters are chosen very broadly at this stage of the procedure to avoid eliminating an area too early due to lack of available data.

The first step ends by rating the subareas from particularly suitable to less suitable according to 11 geoscientific weighing criteria (gwC). These criteria relate to the following requirements: (i) the achievable quality of enclosure, its robustness and predictability with respect to spatial structure and transport with groundwater, (ii) the protection of the insulation capacity, i.e. tendency to form fissures and potential fluid pathways, and (iii) further safety relevant properties, e.g. temperature resilience, hydrochemical quality or retention capacity for radionuclides.

In the second step of stage 1, all sub areas are subject to “representative preliminary safety investigations” which account for possible technical deposition concepts. Planning weighing criteria (pwC) are first introduced while the gwC are reviewed in more detail. The pwC account for (i) surface features e.g. settlements, shallow aquifers, or flooding areas; (ii) nature conservation areas, deep aquifers or important cultural assets, and (iii) potential mining, geothermal facilities or incident-sensitive assets. The earlier EC, which among others refer to existing mining facilities and young groundwater occurrence, are intended to eliminate potential hazards. The pwC only come into effect in order to decide among

geologically equal favorable sites, now considering other future (or past) economic, environmental or cultural value.

Up to the end of stage 1, the site-selection process relies only on existing information, i.e. geological data, bore logs, geophysical sections, maps, 3D-models of the subsurface and—in case that local information is not available—reference data. No specific above ground or subsurface investigations are undertaken yet.

At the end of stage 1, the parliament of Germany has to decide upon sites for above ground exploration in stage 2 based on the recommendation from BGE. Not until the third stage will subsurface explorations and environmental impact assessment be performed. However, stage 2 and 3 will also apply all above sketched criteria with increasing detail, i.e. based on “further” and “comprehensive preliminary safety investigations”. As mentioned earlier, the whole process is intensively supervised by BASE and critically accompanied by NBG. Public participation of citizens is required by law and is achieved through public conferences and various forums.

### 3. Current results and discussion

Having briefly summarized the process scheme and its selection criteria, the following section will present the recent situation after step 1 in stage 1. Some specific issues and obstacles will be discussed without any claim of completeness.



### 3.1. *Recent situation after step 1 in stage 1, now heading into step 2*

On September 28, 2020, the BGE published its first Interim Report on subareas [BGE, 2020b] that are identified as suitable, i.e. areas that have not yet been excluded from the site-selection process. This report is viewable on the internet and comes along with an interactive map, where all 90 sub areas can be viewed and its characteristics as well as weighing results can be reviewed.

The Interim Report surprises in terms of two findings: first, the site that had been preferred in the past and that had intensively been investigated, the salt dome Gorleben, is definitely excluded from further considerations. And second: more than half of the area of Germany (54%) is still considered suitable as potential repository site (Figure 4). The size of identified subareas varies between 37,000 km<sup>2</sup> (crystalline rocks) covering up to six federal states, and 6 km<sup>2</sup> (salt domes). Only the state of Saarland does not have any locations considered suitable. Following the geology of Germany, many clay stone and salt domes occur in the North German basin, whereas crystalline rocks are identified mainly in Saxony, Bavaria and Baden-Württemberg. Prominent clay layers are also described in Baden-Württemberg and Bavaria. Stratiform layered salt formations are found in Thuringia, Brandenburg and Saxony-Anhalt. Some identified subareas extend below the North Sea and the Baltic.

The results presented in the Interim Report have been discussed in public during a series of three sub-area conferences between February and August 2021. A novelty was that the conferences were organized by a preparation group which was built on the basis of democratic election among the potential participants. Another novelty was that due to the pandemic situation, all three conferences were held digitally or at least in a hybrid format. Up to 1000 citizens, scientists, representatives of regional municipalities and environmental or societal organizations debated with the report's authors (BGE) and the other parties involved in the site-selection process (BASE, NBG). Various thematic sessions had been organized and accompanied by experts on various fields, ranging from geoscience to social science. Participants were able to propose motions, which were voted on democratically. Finally, a report on the discussion results and motions was edited, for further attention by the

operative parties BGE and BASE.

The state geological surveys as well as the Federal Geological Survey (BGR) contributed by detailed comments on the identified subareas lying within their responsibility and expertise.

In the meantime, the operative parties BGE and BASE started step 2. One challenging task for the BGE is to further reduce the subareas to regions for above-ground investigations (Figure 1). The new groups of selection tools, preliminary safety investigations and planning weighing criteria (pwC), are used while the previous decision criteria are repeatedly applied. In parallel, additional geological information, that was not yet fully digitally available, needs to be evaluated. Justified criticism of the Interim Report is incorporated into the ongoing work.

According to the Site Selection Act (2017), public participation is not foreseen until areas for above-ground exploration have been designated as Regional conferences. Thus, a participation gap during step 2 in stage 1 exists. Furthermore, only citizens, municipalities and societal organizations in the identified regions are meant here. Therefore, another task resulting from the sub-area conferences has to be resolved by BASE, being responsible for public participation, i.e., a continued format for nationwide public participation which bridges the gap during step 2 in stage 1 needs to be established.

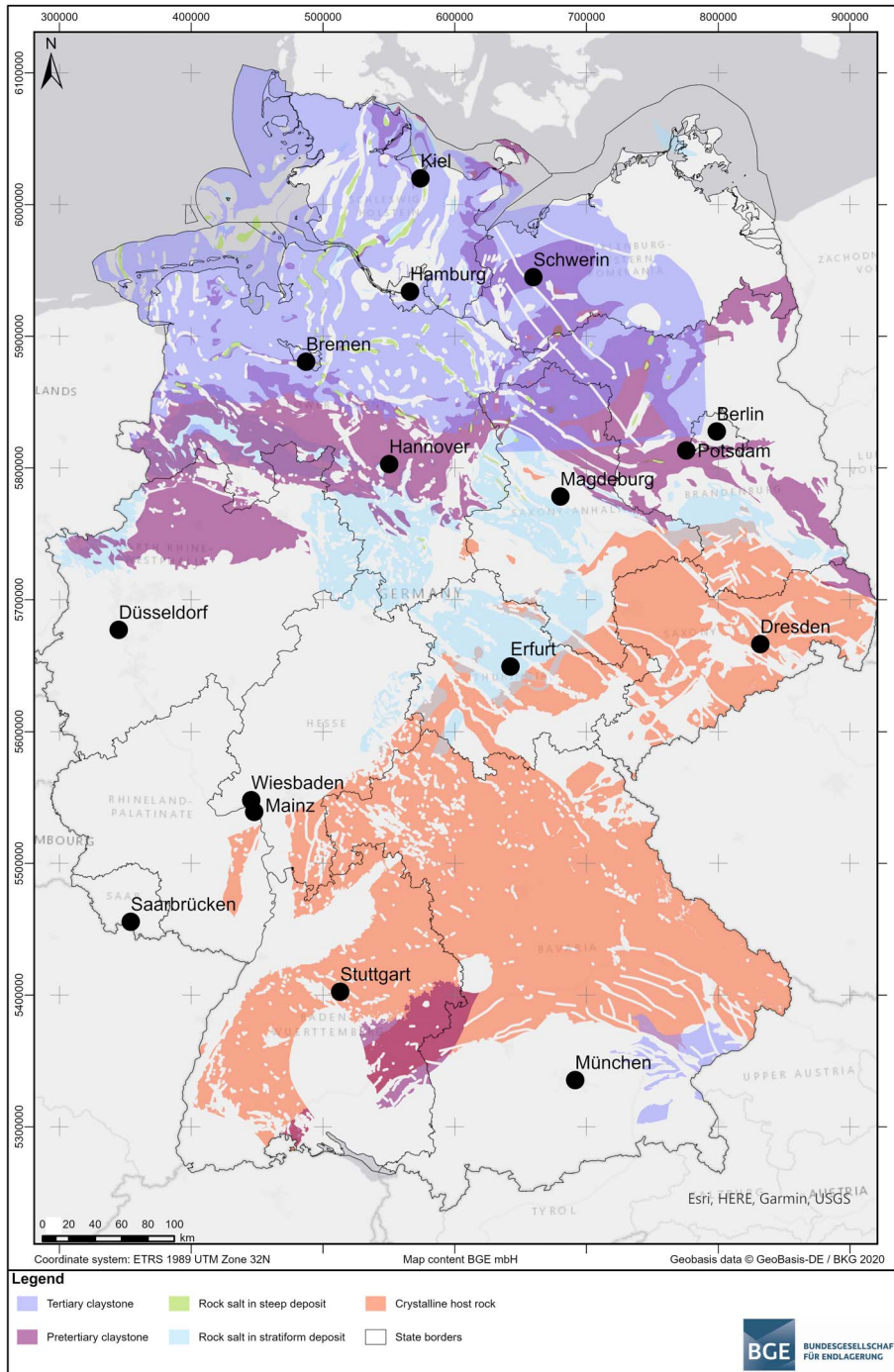
### 3.2. *Selected problems identified at present state of process*

In this section, some selected topics that have been critically debated after the publication of the Interim Report will be briefly addressed, without, however, claiming completeness nor sufficient scientific in-depth discussion—be it in natural science or social science terms. Rather, these examples are presented in order to pinpoint crucial issues which might eventually threaten the site selection process due to misunderstandings and decreasing public trust.

#### 3.2.1. *Problem seismic activity as exclusion criterion*

The Site Selection Act [StandAG, 2017] states that the exclusion criterion seismic activity is applied according to seismic zones defined in [DIN EN1998-1/NA:2011-01, 2011]. The classification considers macro-seismic intensity hazards, calculated as the 10% probability that a macro-seismic intensity is

**Sub-areas pursuant to § 13 Site Selection Act (StandAG)**



**Figure 4.** Subareas, status September 28, 2020 (light lilac—Tertiary clay stone, dark lilac—pre-Tertiary claystone, green—salt domes, blue—stratiform layered salt formations, orange—crystalline rock), after [BGE, 2020b].

exceeded once within 50 years. The classification scheme consists of four classes from 0 to 3 [Grünthal, 1998]. An area is excluded if it falls into a class greater than 1 (i.e. intensity greater than 7). The actual Sub-area map excludes extensive areas in SW Germany (Black Forest and Swabian Alb), NW Germany (Lower Rhine Bay) and two smaller spots in Bavaria (Alpine foothills) and Thuringia.

With this approach, it is worthwhile to take a closer look at the definition of macro-seismic intensity, which corresponds to the strength of tremor and their effects on buildings and people at ground-surface. An expert opinion obtained by the NBG [Wenzel, 2020] concludes that the DIN directive for above-ground buildings cannot assess the potential damage for subsurface structures. Seismic activity may affect the above-ground facilities, which are necessary during the construction phase. Vibrations caused by seismic waves at depths of several 100 m will not affect the subsurface repository except, however, if a fault ruptures the repository directly. Consequently, many countries consider seismic activity as irrelevant to assess the potential risk for subsurface disposal sites, and do not use it as an exclusion criterion [Wenzel, 2020]. Alternative risk estimators can be developed. Since, however, the Site Selection Act explicitly specifies the application of [DIN EN1998-1/NA:2011-01, 2011], it will need a revision of the Site Selection Act, which in turn requires political action.

### 3.2.2. *Problem: depth of repository as minimum requirement and glacial erosion*

One of the MR criteria demands that the eCz must not be at a depth shallower than 300 m (Figure 3). In addition, one of the eleven gwC concerns the protection of the eCz through favorable structure of the overburden against erosion and sub-erosion and their consequences [StandAG, 2017].

The results reported in the Interim Report raise debates among scientists as well as in the public. The depth of the eCz considered over a time span of 1 million years will most likely be subject to changes. Erosion, especially from advancing glaciers in future ice ages expected once every 100 thousand years [Hughes and Gibbard, 2018], or about ten times within the 1 million-year safety margin, must be considered. As mentioned above, sub-glacial channels are observed in North Germany that cut more than

500 m deep into the bedrock. It can be expected that in future advancing glaciers will erode not only areally but also along sub-glacial channels and can thus destroy the barrier effect of the overburden. Ongoing research is currently investigating the controlling geological factors, e.g. lithology and faults of the subsurface, reactivation of existing channel structures, and predictability of the formation of sub-glacial channels [Lang et al., 2021].

### 3.2.3. *Problem: host rocks classification and areal delineation of subareas according to structural zones*

All three types of potential host rocks bear problems when communicated with non-geoscientists. Salt refers not only to Halite and clay stones not only comprise rocks which are composed of clay minerals and particles smaller than 0.002 mm. However, the term “crystalline rock” seems to be the most misunderstood term on the course of the site selection, even among geoscientists. The Dictionary of Mining, Mineral, and Related Terms [AGI, 1997] offers the following definition: “An inexact, but convenient term designating an igneous or metamorphic rock, as opposed to a sedimentary rock.” This in-exactness gives rise to further confusion, especially when communicated to citizens and non-geologists.

Crystalline rocks are any rocks that are made of relatively large mineral grains. These rocks are generally considered as compact and very less permeable. Re-crystallized, coarse-grained marbles, i.e. metamorphosed limestones, however, also belong to crystalline rocks. As soluble calcium carbonate rocks, they are naturally unsuitable as host rocks. Rocks that have experienced high-temperature, high-pressure metamorphosis, are usually referred to as “Gneisses”, and “Granite” is the common term for any plutonic rock, regardless of their mineralogical composition, which ranges from felsic to mafic. Tectonic fracturing and alteration processes by weathering do not affect the classification as “Crystalline rock”, however, they greatly influence the protection capacity against radionuclide transport. Therefore, the identification of subareas as of type “crystalline” bears many obstacles and gives reason for misunderstandings.

Among the seven subareas described as crystalline, the sub-area “Saxothuringikum” is one of the biggest, covering an area of 32,655 km<sup>2</sup>. It

extends from northern Baden-Württemberg eastwards, crossing northern Bavaria, Thuringia, parts of Saxony-Anhalt and finally covering Saxony. The sub-area coincides geographically with the Saxothuringian structural zone. It is only disrupted by small bands or spots, which are not further considered due to EC or MR. This huge structural zone within the Variscan orogeny actually comprises large occurrences of plutonic rocks, mostly granites, as well as high metamorphic rocks, i.e. gneiss. However, it is also characterized by Paleozoic schists, which are not favorable in terms of HLW deposition. Consequently, regional geological surveys criticized the delineation of this sub-area along structural borders rather than to lithological characteristics of rocks that are encountered. For example, the Geological Survey of Baden-Württemberg comments that Paleozoic schists are included in the potentially favorable sub-area “Saxothuringikum”, which should have been excluded already at this stage [LGBR, Geological Survey Baden-Württemberg, 2020].

It can be concluded that the fuzzy understanding of the term “crystalline rock” raises big confusion not only among scientists but even more among citizens, and may thus lower the trust into the selection process and finally the acceptance of a repository. While the problem is most apparent for this big sub-area, the same problems exist for the other two host rock types, which are delineated along stratigraphic extent.

### 3.2.4. *Problem data: availability, suitability for the site selection process and public transparency*

The first stage in the German site-selection process (Figure 1) is based on existing information only, i.e. until now no specific investigations in terms of favorability for an HLW repository have been undertaken, except for the sites that had been assigned in the past: e.g. the salt dome Gorleben, or the Asse and Morsleben (for low- and intermediate-level nuclear waste, LILW). Information on the subsurface comprises measured data, e.g. drill-logs, geochemical and hydrochemical analyses, geophysical sections, as well as interpretations, e.g. geological maps and 3D-geological models. Generally, this information is summarized under the term “geological data”. It is one of the tasks of the state geological surveys in Germany to collect and store all regional geological data

and information. Within the site-selection process, they have to deliver all relevant data to the BGE.

The operating federal company (BGE) faces a number of obstacles here: apart from various data formats including non-digitized data, which still need to be unified, the geological data have been collected over a time span of decades and thus have different qualities. The spatial distribution of available data is also not uniform. It becomes apparent that regions that were promising in terms of geological resources show a much denser data distribution. Since, however, the aim of resource exploration was not to identify an HLW repository, the measurements may often not exactly provide the needed information. For example, the salt domes in the North German basin are considered as potential oil traps. Therefore, their outer shape is more important for hydrocarbon exploration and can well be estimated by geophysical sections. In terms of the repository, the inner structure and mineralogy of the evaporites in the salt dome is relevant, but was not specifically investigated.

The delineation of potentially favorable rock formations in Germany in the first step of stage 1 is based on checking EC and MR and subsequently weighing geological criteria. One criterion is the potential transport of radionuclides by groundwater in the eCz. Here, the knowledge of parameters, e.g. hydraulic conductivity, porosity diffusion coefficient and sorption parameters, and their spatial variability is crucial, but, not known for the specific region. In those cases, reference data sets, derived from original publications, textbooks or from site investigations abroad are used to define data ranges. The idea at this early step is to apply rather wide limits in order to not exclude areas, before more detailed knowledge is gained. If, for instance for a reference rock, a hydraulic conductivity of less than  $10^{-10}$  m/s is reported, the potential host rock is still considered “conditionally favorable”; for less than  $10^{-12}$  m/s, it is seen as “favorable”. Combining the assessment results for all parameters (so called indicators) helps to segregate better favorable from less favorable subareas.

The result of the broad parameter estimation on which the suitability assessment is based is, however, that 54% of Germany’s surface area is currently still identified as potential siting regions (subareas). Whereas the concept was generally accepted by ex-

pert opinion, the quality of reference information has been criticized as “partly not state-of the art”. It is, however, also stated that at this early step the site-selection process allows for a rough view, provided that during the repeated application of geological weighing criteria, site-specific information and updated references are used [Kühn, 2021].

Another problem is the ownership of geological data and interpretations, e.g. models: especially the exploration data are owned by private companies and may not be publicly available. These data are known to the geological state surveys, and are supplied to the BGE to be used in the site-selection process. However, in order to meet all five main principles of the new selection process, i.e. citizens’ participation, science-based, transparency, self-questioning and learning [StandAG, 2017], all selection steps, methods documents, and data must fully be viewable and verifiable to everyone.

This apparent conflict between private ownership and transparency was recently resolved by adopting a new law [GeolDG, 2020], which rules the handling of geological data. This law grants the expeditious publication of state geological data while protecting private ownership of exploration data. However, in case of general welfare also those private geological data, which are relevant for the site selection, must be publicly available. The national task to identify an HLW repository site is considered as a case of general welfare. At the current state of the site-selection process, huge amount of information cannot still be publicly viewed due to non-clarified categorization pursuant the Geological Data Act. It has therefore been ruled by the Site Selection Act that in the meantime the members of the National Citizens’ Oversight Committee and their geological experts are exclusively permitted to view and examine unpublished data in a protected data-room provided by the BGE, thus representing the citizens [StandAG, 2017].

### 3.2.5. *Problem: how to reduce the big extent of subareas to smaller regions for above ground investigation?*

Currently, i.e. in the spring of 2022, the site-selection process is in a crucial step. While the large extent of identified subareas raised serenity on one hand (“The areas are still large, so it is very unlikely that the site will be built in my region!”), it on the other hand raised suspicion of “non-meticulous”

work among citizens and professionals. These concerns have been expressed during the sub-area conferences. As mentioned above, expert opinions [Kühn, 2021, Wenzel, 2020] concur that in the first step of stage 1 the process is justified. However, at the end of the step 2 of stage 1, regions for the above-ground investigations must be proposed, on which the parliament has to decide before the site-selection process can continue in stage 2.

The federal company for radioactive waste disposal (BGE) is currently developing decision strategies on how to further limit the potential areas. Here, the scientists in the BGE benefit from additional digitized data two years after the publication of the Interim-Sub-area Report [BGE, 2020b], additional decision criteria which include the technical configuration of the subsurface disposal mine, and numerous professional and citizens’ comments to the Interim Report. Since all subareas in their full extent need to be assessed, the task is immense. The strategies are currently tested at four selected subareas, one for each host-rock type. These example-sites, so called Regions for the Development of Methods, are chosen according to factors, which can negatively affect the site-selection procedure, e.g. large size of subarea, lack of data, non-uniform data distribution, and lithological heterogeneity.

### 3.2.6. *Problem: comparability of different host rocks*

The German site-selection process aims at absolute fairness for all regions, states and municipalities. Subsurface, geological conditions dominate the selection process, rather than political, economic, cultural or infra-structure concerns—these will only be considered at a later stage. Whereas other countries either do not have all three host-rock types or focused their site selection early on a specific type (e.g. Finland: Olkiluoto, crystalline basement; Switzerland: Jurassic Opalinus Clay), the geological conditions in Germany offer all three types spread all over the country. This may be seen as a fortune; however, it puts the problem of comparability. Each potential host rock bears advantages and disadvantages, which cannot linearly be combined in one decision criterion. Above this, each rock type requires individual containment features, be it in terms of spatial configuration, geotechnical and/or technical barriers. So far, the site-selection procedure does not

offer any scientifically based method on how to answer the question whether site A in host-rock X, given the containment configuration 1, is more favorable than site B in rock Y with configuration 2. It is this open question that may threaten the success of the site-selection process later.

### 3.2.7. *Problem: continued citizens' participation*

Whereas many geoscientific and safety questions still need to be resolved and planning scientific criteria need to be included before the final decision is made, the novel aspect of full and continuous citizen participation based on maximum transparency is one of the most challenging new requirements, which are formulated in the Site Selection Act [StandAG, 2017]. This has been documented by the numerous concerns expressed during the sub-area conferences. While the law is very specific on when and how citizen participation is granted, the meaning of the term “participation” is not clear to all engaged parties.

Participation means the act “of taking part or become involved in something” [Cambridge Dictionary, 2022]. “Something” in this context means the selection process. Participation presupposes but is not limited to information and transparency, rather it means that citizens are part of the selection process. That, however, means that everybody must be conditioned such that he or she understands the process and can judge it. It also means that the citizens are no longer just recipients of information and decisions, rather they can independently contribute to information and decisions by their own judgement. In consequence, it is the task of all professional stakeholders to interest citizens and to provide sufficient technical knowledge. This is truly a challenging task, as the average citizen is not a geologist or radiation physicist and an appropriate but non-academic language needs to be used to bring everyone along. Also, it means a lot of personal engagement from the citizens to spend many free time hours over years to actively participate in conferences, advisory groups and discussion fora. As such, the site-selection process is not just a geoscientific-technical process, it also requires a lot of social science expertise.

It is also a question of who is actually addressed? Prospecting an evaluation time span of one million years it needs the engagement of future generations,

which have never profited from nuclear power nor experienced the political and societal debates in the past. Here, it is our generation's task to develop specific formats and new media, to attract young people's attention, who will soon have to deal with this problem.

As described above, the Site Selection Act provides several participation formats along the process line; however, these are concentrated to specific periods. The first format is the Sub-area conference after the publication of the Interim Report, which was limited to a time span between September 2020 and August 2021. During the sub-area conferences, participants expressed their worries regarding the geoscientific reasoning. However, most concerns are related to the guiding principles of the selection process, i.e. transparency and the possibility of continuous citizens' participation because the next period of participation only starts after regions for above-ground investigations are determined. The so-called Regional conferences are aimed exclusively at citizens and stakeholders from the designated regions and directly adjacent national or transboundary municipalities, i.e. during the step 2 of stage 1, when the narrowing down of the subareas takes place on the basis of gwC and, if applicable, pwC, as participation formats are not explicitly established by the Site Selection act.

Since this so-called “Participation gap” has been identified during the first sub-area conferences a new bridging format is developed in cooperation between BASE (Federal Office for the Safety of Nuclear Waste Management) and other involved parties, citizens, representatives of societal organizations and municipalities, BGE and NBG, the latter in its surveying role. Establishing this new format requires an open willingness to engage in dialogue and compromise among all parties. However, if successful, it is a powerful tool to enhance trust and acceptance in the site-selection process. Moreover, it documents one of the five principles, i.e. that the site selection must be understood as a learning process.

## 4. Conclusion

After a series of unsuccessful attempts by Germany to designate a suitable site for the final storage of nuclear waste, a newly designed selection procedure is now being used to identify a site that, compared to

all others in Germany, offers the most favorable conditions for safe storage for one million years. Rather than taking a top-bottom approach on a site location, the process is based on democratic principles. In this approach, it reacts to bad experiences from the past and follows Switzerland, where a similar strong anti-nuclear resentment exists.

The site-selection procedure puts the geological suitability at first place, rather than regional socio-economic interests. However, the fact that three potential host rock types need to be considered, many relevant data are not yet available and some selection criteria possibly have to be redefined, makes the site selection very complicated. In consequence, the process is accompanied by many research activities. The research questions comprise subjects like radio-nuclide migration, glacial erosion processes, fissure forming in crystalline rocks and its conceptual modeling, development of numerical and stochastic models in order to simulate Hydraulic–Thermal–Mechanical–Chemical (HTMC) processes, age-dating methodologies, characteristics of clay-minerals or salt slack as a geotechnical barrier, just to mention a few.

The site-selection process is designed such that public participation must be warranted. In order to ensure this as well as the other attributes of the site-selection process, a National Citizens' oversight committee (NBG) was formed as a novel, and so far, unprecedented instrument of an independent honorary process surveyor. This committee accompanies the entire search process until the site for a repository has been found and the German parliament has decided on it. The members of the NBG are appointed for three years, but can be extended. The NBG detects possible conflicts, acts as a mediator, issues recommendations, and represents the public. Due to its pluralistic internal composition, which includes citizens, young people, natural and social scientists, theologians, retired politicians and non-governmental environmental associations' representatives, a maximum range of thought impulses is guaranteed.

The site-selection procedure is a task that affects the whole of Germany and its citizens. From a geological perspective, the procedure offers a new wave of geoscientific job opportunities and a new public recognition of geology as an applied science. However, it is also a major societal challenge that car-

ries a strong integrative potential due to the shared commitment. It is the firm conviction of all involved parties that Germany will ultimately operate a repository site that offers optimal safety conditions and that is democratically accepted by German residents. Revisiting Ghislain de Marsily's quote, it is in this sense, that we do not blame the Federal Republic of Germany—nor any other country—for its attempt to build the best protection “wall” against any environmental hazard, which may stem from careless handling of nuclear waste, now or in future.

### Conflicts of interest

The author has no conflict of interest to declare.

### Acknowledgments

The author wishes to acknowledge her co-members of the National Citizens' Oversight Committees and its office for the support while gathering site-selection process documents. The informative web pages of BGE, BASE and BGR were also a great help.

### References

- AGI (1997). *Dictionary of Mining, Mineral, and Related Terms*. American Geological Institute, Alexandria, VA, 2nd edition.
- AtG (2011). Atomgesetz. Gesetz über die friedliche Verwendung der Kernenergie und den Schutz gegen ihre Gefahren (Act on the peaceful use of nuclear energy and protection against its dangers). <https://www.gesetze-im-internet.de/atg/>. Germany, DE.
- BGE (2020a). *Das Ende der weißen Landkarte (The end of the white map)*. Einblicke - Das Magazin der Bundesgesellschaft für Endlagerung, No. 6. EMPUS CORPORATE GmbH, Berlin.
- BGE (2020b). Sub-areas Interim Report pursuant to Section 13 StandAG. <https://www.bge.de/en/sitesearch/sub-areas-interim-report/>. [Online].
- BGR Federal Institute for Geosciences and Natural Resources (2022). BGR-Endlagerung radioaktiver Abfälle: Langzeitsicherheit (BGR-The final disposal of radioactive waste: Long-term safety). <https://www.bgr.bund.de/DE/>

- Themen/Endlagerung/Langzeitsicherheit/langzeitsicherheit\_inhalt.html?nn=1542224. [Online].
- Bundeszentrale für Politische Bildung (2020). BPB Kurz&knapp. <https://www.bpb.de/kurz-knapp/hintergrund-aktuell/310887/vor-40-jahren-raeumung-der-republik-freies-wendland/>. [Online].
- Cambridge Dictionary (2022). *Cambridge Dictionary*. Cambridge University Press, Cambridge, UK, <https://dictionary.cambridge.org/dictionary/english/participation>. [Online].
- Däuper, O., Bosch, K., and Ringwald, R. (2013). Zur Finanzierung des Standortauswahlverfahrens für ein atomares Endlager durch Beiträge der Abfallverursachern (On financing the site selection process for a nuclear repository through contributions from waste producers). In *Zeitschrift für Umweltrecht (ZUR)*, volume 6, pages 329–335. Nomos Verlag, Baden–Baden, DE.
- de Marsily, G. (1982). Predictive geology – with emphasis to nuclear waste disposal. In de Marsily, G. and Merriam, D. F., editors, *Proceedings IAMG at 26th IGC Paris 1980*. Pergamon Press, Oxford, UK.
- DIN EN1998-1/NA:2011-01 (2011). *National Annex - Nationally determined parameters - Eurocode 8: design of structures for earthquake resistance - part 1: general rules, seismic actions and rules for buildings*. DIN Deutsches Institut für Normung e.V. Berlin, Beuth, Berlin, DE.
- EndKom (2016). Abschlussbericht: Verantwortung für die Zukunft - Ein faires und transparentes Verfahren für die Auswahl eines nationalen Endlagerstandorts (Final Report: Responsibility for the Future- A Fair and Transparent Procedure for the Selection of a National Repository Site). In *Komm. Lagerung hochradioaktiver Abfallstoffe Deutscher Bundestag*, page 683. BMUV, Berlin.
- EWN (2022). Entsorgungswerk für Nuklearanlagen (Nuclear Waste Management Plant). <https://www.ewn-gmbh.de>. [Online] [Zitat vom: 25. February 2022].
- GeolDG (2020). Geological Data Act. Gesetz zur staatlichen geologischen Landesaufnahme sowie zur Übermittlung, Sicherung und öffentlichen Bereitstellung geologischer Daten und Zurverfügungstellung geologischer Daten zur Erfüllung öffentlicher Aufgaben (Act on the State Geological Survey and on the Transmission, Safeguarding and Public Provision of Geological Data and Provision of Geological Data for the Fulfilment of Public Duties). Germany, DE., <https://www.gesetze-im-internet.de/geoldg/index.html>.
- Grünthal, G., editor (1998). *European Macroseismic Scale 1998 (EMS\_98)*. Centre Européen de Géodynamique et de Séismologie, Luxemburg, LU.
- Hughes, P. D. and Gibbard, P. L. (2018). Global glacier dynamics during 100 ka Pleistocene glacial cycles. In *Quaternary Research*, volume 90, pages 222–243. Cambridge University Press, Cambridge, UK.
- IAEA (2009). *Classification of Radioactive Waste - General Safety Guide*. IAEA Safety Standards Series No. GSG-1. International Atomic Energy Agency, Vienna, AT.
- Kühn, M. (2021). Entsprechen die Referenzdatensätze, die die BGE zur Anwendung der geowissenschaftlichen Abwägungskriterien teilweise eingesetzt hat, dem Stand von Wissenschaft und Technik? (Do the reference data sets that the BGE has partially used to apply the geoscientific weighing criteria correspond to the state of the art in science and technology?). [https://www.nationales-begleitgremium.de/SharedDocs/Downloads/DE/Downloads\\_Gutachten/Gutachten\\_Referenzdaten\\_Kuehn\\_10\\_5\\_2021.html](https://www.nationales-begleitgremium.de/SharedDocs/Downloads/DE/Downloads_Gutachten/Gutachten_Referenzdaten_Kuehn_10_5_2021.html). Expert opinion commissioned by National Citizens' Oversight Committee (NBG). [Online].
- Lang, J., Bebiolka, A., and Breuer, S. (2021). Was haben Eiszeiten mit der Sicherheit eines Endlagers zu tun? (What do ice ages have to do with the safety of a repository?). In *Oral presentation at DEUQA (German Quaternary Assoc)*. BGR, Hannover, DE.
- LGBR, Geological Survey Baden-Württemberg (2020). Diskussionen und Fachdebatte (Discussions and technical debate). <https://www.bge.de/de/endlagersuche/wesentliche-unterlagen/fachdiskussionen/>. [Online] December 1, 2020. [Cited: March 20, 2022].
- Meschede, M. and Warr, L. N. (2019). *The Geology of Germany - A Process Oriented Approach*. Springer Nature AG, Berlin, DE.
- Pflugbeil, S. (2006). Alle Folgen liquidiert? Die gesundheitlichen Folgen von Tschernobyl (All Consequences Liquidated? The Health Consequences of Chernobyl). *Osteuropa*, 56(4), 81–104.
- Reiche, S., Fink, R., and Nilius, N.-P. (2021). Geoscientific Characterisation and Interpretation



- (Geosynthesis) within the Preliminary Safety Assessment in the German Site-Selection Procedure for a High-Level Nuclear Waste Repository. <https://www.bge.de/de/endlagersuche/wesentliche-unterlagen/vortraege/>. [Online] [Cited: March 4, 2022].
- StandAG (2017). Standortauswahlgesetz (Site Selection Act). Gesetz zur Suche und Auswahl eines Standortes für ein Endlager für hochradioaktive Abfälle (Standortauswahlgesetz - StandAG). [https://www.gesetze-im-internet.de/standag\\_2017/index.html](https://www.gesetze-im-internet.de/standag_2017/index.html). Germany, DE.
- Stude, S. (2019). *Strom für die Republik - Die Stasi und das Kernkraftwerk Greifswald (Electricity for the Republic - The Stasi and the Greifswald Nuclear Power Plant)*, volume 53. Vandenhoeck and Ruprecht GmbH and Co. KG, Göttingen, DE.
- Wenzel, F. (2020). Prüfung Ausschlusskriterium Seismische Aktivität (Audit of Exclusion Criterion Seismic Activity) - Expert opinion commissioned by National Citizens' Oversight Committee (NBG). [https://www.nationales-begleitgremium.de/SharedDocs/Downloads/DE/Downloads\\_Gutachten/Gutachten\\_seismische\\_Aktivitaet\\_14\\_9\\_2020.pdf?\\_\\_blob=publicationFile&v=11](https://www.nationales-begleitgremium.de/SharedDocs/Downloads/DE/Downloads_Gutachten/Gutachten_seismische_Aktivitaet_14_9_2020.pdf?__blob=publicationFile&v=11). [Online].





Research article

Geo-hydrological Data & Models

# The Anthropocene is shifting the paradigm of geosciences and science

Anne Coudrain<sup>Ⓢ,\*, a, b</sup>, Matthieu Le Duff<sup>Ⓢ, b, c</sup> and Danielle Mitja<sup>Ⓢ, a, b</sup>

<sup>a</sup> Institut de Recherche pour le développement, France

<sup>b</sup> UMR 228 Espace-Dev (IRD, Univ Montpellier, Univ Guyane, Univ La Réunion, Univ Antilles, Univ Nouvelle Calédonie), Maison de la Télédétection, 500 rue JF Breton, 34093 Montpellier, Cedex 5, France

<sup>c</sup> Centre Universitaire de Formation et de Recherche (CUFR) de Mayotte, 97660 Dembeni, Mayotte, France

*E-mails:* [anne.coudrain@ird.fr](mailto:anne.coudrain@ird.fr) (A. Coudrain), [matthieu.le\\_duff@univ-mayotte.fr](mailto:matthieu.le_duff@univ-mayotte.fr) (M. Le Duff), [danielle.mitja@ird.fr](mailto:danielle.mitja@ird.fr) (D. Mitja)

**Abstract.** Noting that humans are affecting the Earth's envelopes, with geosciences becoming inseparable from social sciences and humanities, and acknowledging the increasing use of the word Anthropocene, the authors of this paper explore the possibility of a scientific paradigm shift. (1) Since the 17th century, modern science has developed in a context of naturalist worlding that favors its hegemony over other modes of existence and its paradigm of a quest to define the laws of nature. (2) Various manifestations attest to the emergence of a paradigm of knowledge diversity and the unravelling of the naturalist worlding. (3) The current boom in participatory science is a sign of this paradigm shift: through the fundamental changes to the Earth System that refer to it, the Anthropocene forces science to move towards action and interact with a society involved in adapting to changes and concerned with the threats to its territory's habitability.

**Keywords.** Anthropocene, Scientific paradigm, Complexity, Plurality of worldings, Participatory science.

*Manuscript received 29 June 2022, revised 29 August 2022, accepted 6 September 2022.*

## 1. Introduction

The aim of this article is to examine whether the context of the Anthropocene influences science and its universalist paradigm. The authors are neither specialists in philosophy nor in epistemology. Yet curious and aware of the importance of this issue in their own work, they are engaged in a reflection and an analysis of their practices and of specialized publications. The editors of this journal [de Marsily and

Chabaux, 2020] have recently published an editorial reflecting their decision to broaden their editorial policy beyond the traditional fields of Earth sciences through the publication of articles dealing with “planetary sciences” in a broader sense. Encouraged by this editorial, the present authors are delighted to share their thoughts with the readers of the journal “*Comptes Rendus Géoscience*”.

Our first reflection concerns environmental risks which have been present in all areas of life for several decades. As early as 1980, de Marsily [1982] reported that “the rapid progress of technology has made our world increasingly fragile, so that unforeseen local

\* Corresponding author.

events may have deleterious effects on an entire country or even a continent” in his introduction to “Predictive Geology with emphasis on nuclear-waste disposal”. He wrote this statement soon after the Seveso disaster in Italy, which prompted the adoption of legislation on the prevention and control of such accidents. The so-called Seveso-Directive (Directive 82/501/EEC) was later amended in the light of lessons learned from subsequent accidents. The future has proven de Marsily right. Since then, risk is commonly invoked in all scientific fields on a territorial or a global scale. The scientific community has been increasingly aware of the need to address the risks associated with what is commonly known as the Anthropocene.

The word Anthropocene first appeared in the discourse of globalisation during a scientific conference in 2000. The chemist and Nobel laureate Paul Crutzen suggested its use because he found that the extent of human activities was such that the creation of a new geological era was needed [Crutzen and Stoermer, 2000].

The word has followed two distinct paths since then. The first being an academic path within the Subcommission on Quaternary Stratigraphy of the International Union of Geological Sciences. The Anthropocene remains a topic of the active study of a working group charged with determining the distinction of a new geological time unit, based on the human impacts identified in geological records. Members of this working group argue for an Anthropocene epoch distinct from the Holocene [Syvitski *et al.*, 2020, Zalasiewicz *et al.*, 2015]. The argument considers changes in the state and functioning of the Earth System that are (1) beyond the Holocene’s range of variability, and (2) driven by human activities and not by natural variability [Steffen *et al.*, 2015]. This epoch started during early years of the 20th century [Haustein *et al.*, 2019].

The second path is the use of the word by society at large, including scientific communities. The frequency of occurrence of the keyword “Anthropocene” has steadily increased in scientific publications since 2002. In 2018, it reached about 0.4, i.e., 40 occurrences per 100 publications [Gemene *et al.*, 2019]. Its frequency parallels that of “Earth System”, which was at 0.55 in 2018, whereas this keyword was already well established in scientific publications in 1977 with a frequency of 0.02. This popularity il-

lustrates the need for a distinction. Many alternative names for this era have been suggested, including those proposed by Haraway [2015], who, while highly critical of it, recognizes the usefulness of the term “Anthropocene.” This word provides a platform to gather information and current preoccupations. Thus, the absence of an official definition from the academic geological community does not preclude the fact that this term can engage and encourage scientific thought and work.

In this paper, the word Anthropocene is used to define the period of history during which humans have become active agents in the planetary system, capable of hindering the physical mechanisms governing the way the ecosystems function in the general world [Descola, 2015, Steffen *et al.*, 2011]. The issues we raise are the following: the hole in the ozone layer, global warming with all its consequences, the erosion of biodiversity through the extinction of animal and plant species on the planet’s surface, and health considerations, with humans poisoning themselves through their own activities [Grau *et al.*, 2022]. We could then read in the Anthropocene a *mise en abyme* of our own vulnerability, or to quote Serres [2020, p. 8]: « Nous avons récemment appris, au risque de notre survie, que nous dépendons désormais de choses qui dépendent de nous ».<sup>1</sup>

This article is divided into three parts. The first part reviews few scientific paradigm shifts in the past. It highlights the extent to which science and the work of researchers are embedded in the way societies deal with technology and the environment. It reminds us that abstract universalism dominates the paradigm of Western science and the relationship between humans and environment. Part one leads the authors of this article to conclude that a shift in the scientific paradigm is a valid assumption in view of the extraordinary impact that human activities have on the environment.

The second part “Towards a new scientific paradigm: knowledge diversity” explores how the encounter with other non-Western cultures leads us to question the universal, where it seemed to be most established, i.e., that science excludes contingency, and to explore how the context of the Anthropocene

---

<sup>1</sup>Our translation: “We have recently learned, at the risk of our survival, that we are now depending on things that depend on us”.

leads towards a more humble, more inclusive science, more responsive to the singularity of beings and the diversity of forms of knowledge.

The third part “Strengthening participative science in the context of the Anthropocene” considers citizen involvement in the changing conditions of the Earth’s habitability as it seems necessary to mobilize all spheres of society. It is based on a review of publications and on our own practice within a recent group dedicated to “Science and Society” [Mitja *et al.*, 2021] established in the Espace-Dev research unit.<sup>2</sup>

## 2. Review of few previous scientific paradigm shifts

Targeting the scientific paradigm shift on the face of Anthropocene calls for the exploration of few of the paradigm shifts that have preceded this period. This part spells out a few examples. It demonstrates the relative banality of such shifts as their occurrence is relatively high. It also highlights how the legacy of some of these changes constitutes obstacles to change (the concept of the universal, positivism) while, on the contrary, others are breakthroughs that favor it (the theory of evolution, the concept of paradigm shift).

### 2.1. *From the concept of the universal to the laws of Nature*

In the philosophical tradition associated with Plato and Aristotle, intellect had the power to reach the very essence of things beyond their contingent appearances, and translated this power into concepts, explains Stengers [1991], an expert in the philosophy of science. Intellectual knowledge was therefore naturally relevant, free from passions and doubts, and accessible to any being endowed with intellect.

At the height of modern science, the concept of the universal was developed by many authors [Jullien, 2016, p. 8]. That it was a requirement of thought developed by classical Greeks was the source of its immense efficiency in Europe. The Greek “logos” of the centuries preceding the Common Era, writes Cassin [2016], a philologist and specialist in

the rhetoric of modernity, is the starting point for the claim to universality. François Jullien explains that science in Europe was founded on the fact that the abstract universal constituted knowledge. He specifies the following points: (i) “l’opinion commune envisage les choses sur le mode du contingent, c’est-à-dire de ce qui peut être autrement qu’il est” while (ii) “la science envisage les choses sur le mode du nécessaire, donc de l’universel, c’est-à-dire de ce qui ne peut être autrement”<sup>3</sup> [Jullien, 2016, p. 18]. The universal laws of nature have become an essential foundation of science in connection with this concept of the universal and of the evolution of a geographical view of the world. In the Western Middle Ages, the importance of religion and Catholicism made Jerusalem the centre of the world [Clerc *et al.*, 2019]. The importance of this vision is necessary to understand the major role played by subsequent discoveries. Copernicus (16th century), followed by Galileo, Kepler and later by Newton reversed this perspective by considering the Earth as part of a larger system. To define this larger context, Galileo used the special language of mathematics. This language, later mobilised by physics, makes it possible to decode this universe and extract its fundamental laws. It is from this universal that classical Europe, transposing it from mathematics to physics (Galileo, Newton), conceived the “universal laws of nature” with the success we know.

### 2.2. *Positivism driven by industrialisation and faith in progress*

Developed during the 19th century, positivism continues to a large extent to permeate present-day science. Auguste Comte [Mill, 1868] developed this philosophical doctrine in the early days of the Industrial Revolution. As presented by Stengers [1991], the example of chemistry shows its transformation from

<sup>2</sup>Espace-Dev research unit: Space observation, models and actionable science, <https://www.espace-dev.fr/>.

<sup>3</sup>Our translation: (i) “common opinion envisions things in the mode of contingency, that is to say, of what can be otherwise than it is”, and (ii) “science envisions things in the mode of what is necessary, and therefore of the universal, that is to say, of what cannot be otherwise”.

a science of nature into pure science. Chemistry became a “real” and powerful science in the 19th century after discovering the idea of chemical combination between a limited number of unit types. Prior to this, chemistry was a labyrinth of varied, poorly reproducible processes because it used non purified products. Industrial development enabled chemistry to work with pure products—it created its own components with properties that were independent from circumstances—after economic and social transformation led from craft to large-scale industry.

Positivism went beyond the strict framework of science to spread to the institutions of the Third Republic and into society itself [Weber, 1899]. According to positivism, scientific knowledge is made up of laws based on experiments, which one tries to generalize.

With the Industrial Revolution, the Western model imposed itself everywhere. With the help of science and technology, it sought to encourage humans to transcend their dependence on nature. Latour’s illuminating exploration of Hobbes’ *Leviathan* [2015, pp. 194–198] is quite explicit on this point.

Since the Renaissance, European powers have gradually built up a power structure that dominates the planet, with transportation and trade networks gradually covering oceans and continents. This was accomplished by an extremely small fraction of the population within the power structure, as well as by harnessing available energy. When the mastery of fossil fuels allowed European nations to strengthen their dominant positions, acquired through several centuries of colonial domination and capitalism [Chakrabarty and Chalier, 2018], a new relationship with the world came about.

The idea of progress became a source of salvation, and the human became superhuman in the West, and beyond. The impact on the planet then became increasingly important, leading to the modification of the planetary system, which had been in a state of relative equilibrium throughout the Holocene. Today the system is about 1 °C warmer with a degraded biosphere. If crossed, the threshold of around 2 °C makes self-reinforcing feedbacks possible. Even if human emissions were reduced, these feedbacks could then cause continued warming, the boundaries and implications of which are still formidable and unknown [IPCC, 2022a, Steffen *et al.*, 2018].

### 2.3. *Time, chance, and interdependence come into play*

In the 19th century, a major revolution in scientific thinking took place [Nouvel, 2020, p. 279]. Science acquired a new paradigm. Time changed Nature and gave it a narrative pieced together in Lyell’s “Principles of Geology (vol 3)”, [2019, first published in 1833], captioned “Being an Inquiry How Far the Former Changes of the Earth’s Surface are Referrable to Causes Now in Operation.”

Darwin, who entertained a friendly relationship with Lyell, incorporated this new paradigm into his theory of evolution. With his *On the Origin of Species*, Darwin [1859] wrote the first scientific discourse on origin. Pascal Nouvel specifies that, before Darwin, “le vivant ne peut s’expliquer qu’en faisant intervenir une causalité finale (descendante), une intention... parce qu’il est inconcevable que le temps, si étendu qu’on puisse imaginer qu’il fut, ait pu

produire les formes vivantes que nous connaissons”<sup>4</sup> [Nouvel, 2020, p. 269].

Darwin’s work initiates several breaks [Hoquet, 2009]. Nature obeys laws that are not entirely reducible to mathematical laws, and chance comes into play, a step away from the trend initiated in the 17th century. Moreover, humans lose their exceptional status: species emerge only through a process of progressive differentiation. Finally, this differentiation establishes a kinship between species. We believe that these breaks are essential to build a new scientific paradigm in the face of the Anthropocene (cf. part 2).

#### 2.4. *“Paradigm shift”: a concept that facilitates change*

In the second half of the 20th century, the physicist and philosopher Kuhn [1970] introduced the concept of “paradigm shift”. The paradigm shift may concern a minor modification, such as the use of an instrument. It can also concern a major change, requiring a new mode of narration, which is what we are exploring here in the face of the Anthropocene.

The depth of change brought by this concept can be substantiated in several ways. First, this concept appears today as necessary to the scientific community which uses it widely. For example, this expression has an occurrence of more than 400 per year since 2017 in titles of documents referenced in the “Web of Science” database.<sup>5</sup>

In essence, this concept encourages a rethink. It is no longer a value judgment of what true science is. It is not (or no longer) a question of moral order or of the absolute opposition between science and opinion. It is the context of production itself that is considered [Stengers, 1991].

For Kuhn, scientific discourse is linked to the community and the context in which it is produced. A paradigm revolution is therefore an evolution of the way of looking at an object from a new conceptual frame of reference. For the scientific community,

sharing a paradigm means “seeing as” and this community that “sees like” will then come together to share an intense interest and work.

A paradigm shift occurs when the framework of thought is confronted with too many anomalies that can no longer be ignored and that lead to the formulation of new hypotheses and concepts to apprehend differently the questions that remain unresolved. A phase of crisis is characterized by an effervescence of new common frameworks proposed and questioned until a new paradigm gathers enough members of the concerned community to succeed in its adoption by the majority, to then return to a phase of stabilization. In this sense, the shift from geocentrism to heliocentrism mentioned earlier, or the theory of evolution, constitute paradigm shifts, because they upset our representation of the world and of the place of humans in it.

The establishment of this concept of paradigm shift is part of a period in which systems thinking and the complexity approach have developed as a critique of approaches that had been inherited from positivism. Considered too linear and analytical, these approaches nonetheless persist. Bourdieu [1976, 1997] has widely discussed and underlined the limits of a science that was considered objective, as stemming from the heritage of 19th century positivism [Mill, 1868]. The knowledge produced must be considered in its context: politically, socially, and historically. Furthermore, Daston and Galison’s [2012] magisterial contribution to objective science leads us to identify how each major regime of thought defines in its own way the aspects necessary for the acquisition of proven knowledge [Latour, 2012b].

Popper [1963] argues that the aim of science is not so much to state the truth as to construct certainties from the demonstration of what is false. Thus, the scientific community does not have to claim to hold the truth, but can, on the other hand, object that such and such a proposition is false: a scientific theory is then true until the demonstration of its errors is established, or to use Kuhn’s terms [1970], until a paradigm change intervenes.

#### 2.5. *Conclusion on previous paradigm shifts and transition*

Today scientific knowledge is still mainly based on the Western paradigm of the abstract universal;

<sup>4</sup>Our translation: “the living can only be explained by involving a final (descending) causality, an intention... because it is inconceivable that time, however extensive we may imagine it to have been, could have produced the living forms we know.”

<sup>5</sup><https://www.webofscience.com/wos/woscc/basic-search>.

however, scientific knowledge has undergone profound changes proving its capacity to develop new frameworks. Not least of these changes is that science is one mode of existence among others—with its own forms of veridiction—and that this mode of existence does not have to be hegemonic [Latour, 2012a, p. 79–104; Latour, 2022]. The environmental history of political ideas shows that the will to modernize has been expressed in the form of a double injunction since the 17th century: one oriented towards abundance, the other towards freedom [Charbonnier, 2020, p. 41]. The challenges faced by science today require a fine understanding of the complexity of the situation, as this article aims to demonstrate.

In the following Section 3, we highlight a few milestones leading to a new paradigm, capable of making science a lever for action in the context of the Anthropocene. The current period marks an essential breaking point for humanity [Hamilton, 2016].

### 3. Towards a new scientific paradigm: knowledge diversity

In the present section, we deliberately examine the scientific discourses and emerging concepts likely to fuel a profound and far-reaching change in scientific thinking and practice. If, as Bouleau [2017] attests, 90% of science is still positivist, the present exploration deals with the remaining 10% that opens the horizon. The evolution of the positioning of science in the context of the Anthropocene and the conceptual and paradigmatic developments that are taking place, open up a necessary dialogue with the other terrestrials, to use Latour's word [2021]. Whatever the scale envisaged, it will be a matter of questioning science in the face of otherness.

#### 3.1. *Geographicity and recognition of the other*

We have mentioned Kuhn's framework of thought about paradigm shifts in science earlier; we now focus on the case of "Man and the Earth" [Dardel, 1990, first ed. 1952]. As early as the 1950s, Dardel wrote that the superiority of modern man over the surrounding world seems an insurmountable obstacle to a sincere harmony with the forest, the sea, or the mountains [quoted by Raffestin, 1987].

Unnoticed at the time of its publication, this statement was later questioned in the 1970s. It now deserves to be considered in the corpus of resources

facing the Anthropocene. Why did Dardel's work go unnoticed? It emerged in the post-war context when geography was then largely oriented towards the neo-positivist theses developed by the Vienna Circle, who at the time exercised a strong influence on the social sciences [Hempel, 1942]. This movement, born after the First World War, continued in France, Dardel's country, with the development of the "New Geography" after the Second World War. The aim was then to contribute to the reconstruction of the country. Legitimizing the scientific character of geography involved nomothetic approaches that made it possible to establish universal laws. Raffestin [1987] explains that Dardel's tragedy is to have been one paradigm ahead of his contemporaries: « Formé au paradigme du « voir », il a écrit au moment où triomphait celui de l' « organiser » alors qu'il postulait celui de l' « exister ».<sup>6</sup>

This paradigm proposed by Dardel [1990, first ed. 1952] is what he calls geographicity, which is defined as: « connaître l'inconnu, atteindre l'inaccessible, l'inquiétude géographique précède et porte la science objective. (...) une relation concrète se noue entre l'homme et la terre, une géographicité de l'homme comme mode de son existence et de son destin »<sup>7</sup> [quoted by Raffestin, 1989]. Dardel [1955], influenced by the work of the anthropologist Leenhardt [1985, first ed. 1947] in New Caledonia, and by Heidegger's philosophy, put forward the idea that the human being is inseparable from his space; meaning there is "an existential relationship" in which everything that surrounds the human being participates in his structure and his substance. These ideas obviously distanced Dardel from the positivist postures of the science of his time. Years later, his proposition to rethink the relationship between humans and the world by questioning this link was taken up by human and cultural geography [Berque, 1994, Bonnemaïson, 1992, Collignon, 2002, Frémont, 1972,

<sup>6</sup>Our translation: "Trained in the paradigm of 'seeing', he wrote at a time when the paradigm of 'organizing' was triumphing, whereas he postulated the paradigm of 'existing'."

<sup>7</sup>Our translation: "To know the unknown, to reach the inaccessible, geographical concern precedes and supports objective science. (...) a concrete relationship is established between man and the earth, a geographicity of man as a mode of his existence and his destiny."



Lussault, 2007]. The critique and analytical contributions that followed made it possible to understand that the positivist foundations of science were not sufficient to grasp reality in all its complexity.

Today the question of the relationship between humans and the Earth—and the need to escape from an all-technological world, as analysed by Hoquet [2021]—are central to reflections on the future of societies, and on the place of science in this articulation between the particular and the general, the local and the global, the specific and the systemic.

### 3.2. *Complexity science, a revolution since the 1980s*

For several decades, the idea has been developed that natural or social systems are for the most part complex systems and can be studied as such. The reduction to simple linear causalities, practiced for centuries, is no longer relevant. A complex system is a system composed of many differentiated elements interacting with each other in a non-trivial way (non-linear interactions, feedback loops, etc.). It is characterized by the emergence at the global level of new properties, unobservable at the level of the constituent elements and by global operating dynamics difficult to predict from the observation and analysis of elementary interactions [Guespin-Michel, 2016].

This evolution emerged in the last quarter of the 20th century and opened avenues for scientific, social, philological, and philosophical ventures to deal with the intertwined crises of climate change, ecosystems, and human societies.

While Morin [1982] introduced the idea of complexity into the human sciences independently from mathematics, according to Guespin-Michel [2016], the complexity revolution was made possible by advances in computer science. Guespin-Michel [2016, p. 12] explained that the non-linearity of systems was initially ignored, not only because mathematics lacked the means to address it, but also because Cartesian thinking was an obstacle to complex thinking. The author concluded that a new rationality was emerging, based on a dialectical thinking of complexity, capable of studying natural systems and of fighting against the danger of irrationalism.

The complexity revolution may also be observed in the field of physics. In the introduction to his colleagues' interventions, the academic Derrida [2008]

stated that, since its origins at the end of the 19th century, statistical physics has attempted to explain the collective behavior of many elementary objects based on their interactions. The aim was to predict whether a body was a gas, a liquid, or a solid. It gradually became apparent that increasingly complex phenomena could result from the collective behavior of many interacting objects, both in physics (fractures, avalanches, etc.) and in other fields (neural networks, gene and protein networks, the evolution of species, sociological networks, road traffic, financial markets).

However, many works that adopt the complex method encounter strong opposition. This has been the case with the Gaia theory “that views the evolution of the biota and of their material environment as a single, tightly coupled process, with the self-regulation of climate and chemistry as an emergent property” [Lovelock, 1989]. The philosopher and sociologist of science Latour [2015] recalls how this Gaia theory, that forces one to accept, or at least to explore the limits of the Earth, has found many detractors.

Today scientific discussions tend to abate, with articles such as Lenton and Wilkinson's [2003] and its discussion of all the specific terms of the controversy and its conclusion on how Gaia has contributed to the emerging field of “Earth System Science.” Žukauskaitė [2020] synthesizes Haraway's [2016], Latour's [2015] and Stengers's [2015] work and invites us to rethink Gaia, not as an autopoietic unit, but as a complex and dynamic system of living things, including humans. Lenton and Latour [2018] present a convergence between the Gaia hypothesis, Earth System Science and humans' reflexivity as called for by Beck [1986]. They offer to create an infrastructure of sensors that helps track environmental changes along with social responses.

As the world is essentially complex, the concepts of complex real systems should be part of the arsenal of every and any scientist, but also of the thinking of every citizen [Guespin-Michel, 2016, Lévy and Lussault, 2003]. The very organization of scientific institutions based on disciplines must be revisited. The objective of reaching reality through the universal laws of a discipline is replaced by the objective of apprehending possible risks that can only be approached through the coupling of ecosystem dynamics and socio-cultural dynamics (i.e., all

disciplines taken together). This advance towards complex thinking is therefore in the direction of a change in the posture of science and of its aims.

### 3.3. *Breaking down the silos of thinking*

This part explores the emergence of breaking down siloed thinking as attested in diverse studies: figuration [Descola, 2021], the concept of the “in-between” [Jullien, 2016, Cassin, 2016], origin and legal narratives [Nouvel, 2020, Bourgeois-Gironde, 2020, Notre affaire à tous, 2022].

The masterly synthesis of contemporary anthropological knowledge “Beyond Nature and Culture” [Descola, 2005 first ed., English ed. 2013] establishes four forms of worlding: i.e., four ways of composing worlds from the salient elements that the members of a collective detect or actualize in their surroundings. The key lies in the elementary mechanism of identification of the other. For any being that falls under his perception (fellow human being, animal, plant, thing, technical object, or other), every individual performs a process of “identification” that leads him to two questions. Does the other have the same interiority as me (or us)? Does it have the same physicality as me (or us)? The two possible answers (yes or no) to these two questions (interiority/physicality) lead to four possible combinations. Descola [2005] then establishes four categories of worlding: animism (yes/no), analogism (no/no), naturalism (no/yes) and totemism (yes/yes).

This synthesis not only presents a plurality of forms of worlding, but also shows that all four ways of composing the world are respectable. Science has so far been based on only one of these forms and is therefore invited to construct its approaches differently in the face of this plurality. Until recently, science, i.e., Western science, led to the view that non-naturalist populations were unenlightened and lacked the necessary basis for understanding the world in which they lived. In the rush to colonise, it was thought that these populations had to be civilised by teaching them the supposedly universal naturalist bases of knowledge.

Today, the challenge is to try to understand the multiple ways in which humans describe the world and what they do in it [Descola, 2011]. In response to Lacroix [2021], Descola states that it is interesting that the siloed nature of ontologies is now breaking down. The hybridity we experience today is only un-

derstandable if we can redesign the constituent parts that it combines. To do so, Descola’s recent book [2021] led an investigation to identify the four ontologies of worlding in figurations, i.e., the objects they represent, and the relationships they depict. Indeed, figuration is, on the one hand, an operation common to all humans [Descola, 2021, p. 29] and, on the other hand, a doubly significant representation: as an icon and as an index of intentionality [Descola, 2021, p. 30].

A striking result of this investigation is the capacity of images to prefigure ontological and cosmological changes that are subsequently evident in texts that appear much later [Descola, 2021, p. 18]. In that respect, Europe offers a remarkable example. The dominant ontology in Europe had been analogism since Antiquity [Descola, 2021, p. 57] until naturalism took shape with the Renaissance, as attested in the writings of scholars and philosophers of the 17th century. However, naturalist ontology already appeared in painting at the beginning of the 15th century.

What about the current period? Do figurations of the 20th century augur new ways of seeing the world? Although we still lack the necessary hindsight, tangible leads are enlightening. With Cubism, at the beginning of the 20th century, entire sections of European images began to break free from the iconographic canons of naturalism, heralding the probable end of that cycle [Descola, 2021, p. 562]. Many contemporary artists evidence a great deal of ontological eclecticism in their work and their figurations fall under a plurality of worlding. Today’s renewed success of Arcimboldo (1526–1593), qualified as an analogist bubble [Descola, 2021, p. 563], also illustrates our societies’ appetite for a renewed way of seeing people and things [Parisi and Horvath, 2021]. Figurations are thus auguring the breaking down of siloed worlding.

The richness of the “in-between” emerges from sciences as diverse as biology, anthropology, Earth system science, philosophy, or philology. Jullien [2016] points out that in a world becoming globalized, there is no longer a “beyond” to dream of. It is in the “in-between” that resources are being discovered. The present resource is not that of identification, but of exploration, allowing another possible world to emerge. It will be necessary, he specifies, to leave the thought of being (ontology) and begin to think of the “in-between”. Coviability, the concept of sustainable life for human societies within ecological systems, in-

tegrating interdependence and respect for others, fits well into this trend: it is in the “in-between” that the sense of self, respect for the other and interaction or interdependence develops [Coudrain, 2019].

The philologist Cassin [2016] praises translation for creating the passage between languages. As a competence in dealing with differences, translation can constitute the new paradigm of the humanities, each language being a web of equivocations. A single sentence, with its syntax and semantics, is indeed rich in perception, direction, and meaning.

Breaking down siloed thinking is also emerging from the analysis of discourses of the origin. Nouvel [2020] shows that there are four types of discourses: (1) descending from the complex to the simple: mythological discourses; (2) ascending from the simple to the complex: scientific discourses inaugurated by Darwin [1859]; (3) ascending and descending: rational discourses; and (4) neither ascending nor descending: phenomenological discourses. Nouvel [2020] links these four types of discourses to the four forms of worlding Descola [2005]. He emphasises the circularity of these discourses, as Descola [2005, 2021] did for the four types of worlding: there is no continuous evolution from one to the other, nor any superiority of one over the other.

The breaking down of the silos of worlding opens the way to thinking about the gap between the different ways of posing and solving scientific questions. Researchers can try to disrupt the scientific vision of the way we inhabit the Earth through three processes that are not totally utopian since they have already existed: how humans adapt to their environments, how they appropriate them, and how they express them politically [Descola, 2015]. Adaptation involves propagating the idea that our destiny is entirely dependent on billions of actions and feedbacks through which we generate the environmental conditions that allow us to inhabit the Earth; this theme is also widely developed by Latour [2021] and by Stengers [2019]. Appropriation means stressing that it is rather the ecosystems that are the bearers of rights and not solely human beings. Political representation is that of ecosystems.

In the wake of the “Natural contract” [Serres, 2020, first ed. in French 1990; English ed. 1995] which focused on the recognition of nature as a subject of law, there has been a proliferation of initiatives in recent years throughout the world, aimed at establish-

ing legal rights of the environment [Notre affaire à tous, 2022, p. 141–414]. In the anthropocentric Western culture, characterized by a naturalist logic, this is a new positioning compared, for example, to the Universal Declaration [UN General Assembly, 1948], characterized by the absence of the word “nature” and which concerns only humans, in which non-humans do not benefit from a legal personality, nor from rights of their own. These changes are linked to the environmental crisis and to the position of civil society in favor of environmental protection [Notre affaire à tous, 2022, p. 7, p. 13–14]. Recent work on plant intelligence [Bouteau *et al.*, 2021] provides evidence of the ability of non-humans to adapt and solve problems. The controversy surrounding the work on plants is reminiscent of decades of debate in the Earth sciences around the Gaia hypothesis [Love-lock, 1989]. Indigenous people who see humans as part of Nature are also the ones who protect it the most and lead the way towards its recognition as a subject of law [Notre affaire à tous, 2022, p. 8 and p. 105]. For example, Maori claims in 2017 led to the recognition of the Whanganui River as a subject of law. This case attests to the consideration of customary cultural values and a worlding different from Western universalism [Bourgeois-Gironde, 2020] in the establishment of law. There is something fundamental in this movement: the capacity to exercise a right for a human being in a place is no longer linked to his or her person. It is linked to his dependence on a place, that dependence being the legal source of the legitimacy of the occupation of space. This effects a reversal of the theory of appropriation.

Positions are emerging for science with the awareness that dialogue, understood as an attention to the “in-between”, is necessary both between existing people and different ontologies. This can also be seen in traditional ecological knowledge (TEK). The rise of citizen science and the place it takes in the field of research is a matter of concern. In the following section, we develop the idea that it corresponds to a form of adaptation of science to the context of the Anthropocene.

#### **4. Strengthening participative science in the context of the Anthropocene**

The paradigm shift in science is also a consequence of the call from citizens to find in science, usable

matter, questions, and answers to their own preoccupations. In this way, it is remarkable that IPCC reports changed from theoretical climate science to actionable science, especially in the last two assessment reports [IPCC, 2014, 2022b]. It is now more a question of delineating the gap between different ways of approaching one's existence and one's relationship to others (humans, groups of humans or non-humans) than of seeking a universal truth which in any case is not universal.

Scientific knowledge on the state of the environment has an impact only if society is informed or concerned by this scientific knowledge and puts in place decision-making mechanisms [de Marsily and Lallier-Vergès, 2015]. This calls to mind the deeply politicized nature of the environmental issue at a time when the need to reform our lifestyles and thinking patterns must take place without further delay. Humanity, a victim of itself, must face these changes while realizing that the most socially and economically vulnerable are the first victims of these changes. If the notion of a risk society, born in the 1980s [Beck, 1986, Perretti-Watel, 2010], called for risk to be considered, this notion is now outdated.

Uncertainty [Reghezza, 2015] and threat are now taking hold [Bourg *et al.*, 2013]. They are part of everyone's daily call for collective action and the scientific community has a role to play. This section presents the specificity of citizen science, its growth, and examines the opportunity of this activity to face current challenges.

#### 4.1. *The specificity of citizen science*

Dias da Silva *et al.* publication, "Review on citizen science in ecology and the environment" [2017] is a foundational document that clarifies the variations of terms used to designate these sciences, or the diversity of forms that they can take. In this paper, we focus more specifically on the usefulness of citizen science and the very concrete reasons that lead us to see them as a form of adaptation to the context of the Anthropocene.

Citizen science projects combine the knowledge of scientists (expert knowledge) with the knowledge of non-scientist professionals (local or lay knowledge) [D'Arripe and Routier, 2013, Wynne, 1999]. Non-scientist professionals are citizens who may be involved in one or more stages of the scientific

research process: the definition of the research project, the development of the methods used to answer the questions posed by the project, the collection of data and/or their processing, the publication of the results, and finally the dissemination of results [Godrie and Heck, 2021]. The co-production of knowledge between citizens and scientists begins by a negotiation and results in a different way of doing science [Billaud *et al.*, 2017].

Participative research projects need time and appropriate means and methods [Carrel, 2006]. They can be hindered by certain obstacles, such as the abandonment by, or the disinterest of citizens [Conrad and Hilchey, 2011]. To overcome these problems, participatory research projects should respect certain operating rules: consensus on objectives, clear definition of everyone's participation and funding. The specific challenge lies in refraining from depreciating the knowledge produced by the different participants, so that trust and dialogue allow the most relevant questions to be raised by the group as a whole. Researchers are, for their part, responsible for ensuring that scientific rigor is respected so that the results can be operative.

Citizen science covers a very wide variety of situations. In France, the greatest number of publications resulting from participatory research are in the fields of health and the environment [Storup, 2013]. Research projects dealing with concrete problems affecting civil society, with results likely to improve living conditions, trigger or accelerate citizen participation. In these cases, the knowledge of local citizens confronted with these problems can provide a complementary dimension with a better understanding of issues [Tengö *et al.*, 2014].

In some cases, citizen science is essential as for example in New Caledonia (Loyalty Islands Province, French Republic) that has adopted an Environmental Code of the Loyalty Islands Province [CEPIL, 2019]. This code is based on the specific conception of life of the Kanak populations which stipulates that humans belong to the natural environment just like non-humans. As a result, certain elements of nature can be endowed with a legal personality with rights and can be represented by Kanak clans [Notre affaire à tous, 2022, p. 341–342]. Thus, lands under customary law are legally recognized. Research projects concerning these lands are enriched by a dimension of "negotiated knowledge" which should be considered

in terms of otherness or intercultural dialogue. With the management of coastal areas, the classic framework for public action is inoperative since customary law supersedes state intervention on matters related to land. The involvement of local populations in management decision-making goes beyond any simple institutional “good practice”; it is a necessary precondition for action. In this context, work has been carried out on coastal erosion monitoring with the involvement of local populations. This is part of a science-participatory approach [Le Duff *et al.*, 2020]: a more global reflection on coastal risk management, within the context of global warming and the proximity of the Pacific Ring of Fire, implying risks of tsunamis [Le Duff *et al.*, 2016, Le Duff, 2018]. The population was waiting for answers and the exchanges that fuelled this work made it possible to develop a dialogue, both with the custodians and with the population. Customary authorities could then move forward on the shared understanding that discussions were ongoing about the future of the tribes established as close as possible to the coast. This joint reflection led to defining a management strategy shared by the customary, municipal, and provincial authorities and to the commitment of financial support for various development operations on the territory.

#### 4.2. *Recent growth of citizen science*

To take into account the knowledge of the inhabitants of a particular location is nothing new. Empirical sciences from the 16th to the 19th centuries (botany, entomology, zoology, or astronomy) incorporated this knowledge into their corpora and datasets [Houllier *et al.*, 2017]. At that time, they did not necessarily take into consideration the framework in which this knowledge was produced and mobilized, nor the full extent of its functions in the societies in question. More recently, social sciences have made individuals and societies their object of study. They have developed methods and protocols to encourage interaction and exchange with populations in order to apprehend this knowledge and to better understand its position, its representation, its use, and its practice in the social sphere. It meant trying to understand the meaning of this knowledge. But this knowledge is not, as we understand it today, “citizen science”, although these projects involved the participation of

people. Citizen science is characterized by the reflexive positioning of the participants, and by a mutual and objective will to build a dialogue.

The strong growth of scientific publications on citizen science since the 2000s testifies to an increasing interest in the scientific community for this type of approach [Houllier *et al.*, 2017, Juan, 2021, Storup, 2013] and highlights its usefulness in the current context. This growing interest could be read as converging with a concomitant societal desire to expand frameworks for democratic expression. Indeed, the 1980s were marked by the development of decentralization, social dialogue bodies, and administrative procedures for considering public opinion, particularly on environmental issues. This movement led in the early 2000s to the idea of participatory democracy that left more room for deliberation [Chlous *et al.*, 2017]. During that same period, the discourse of international bodies on a community-based management of natural resources and, more generally, of the environment developed. The rise of citizen science was thus part of a context that helped shape it.

Thanks to the mobilization of the participatory approach, the OdyséYeu project on risks affecting coastal areas, such as coastal erosion and marine submersion, has made possible the acquisition of data over larger areas at higher frequencies, and with a greater reactivity than what conventional scientific production could have achieved [Cariou *et al.*, 2021]. The usefulness of the approach goes even further because it also values the vernacular knowledge of the participants on local issues. What takes place between the participants is dialogue, listening, consideration, and an understanding of the respective positions of the participants. The COSACO project, which also deals with coastal risk [Ruz *et al.*, 2021] emphasizes this valuing of vernacular knowledge.

On a national level, debates led by French representatives of different parts of society resulted in guidelines issued by the Economic, Social and Environmental Council [Blanchet and Jouzel, 2017]. At an international level, IPCC’s sixth report [2022a, 2022b] clearly recognizes the value of various forms of knowledge, such as scientific knowledge, but also indigenous knowledge and local knowledge, in understanding and assessing climate adaptation processes and actions to reduce the risks of human-induced climate change.

### 4.3. *Citizen science and the challenges of the Anthropocene*

Lenton and Latour [2018] clarify this challenge by putting forward an “infrastructure of sensors” to track the environmental changes and the latency of societal responses to these changes. Citizens, activists, and politicians would then collaborate with scientists, the goal being to assess where things are going wrong.

Climate change best illustrates how much the actions of everyone on Earth have an impact on the life of everyone else on Earth. Masson-Delmotte’s review article [2020] provides a succinct scholarly summary on the current state of knowledge on climate change. The Club of Rome warned the world of the looming catastrophe in 1972 by a regularly updated publication [Meadows *et al.*, 2004], and in 2017 more than 15,000 scientists confirmed the seriousness of the situation [Ripple *et al.*, 2017].

The recent publications of IPBES and IPCC reaffirm that risks are still omnipresent in the lives of all humans and must be taken on board for the future habitability of the Earth. IPCC [2022a] sets the tone in its opening pages: “This report has a strong focus on the interactions among the coupled systems climate, ecosystems (including their biodiversity) and human society. These interactions are the basis of emerging risks from climate change, ecosystem degradation and biodiversity loss and, at the same time, offer opportunities for the future.” IPBES [2019] delivers a similar message: Nature and its vital contributions to people are deteriorating worldwide.

In the tangible field of water, de Marsily [2009] explained the inescapable risk of famine with a population expected to reach 10 billion by 2100. Deficient food production in several regions highlights the need to share a common vision for the future and to plan for huge transfers of virtual water (agricultural production) to avoid inevitable famines.

The question is to move away from the current system, which is recognized as unsustainable in the long term, to another system which has yet to be invented, through sustainable transitions in sectors such as energy, food, and transport [Markard *et al.*, 2020].

This challenge requires a global mobilisation, and to do so demands a detailed understanding of the complexity of the problem. Bouleau [2017], like

Serrao-Neumann and Coudrain [2017], calls for a re-orientation of science in its aims and methods. He proposes to introduce these fears into scientific work. This position of concern highlights the necessity of gathering knowledge from all available interpretative sources. The method then becomes participatory, between scientists from various disciplines and between scientists, citizens, and decision-makers.

Both citizens and researchers, as they share their different knowledge with each other, gain competence [Houllier *et al.*, 2017]. For projects with data collection, citizen engagement allows covering a greater diversity of situations much faster (time of day, survey periodicity, seasons), which a research project cannot always achieve [Houllier *et al.*, 2017]. The diversity of stakeholders brings added value through the multiplication of points of view [Brun, 2017]. Sauermann *et al.* [2020] argue that citizen science has the potential to help solve sustainability problems. This combination of diversified knowledge makes it possible to ask the most relevant questions, sometimes bringing to light new decisive elements that will be of great importance during the project [Juan, 2021].

Moreover, the knowledge of the citizens involved in participatory research who have been attached to a territory for many years can help in the interpretation of the results of the research project [Brun, 2017]. Beyond these appreciable advantages, motivated and enthusiastic citizens will be the bearers of the results acquired in collaboration with the scientists. The application of the results will be easier as these committed citizens become mediators. They are able to convince the general public and to participate in the urgent implementation of the solutions. Obviously, this participatory research does not exclude other types of scientific research [Godrie and Heck, 2021], but feeds off them, and vice versa.

## 5. Conclusions

The authors of this article have approached this work on the issue of paradigm shift in science and geoscience in the context of the Anthropocene as non-specialists. Bringing together the skills as biologist, geographer, and hydrologist, have placed ourselves in the “in-between” [Jullien, 2016] to explore the gap in our thinking habits in the face of an unprecedented upheaval: that of humans who have

become a geological force modifying the Earth System through their actions.

The scientific syntheses on the climate or the biosphere by large organizations (IPCC, IPBES) insist on the urgency and the need to develop an intelligent and conscious society, capable of reacting. This is complemented by the contemporary work of human and social scientists, such as Descola, Haraway, Latour, Morin, Oreskes, Serres and Stengers. It is true that there is an increasing number of scientific programs and publications on the upheaval of the Earth's envelopes, but communication is clearly not enough to bring about necessary changes. The question becomes that of appropriating new worldviews.

The authors of this paper have examined the scientific paradigm established since the 17th century. It led to the establishment of universal laws of nature inherited from the classical Greeks. It also encouraged a sense of hierarchy with the hegemony of science over other modes of existence, such as art, law, politics, or religion. It is this form of power, based on a limited vision of a universal, that is questioned.

In today's scientific communities, the universal is gradually being replaced by the idea of a complex universe. The scientific community should meet the challenges posed by the Anthropocene. This situation calls for a collective responsibility for the current state of the planet. It demands a change in our vision, our practices, our relationship to the world and therefore our way of scientific work. History shows us that our scientific community has not only been able to evolve over time but has also been an actor in these changes. Different elements interacting with each other in a meaningful manner compose the world. We are witnessing conditions favorable to the emergence of a new paradigm and a gradual withdrawal of "normal science" to use Kuhn's words [1970].

Several developments attest to the emergence of a global, albeit still small scale, change in the scientific paradigm within scientific communities. The concept of paradigm shift developed by Kuhn [1970], has achieved great success beyond the scientific sphere. It facilitates the intrinsic transformation of research practices. It liberates from value judgement and morality. The complexity revolution is another thinking tool: a system composed of many different elements which interact with one another, in a non-linear way and feedback loops. This thinking tool is spreading to many scientific fields, from the human

sciences to physics.

The anthropologist Descola [2005, 2021] provides a basis for understanding that all humans conceive a world based on what they perceive as other. This author established four categories of worlding. Modern science developed in the 17th century based on only one of these worlding: naturalism, according to which what is non-human is devoid of interiority. His approach makes us understand how and to what extent it has become "normal" for humans to look down on objects and non-human living things. Descola's most recent work on images [2021] shows that the worlding limited to naturalism may now be coming to an end in favor of a plurality of ontologies. We consequently focused on exploring the other ontologies of worlding without value judgements and an openness to a multiplicity of ways of seeing the world.

Changing the frame of reference is no easy task, given that the time window is only a decade long and that what is left to be invented is considerable [Chakrabarty, 2018, IPCC, 2022a]. Citizen science could be essential, allowing many citizens to see science in action, with its complexity, its temporality, its challenges, its limits, but also its strength. It seems to pave the way for blending the efforts of the scientific community with the rest of society.

To the question we initially posed, the answer is that the sudden growth of citizen science in recent years responds to the need for a shared reality. Global warming and the so-called "crises" of the Anthropocene are still too often seen as catastrophic when in fact there is no turning back. There will be no post-disaster or "relief phase", no "return to normal". However, to initiate the necessary change, to initiate collective action, we need a narrative to drive us to a viable future.

## Conflicts of interest

Authors have no conflict of interest to declare.

## Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## Acknowledgements

The authors are very grateful to the four anonymous referees whose valuable comments helped them in the revision of this article. They also thank Professor Claudine Raynaud for her intense work on the revision of the English language for greater fluency.

## References

- Beck, U. (1986). *Risikogesellschaft. Auf dem Weg in eine andere Moderne*. Suhrkamp, Germany. ISBN:978-3-518-11365-3. French ed., 2001. La société du risque : sur la voie d'une autre modernité. Aubier, France, ISBN:9782700736793.
- Berque, A. (1994). *Cinq propositions pour une théorie du paysage*. Champ Vallon, Suisse. ISBN:9782876731912.
- Billaud, J.-P., Hubert, B., and Vivien, F.-D. (2017). Les recherches participatives : plus de science ou une autre science ? *Nat. Sci. Soc.*, 25(4), 325–326. ISSN 1240-1307.
- Blanchet, J. and Jouzel, J. (2017). *Sciences et société : Répondre ensemble aux enjeux climatiques*, volume 19. Avis du Conseil économique et social, CESE, France. Accessed 7 March 2022, [https://www.lecese.fr/sites/default/files/pdf/Avis/2017/2017\\_19\\_contribution\\_dialogue\\_sciences\\_societe.pdf](https://www.lecese.fr/sites/default/files/pdf/Avis/2017/2017_19_contribution_dialogue_sciences_societe.pdf).
- Bonnemaison, J. (1992). Le territoire enchanté. Croyances et territorialités en Mélanésie. *Géogr. Cult.*, 3, 71–88.
- Bouleau, N. (2017). *Penser l'éventuel. Faire entrer les craintes dans le travail scientifique*. Quae, France. ISBN:978-2-759-22544-6.
- Bourdieu, P. (1976). Le champ scientifique. *Actes Rech. Sci. Soc.*, 2(2), 88–104.
- Bourdieu, P. (1997). *Les usages sociaux de la science : pour une sociologie clinique du champ scientifique*. Sciences en question. INRA, France. ISBN:9782738007933.
- Bourg, D., Joly, P.-B., and Kaufmann, A. (2013). *Du risque à la menace. Penser la catastrophe*. PUF, France. ISBN:978-2-13-060631-4.
- Bourgeois-Gironde, S. (2020). *Être la rivière*. PUF, France. ISBN:978-2-13-082094-9.
- Bouteau, F., Gresillon, E., Chartier, D., Arbelet-Bonnin, D., Kawano, T., Baluška, F., Mancuso, S., Calvo, P., and Laurenti, P. (2021). Our sisters the plants? notes from phylogenetics and botany on plant kinship blindness. *Plant Signal. Behav.*, 16(12), article no. 2004769.
- Brun, P. (2017). Le croisement des savoirs dans les recherches participatives, Questions épistémologiques. *Vie sociale*, 20, 45–52.
- Cariou, E., Baltzer, A., Leparoux, D., and Lacombe, V. (2021). Collaborative 3D monitoring for coastal survey: conclusive tests and first feedbacks using the SELPhCoAST workflow. *Geosciences*, 11(3), article no. 114.
- Carrel, M. (2006). Politisation et publicisation : les effets fragiles de la délibération en milieu populaire. *Politix*, 19(75), 33–51.
- Cassin, B. (2016). *Éloge de la traduction, compliquer l'universel*. Librairie Arthème Fayard, France. ISBN:978-2-213-70077-9.
- CEPIL (2019). *Code de l'environnement de la province des îles Loyauté*. Province des îles Loyauté, 1ère édition. Accessed 12 August 2022, <http://www.province-iles.nc/sites/default/files/2019-03/19-033-Code-de-l-environnement-PIL.pdf>.
- Chakrabarty, D. (2018). Whispers troubling the human heart / Troublants murmures à l'oreille des hommes. In Simenel, R. and Sabrié, M.-L., editors, *L'écologie des mondes / Ecology of our worlds*, pages 18–21. IRD, France. ISBN:978-2-7099-2883-0.
- Chakrabarty, D. and Chalier, J. (2018). Changement climatique et capitalisme. *Esprit*, 1–2, 153–168.
- Charbonnier, P. (2020). *Abondance et liberté. Une histoire environnementale des idées politiques*. La Découverte, France. ISBN:978-2-348-04678-0.
- Chlous, F., Dozières, A., Guillaud, D., and Legrand, M. (2017). Introduction. Foisonnement participatif : des questionnements communs ? *Nat. Sci. Soc.*, 25(4), 327–335.
- Clerc, P., Deprest, F., Labinal, G., and Mendibil, D. (2019). *Épistémologie et histoire des savoirs sur l'espace*. Collection U Géographie. Armand Colin, France, 2nd edition. ISBN:978-2-200-62478-1.
- Collignon, B. (2002). Les toponymes inuit, mémoire du territoire : étude de l'Histoire des Inuinnait. *Anthrop. Soc.*, 26(2–3), 45–69.
- Conrad, C. C. and Hilchey, K. G. (2011). A review of citizen science and community-based environmental monitoring: issues and opportunities. *Environ. Monit. Assess.*, 176, 273–291.
- Coudrain, A. (2019). Climate change, a catalyst for a new utopia towards coviability. In Barrière,



- O., Behnassi, M., David, G., Douzal, V., Fargette, M., Libourel, T., Loireau, M., Pascal, L., Prost, C., Voynier, Ravena-Cañete, V., Seyler, F., and Morand, S., editors, *Coviability of Social and Ecological Systems: Reconnecting Mankind to the Biosphere in an Era of Global Change*, volume I of *The Foundation of a New Paradigm*, pages 533–552. Springer, Switzerland. ISBN:978-3-319-78496-0.
- Crutzen, P. J. and Stoermer, E. F. (2000). The Anthropocene. *Glob. Chang. Newsl.*, 41, 17–18.
- Dardel, E. (1955). L'Homme dans l'univers mythique. D'après l'œuvre de Maurice Leenhardt. *Rev. d'histoire Philos. religieuses*, 35(2), 159–173.
- Dardel, E. (1990). *L'Homme et la terre : Nature de la réalité géographique*. Editions du CHTS, France. ISBN:2-7355-0200-7 (first ed. 1952).
- D'Arripe, A. and Routier, C. (2013). Au-delà de l'opposition entre savoir profane et savoir expert : une triangulation des méthodes. Recherches qualitatives. *Hors série « Les Actes »*, 15, 221–233. Du singulier à l'universel, Canada. ISSN 1715-8702. <http://www.recherche-qualitative.qc.ca/Revue.html>.
- Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection, or The Preservation of Favoured Races in the Struggle for Life*. John Murray, United Kingdom.
- Daston, L. and Galison, P. (2012). *Objectivité*. Les presses du réel, France. ISBN:978-2-84066-334-8. 2007 first ed. Objectivity, Zone books, Princeton Univ Press, USA. ISBN: 9781890951795.
- de Marsily, G. (1982). Introduction. In de Marsily, G. and Merriam, D. F., editors, *Predictive Geology with Emphasis on Nuclear-waste Disposal*, pages xi–xv. Pergamon Press, England. ISBN:9780080262468.
- de Marsily, G. (2009). *L'eau, un trésor en partage*. Dunod, France. ISBN:978-2-10-051687-2.
- de Marsily, G. and Chabaux, F. (2020). Des géosciences aux sciences de la Planète / From geoscience to sciences of the planet. *C. R. Géosci.*, 352, 1–6.
- de Marsily, G. and Lallier-Vergès, E. (2015). Les recherches sur l'eau : de nombreux défis pour l'avenir. In *L'eau à découvert*, pages 320–321. CNRS, France. ISBN:9782271119117.
- Derrida, B. (2008). Introduction. In *Conférence débat et controverses « Physique des systèmes complexes »*. Académie des sciences, Ecole normale supérieure. 1er juillet, Accessed 26 February 2022, [https://www.academie-sciences.fr/archivage\\_site/activite/conf/debat\\_010708.pdf](https://www.academie-sciences.fr/archivage_site/activite/conf/debat_010708.pdf).
- Descola, P. (2005). *Par-delà nature et culture*. Galilimard, France. ISBN:9782070772636. English ed., (2013). Beyond Nature and Culture. Univ. of Chicago Press, USA. ISBN:978-0-226-14445-0.
- Descola, P. (2011). Cognition, perception et mondialisation. *Cah. Philos.*, 4(127), 97–104.
- Descola, P. (2015). Humain, trop humain. *Esprit*, 420, 8–22.
- Descola, P. (2021). *Les Formes du visible*. Collection « Les livres du nouveau monde ». Seuil, France. ISBN:978-2-02-147698-9.
- Dias da Silva, P., Heaton, L., and Millerand, F. (2017). Une revue de littérature sur la « science citoyenne » : la production de connaissances naturalistes à l'ère numérique. *Nat. Sci. Soc.*, 25, 370–380.
- Frémont, A. (1972). La région : essai sur l'espace vécu. In *Mélanges offerts au professeur A. Meynier, La pensée géographique contemporaine*, pages 633–678. Presses univ Bretagne, France.
- Gemenne, F., Rankovic, A., and Atelier de cartographie de Sciences Po (2019). *Atlas de l'Anthropocène*. Presses de Sciences Po, France. ISBN:978-2-7246-2415-1.
- Godrie, B. and Heck, I. (2021). L'approche participative, la recherche-action et leurs principales stratégies d'enquête et d'inclusion des groupes subalternisés. In Piron, F. and Arsenault, E., editors, *Guide décolonisé et pluriversel de formation à la recherche en sciences sociales et humaines*. Science et bien commun, Canada, <https://scienceetbiencommun.pressbooks.pub/projetthese/chapter/methodes-de-recherche-participative-recherche-action-et-sciences-citoyennes/>.
- Grau, D., Grau, N., Gascuel, Q., Paroissin, C., Stratonovitch, C., Lairon, D., Devault, D. A., and Cristofaro Di, J. (2022). Quantifiable urine glyphosate levels detected in 99% of the French population, with higher values in men, in younger people, and in farmers. *Environ. Sci. Pollut. Res.*, 29, 32882–32893.
- Guespin-Michel, J. (2016). La révolution du complexe, Sciences, dialectique et rationalité. Accessed 26 February 2022, <http://www.revolutionducomplexe.fr/images/downloads/revolutionducomplexe-guespin.pdf>.
- Hamilton, C. (2016). The Anthropocene as rupture.

- Anthr. Rev.*, 3(2), 93–106.
- Haraway, D. J. (2015). Anthropocene, capitalocene, plantationocene, chthulucene: Making kin. *Environ. Humanit.*, 6(1), 159–165.
- Haraway, D. J. (2016). *Staying with the Trouble: Making Kin in the Chthulucene*. Duke Univ Press, UK. ISBN:978-0-8223-6224-12020. French ed., 2020. *Vivre avec le trouble*. Les Editions des mondes à faire, France. 978-2-9555738-4-6.
- Haustein, K., Otto, F. E. L., Venema, V., Jacobs, P., Cowtan, K., Hausfather, Z., Way, R. G., White, B., Subramanian, A., and Schurer, A. P. (2019). A limited role for unforced internal variability in twentieth-century warming. *J. Climate*, 32, 4893–4917.
- Hempel, C. G. (1942). The function of general laws in history. *J. Philos.*, 39(2), 35–48.
- Hoquet, T. (2009). *Darwin contre Darwin : comment lire L'origine des espèces*. Seuil, France. ISBN:978-2-02-098299-3.
- Hoquet, T. (2021). Nature, l'impossible ailleurs. In Bossi, L., editor, *Les Origines du monde. L'invention de la nature au XIX<sup>e</sup> siècle*, pages 348–351. Musée d'Orsay / Musée des beaux-arts de Montréal / Muséum d'histoire naturelle / Gallimard, France. ISBN:978-2-35433-318-8.
- Houllier, F., Joly, P.-B., and Merilhou-Goudard, J.-B. (2017). Les sciences participatives : une dynamique à conforter. *Nat. Sci. Soc.*, 25(4), 418–423.
- IPBES (2019). *Summary for Policymakers of the Global Assessment Report on Biodiversity and Ecosystem Services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*. Díaz, S., Settele, J., Brondízio, E. S., Ngo, H. T., Guèze, M., Agard, J., Arneeth, A., Balvanera, P., Brauman, K. A., Butchart, S. H. M., Chan, K. M. A., Garibaldi, L. A., Ichii, K., Liu, J., Subramanian, S. M., Midgley, G. F., Miloslavich, P., Molnár, Z., Obura, D., Pfaff, A., et al., editors, IPBES Secretariat, Germany.
- IPCC (2014). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Core Writing Team, R. K. Pachauri and L. A. Meyer, editors, IPCC, Geneva, Switzerland. ISBN:978-92-9169-143-2.
- IPCC (2022a). *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Shukla, P. R., Skea, J., Slade, R., Al Khourdajie, A., van Diemen, R., McCollum, D., Pathak, M., Some, S., Vyas, P., Fradera, R., Belkacemi, M., Hasija, A., Lisboa, G., Luz, S., and Malley, J., editors, Cambridge University Press, UK and USA.
- IPCC (2022b). Summary for policymakers. In Pörtner, H.-O., Roberts, D. C., Poloczanska, E. S., Mintenbeck, K., Tignor, M., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., and Okem, A., editors, *Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, UK and USA.
- Juan, M. (2021). Les recherches participatives à l'épreuve du politique. *Sociol. Trav.*, 63, 1–26.
- Jullien, F. (2016). *Il n'y a pas d'identité culturelle, mais nous défendons les ressources d'une culture*. Editions de l'Herme, France. ISBN:97828519783295.
- Kuhn, T. S. (1970). *The Structure of Scientific Revolutions*. University of Chicago Press, USA. ISBN:978-0-226-45803-8 (first ed., 1962).
- Lacroix, A. (2021). Rendre visibles des choses invisibles. Entretien avec Philippe Descola. *Philos. Mag.*, 152, 64–70.
- Latour, B. (2012a). *Enquête sur les modes d'existence. Une anthropologie des Modernes*. La Découverte, France. ISBN:978-2-7071-7347-8.
- Latour, B. (2012b). Les mille aventures de la connaissance objective. Préface. In Daston, L. and Galison, P., *Objectivité*, pages 9–14. Les presses du réel, France. ISBN:978-2-8466-334-8.
- Latour, B. (2015). *Face à Gaiä, huit conférences sur le nouveau régime climatique*. Les empêcheurs de penser en rond. La Découverte, France. ISBN:978-2-35925-108-1.
- Latour, B. (2021). *Où suis-je ? Leçons du confinement à l'usage des terrestres*. Les empêcheurs de tourner en rond. La Découverte, France. ISBN:978-2-35925-200-2.
- Latour, B. (2022). Introduction générale sur les modes d'existence. Entretien avec Adèle Van Reeth, radio France Culture. Episode 2 : la pratique des sciences et la question de la nature 22 mars. <https://www.franceculture.fr/emissions/les-chemins-de-la-philosophie/les-chemins-de-la-philosophie-du-mardi-22-mars-2022>.
- Le Duff, M. (2018). Les risques naturels côtiers en Nouvelle-Calédonie : contribution pour une

- gestion intégrée : de la caractérisation du risque à la participation citoyenne, quelques perspectives pour la prévention des risques aux îles Loyauté. Thèse de doctorat. Univ Nouvelle-Calédonie, France. Accessed 23 March 2022, <https://www.theses.fr/2018NCAL0005/document>.
- Le Duff, M., Dumas, P., Allenbach, M., and Cohen, O. (2020). An orientation for coastal disaster risks management and prevention policy in a global warming context: Case study in Ouvea (New Caledonia). *Mar. Policy*, 117, 1–13.
- Le Duff, M., Dumas, P., Sabinot, C., and Allenbach, M. (2016). Le risque tsunami en Nouvelle-Calédonie : Evolutions des facteurs de vulnérabilités et de résiliences à Lifou en territoire coutumier kanak. *VertigO - la revue électronique en sciences de l'environnement*, 16(3), 32p.
- Leenhardt, M. (1985). *Do Kamo, la personne et le mythe dans le monde mélanésien*. Gallimard, France. ISBN:2-07-070412-2 (first ed., 1947).
- Lenton, T. M. and Latour, B. (2018). Gaia 2.0. *Science*, 361, 1066–1068.
- Lenton, T. M. and Wilkinson, D. M. (2003). Developing the Gaia theory. A Response to the Criticisms of Kirchner and Volk. *Clim. Change*, 58, 1–12.
- Lévy, J. and Lussault, M. (2003). *Dictionnaire de la Géographie et de l'espace des sociétés*. Belin, France. ISBN:978-2-7011-2645-6.
- Lovelock, J. E. (1989). Geophysiology, the science of Gaia. *Rev. Geophys.*, 27(2), 215–222.
- Lussault, M. (2007). *L'Homme spatial. La construction sociale de l'espace humain*. La couleur des idées. Seuil, France. ISBN:9782020937955.
- Lyell, C. (2019). *Principles of Geology: Being an Inquiry How far the Former Changes of the Earth's Surface Are Referrable to Causes Now in Operation*, volume 3. Hardpress Publishing, USA. (first ed., 1833).
- Markard, J., Geels, F. W., and Raven, R. (2020). Challenges in the acceleration of sustainability transitions. *Environ. Res. Lett.*, 15, article no. 081001.
- Masson-Delmotte, V. (2020). Réchauffement climatique : état des connaissances scientifiques, enjeux, risques et options d'action. *C. R. Geosci.*, 352, 251–277.
- Meadows, D., Randers, J., and Meadows, D. (2004). *Limits to growth: The 30-Year Update*. Chelsea Green Pub, USA. ISBN 10 :193149858X. French ed., 2022. Les limites à la croissance (dans un monde fini), le rapport Meadows 30 ans après. Ed. Rue de l'échiquier, France. ISBN:978-2-37425-332-9.
- Mill, J. S. (1868). *Auguste Comte et le positivisme*. Libraire-éditeur Baillière Germer, France.
- Mitja, D., Coudrain, A., Barrière, O., Begue, A., Bonnet, M. P., Cubizolles, S., David, G., Delaître, E., Delaplace, J. M., Dessay, N., Faure, J. F., Gervet, C., Le Duff, M., Longépée, E., Patel, P., Rousseau, V., Sabinot, C., Saito, H. C., Tabau, A. S., and Viale, F. (2021). Dialogue entre les parties prenantes : un levier dans la mutation de la science chahutée par l'Anthropocène. In *Actes des sessions parallèles, Univ. Lorraine, France. Colloque Science & you 2021 "un temps de réflexion pour avancer"*, Nov 2021, pages 314–316. Metz, France. (hal-03799523).
- Morin, E. (1982). *Science avec conscience*. Fayard, France. ISBN:9782213011516.
- Notre affaire à tous (2022). *Les droits de la Nature. Vers un nouveau paradigme de protection du vivant*. Le Pommier, France. Yzquierdo, M. (coordinator). ISBN:978-2-7465-2491-0.
- Nouvel, P. (2020). *Avant toutes choses, Enquête sur les discours d'origine*. CNRS, France. ISBN:978-2-271-12252-0.
- Parisi, C. and Horvath, A. (directors). (2021). *Face à Arcimboldo*. Centre Pompidou-Metz, France. ISBN:978-2-35983-062-0.
- Perretti-Watel, P. (2010). *La société du risque*. La découverte, France. ISBN:9782707164568.
- Popper, K. R. (1963). *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge and Kegan Paul, USA. ISBN:9780710065070.
- Raffestin, C. (1987). Pourquoi n'avons-nous pas lu Eric Dardel? *Cah. Géogr. Qué.*, 31(84), 471–481.
- Raffestin, C. (1989). Théorie du réel et géographicité. *Espace-Temps*, 40–41, 26–31.
- Reghezza, M. (2015). *De l'avènement du Monde à celui de la planète : le basculement de la société du risque à la société de l'incertitude. Mémoire d'habilitation à diriger les recherches*. Univ Paris 1-Panthéon Sorbonne. Accessed 19 March 2022 on <https://tel.archives-ouvertes.fr/tel-01255031>.
- Ripple, W. J., Wolf, C., Newsome, T. M., Galetti, M., Alamgir, M., Crist, E., Mahmoud, M. I., Laurance, W. F., and 15,364 scientist signatories from 184 countries (2017). World scientists' warning to humanity: a second notice. *BioScience*, 67(12), 1026–1028.

- Ruz, M.-H., Rufin-Soler, C., Hellequin, A.-P., Hequette, A., Deboudt, P., Cohen, O., Herbert, V., Lebreton, F., Schmitt, F., Le Goff, L., Révillon, R., and Marin, D. (2021). COSACO : Quel littoral dans cinquante ans ? CO-construction de Stratégies d'Adaptation au changement climatique en Côte d'Opale. Rapport de synthèse. hal-03373431.
- Sauermann, H., Vohland, K., Antoniou, V., Balázs, B., Göbel, C., Karatzas, K., Mooney, P., Perelló, J., Ponti, M., Samson, R., and Winter, S. (2020). Citizen science and sustainability transitions. *Res. Policy*, 49, article no. 103978.
- Serrao-Neumann, S. and Coudrain, A. (2017). Science and knowledge production for climate change adaptation: challenges and opportunities. In Serrano-Neumann, S., Coudrain, A., and Coulter, L., editors, *Communicating Climate Change Information for Decision-making*, pages 3–11. Springer, Switzerland.
- Serres, M. (2020). *Le Contrat Naturel*. Flammarion, France. ISBN:978-2-0815-1269-6. First ed. 1990, François Bourin, France. ISBN:978-2-87686-041-4. English ed. 1995, The Natural Contract, Univ of Michigan Press, USA. ISBN:9780472095490.
- Steffen, W., Broadgate, W., Deutsch, L., Gaffney, O., and Ludwig, C. (2015). The trajectory of the Anthropocene: The Great Acceleration. *Anthr. Rev.*, 2(1), 81–98.
- Steffen, W., Persson, A., Deutsch, L., Zalasiewicz, J., Williams, M., Richardson, K., Crumley, C., Crutzen, P., Folke, C., Gordon, L., Molina, M., Ramanathan, V., Rockstrom, J., Scheffer, M., Schellnhuber, H. J., and Svedin, U. (2011). The anthropocene: from global change to planetary stewardship. *Ambio*, 40(7), 739–761.
- Steffen, W., Rockström, J., Richardson, K., Lenton, T. M., Folke, C., Liverman, D., Summerhayes, C. P., Barnosky, A. D., Cornell, S. E., Crucifix, M., Donges, J. F., Fetzer, I., Lade, S. J., Scheffer, M., Winkelmann, R., and Schellnhuber, H. J. (2018). Trajectories of the Earth System in the Anthropocene. *Proc. Natl. Acad. Sci.*, 115, 8252–8259.
- Stengers, I. (1991). Le pouvoir des concepts. In Stengers, I. and Schlanger, J., editors, *Les concepts scientifiques : invention et pouvoir*, pages 29–66. Gallimard, La Découverte, France. ISBN 10 : 2-07032624-1.
- Stengers, I. (2015). Accepting the Reality of Gaia: A fundamental shift? In Hamilton, C., Bonneuil, C., and Gemenne, F., editors, *The Anthropocene and the Global Environmental Crisis: Rethinking Modernity in a New Epoch*, pages 77–93. Routledge, UK. ISBN:9781138821248.
- Stengers, I. (2019). *Résister au désastre*. Wildproject, France. ISBN:9782918490920.
- Storup, B. (coordinator). (2013). La recherche participative comme mode de production de savoirs. Synthèse du rapport Fondation sciences citoyennes, rapport, 8 p. [https://sciencescitoyennes.org/wp-content/uploads/2014/10/Synth%C3%A8se\\_Rapport\\_FdF\\_FSC.pdf](https://sciencescitoyennes.org/wp-content/uploads/2014/10/Synth%C3%A8se_Rapport_FdF_FSC.pdf).
- Syvitski, J., Waters, C. N., Day, J., Milliman, J. D., Summerhayes, C., Steffen, W., Zalasiewicz, J., Cearreta, A., Gałuszka, A., Hajdas, I., Head, M. J., Leinfelder, R., McNeill, J. R., Poirier, C., Rose, N. L., Shotyk, W., Wagnreich, M., and Williams, M. (2020). Extraordinary human energy consumption and resultant geological impacts beginning around 1950 CE initiated the proposed Anthropocene Epoch. *Commun. Earth Environ.*, 1, article no. 32.
- Tengö, M., Brondizio, E. S., Elmqvist, T., Malmer, P., and Spierenburg, M. (2014). Connecting diverse knowledge systems for enhanced ecosystem governance: the multiple evidence base approach. *Ambio*, 43, 579–591.
- UN General Assembly (1948). Universal Declaration of Human Rights. United Nations, 217 (III)A, Paris. <https://www.un.org/en/about-us/universal-declaration-of-human-rights>.
- Weber, L. (1899). Positivisme et rationalisme. *Rev. Métaphys. Morale*, 7(4), 426–445. <https://www.jstor.org/stable/40892250>.
- Wynne, B. (1999). Une approche réflexive du partage entre savoir expert et savoir profane. In *Les cahiers de la sécurité intérieure*, volume 38, pages 219–236. ISSN:1150-1634.
- Zalasiewicz, J., Waters, C. N., Williams, M., Barnosky, A. D., Cearreta, A., Crutzen, P., Ellis, E., Ellis, M., Fairchild, I. J., Grinevald, J., Haff, P. K., Hajdas, I., Leinfelder, R., McNeill, J., Odada, E. O., Poirier, C., Richter, D., Steffen, W., Summerhayes, C., Syvitski, J. P. M., et al. (2015). When did the Anthropocene begin? A mid-twentieth century boundary level is stratigraphically optimal. *Quat. Int.*, 383, 196–203.
- Žukauskaitė, A. (2020). Gaia theory: between auto-poiesis and sympoiesis. *Problemos*, 98, 141–153.

# **Climate Change**





Research article

Geo-hydrological Data & Models

# Climate reconstruction of the Little Ice Age maximum extent of the tropical Zongo Glacier using a distributed energy balance model

Philémon Autin <sup>\*,a</sup>, Jean Emmanuel Sicart <sup>a</sup>, Antoine Rabatel <sup>a</sup>, Regine Hock <sup>b,c</sup>  
and Vincent Jomelli <sup>d</sup>

<sup>a</sup> Univ. Grenoble Alpes, CNRS, IRD, Grenoble-INP, Institut des Géosciences de l'Environnement (IGE, UMR 5001), F-38000 Grenoble, France

<sup>b</sup> Department of Geoscience, Oslo University, Oslo, Norway

<sup>c</sup> Geophysical Institute, University of Alaska, Fairbanks, AK, USA

<sup>d</sup> CNRS Cerege, Technopôle de l'Environnement Arbois-Méditerranée, Aix en Provence, France

*Current address:* Institut des Géosciences de l'Environnement, Bâtiment Glaciologie, 54, rue Molière, 38400 Saint Martin d'Hères, France (P. Autin)

*E-mails:* [philemon.autin@univ-grenoble-alpes.fr](mailto:philemon.autin@univ-grenoble-alpes.fr) (P. Autin), [jean-emmanuel.sicart@ird.fr](mailto:jean-emmanuel.sicart@ird.fr) (J. E. Sicart), [antoine.rabatel@univ-grenoble-alpes.fr](mailto:antoine.rabatel@univ-grenoble-alpes.fr) (A. Rabatel), [regine.hock@geo.uio.no](mailto:regine.hock@geo.uio.no) (R. Hock), [jomelli@cerege.fr](mailto:jomelli@cerege.fr) (V. Jomelli)

**Abstract.** This study assessed the climate conditions that caused the tropical Zongo Glacier (16° S, Bolivia) to reach its Little Ice Age (LIA) maximum extent in the late 17th century. We carried out sensitivity analyses of the annual surface mass balance to different physically coherent climate scenarios constrained by information taken from paleoclimate proxies and sensitivity studies of past glacier advances. These scenarios were constrained by a 1.1 K cooling and a 20% increase in annual precipitation compared to the current climate. Seasonal precipitation changes were constructed using shuffled input data for the model: measurements of air temperature and relative humidity, precipitation, wind speed, incoming short and longwave radiation fluxes, and assessed using a distributed energy balance model. They were considered plausible if conditions close to equilibrium glacier-wide mass balance were obtained. Results suggest that on top of a 1.1 K cooling and ~20% increase in annual precipitation, only two seasonal precipitation patterns allow LIA equilibrium: evenly distributed precipitation events across the year and an early wet season onset.

**Keywords.** Tropical glacier, Surface energy balance modeling, Little Ice Age climate, Andes, Climate reconstruction.

*Manuscript received 29 March 2022, revised 22 June 2022 and 27 June 2022, accepted 18 July 2022.*

## 1. Introduction

The Little Ice Age (LIA) was a multi-secular cold period, lasting approximately from 1350 to 1850 AD,

during which glaciers advanced worldwide [e.g., Solomina et al., 2015, 2016]. The magnitude and timing of the glacial advances was not synchronous worldwide [Grove, 1988, Rabatel et al., 2008, Neukom et al., 2014, Solomina et al., 2015, 2016]. The evolution of glaciers since the Little Ice Age has been most

\* Corresponding author.

frequently studied in mid-latitudes. For example, in the French Alps glaciers reached their maximum extent around the first half of the 17th century, and their front oscillated in a position slightly up-valley before re-advancing close to their maximum extent throughout the first half of the 19th century [e.g., Grove, 1988, Vincent *et al.*, 2004, 2005, Solomina *et al.*, 2015, 2016]. [Paul and Bolch, 2019] reviewed a number of studies of the evolution of Alpine glaciers since the LIA and concluded that it could be summarized by a km scale retreat, a rise in the equilibrium-line altitude (ELA) of 100–200 m, and an up-valley retreat of several hundred meters in glacier terminus positions. In Pyrenean glaciers, Campos *et al.* [2021] reported ELA rises of 220 m.

On the other hand, the evolution of glaciers in Tropical South America since the LIA has been the subject of few studies. The majority of glaciers reached their maximum extent between 1630 and 1720 and have been retreating ever since; their retreat has been interrupted several times by minor advances [Rabatel, 2005, Rabatel *et al.*, 2005, 2008, Jomelli *et al.*, 2009, 2011]. Based on the dating of 15 Bolivian glaciers using lichenometry, radiocarbon or  $^{10}\text{Be}$  techniques, Rabatel *et al.* [2008] estimated that their ELA rose by  $285 \pm 50$  m since the LIA while, based on the Telata Glacier also in Bolivia, Jomelli *et al.* [2011] estimated that the ELA rose by 320 m.

Understanding past glacier variations is key to better insights into future changes. In the case of tropical glaciers, this is of particular interest as they represent an important water resource for Andean populations [e.g., Soruco *et al.*, 2015]. Tropical glaciers react rapidly to climate perturbations [by advancing or retreating, e.g., Kaser, 2001, Thompson *et al.*, 2006] making them suitable paleoclimate proxies to evaluate past climate scenarios.

Over tropical South America, paleoclimate proxies such as ice cores, sediments, tree rings or speleothems [e.g., Thompson *et al.*, 2006, Ledru *et al.*, 2013, Apaéstegui *et al.*, 2018, Campos *et al.*, 2019, Bird *et al.*, 2011, Morales *et al.*, 2012] and glacier sensitivity analyses [e.g., Rabatel *et al.*, 2005, 2008, Malone *et al.*, 2015] provide valuable temperature and precipitation data on the LIA. The main drawback of these methods is the temporal scale (annual to decadal). For example, glacier studies usually use simple approaches [e.g. simplified energy balance

methods, Hastenrath and Ames, 1995] which only allow comparison of glacier steady state conditions at two different periods, but do not provide detailed insights into seasonal changes in climatic forcing on the glacier mass balance.

The annual mass balance of South American tropical glaciers depends primarily on the timing and length of the wet season during the austral summer [Sicart *et al.*, 2011] and they are highly sensitive to the temporal distribution of precipitation between September and the onset of the wet season [Autin *et al.*, 2022]. Several studies suggest that the onset and length of the wet season may have changed during the Little Ice Age [e.g. Diaz and Vera, 2018], a feature that cannot be captured by models using annual temperature and precipitation amounts. Hence, such models may largely underestimate past mass balance variability [Sicart *et al.*, 2011]. In addition, sub-annual precipitation records are rare over low latitude mountains and precipitation predictions are unreliable, making reconstruction of past glacier advances and estimating future retreat challenging in these regions.

In this study, we used the distributed energy balance model [DEBAM, Hock and Holmgren, 2005] which had been adapted to tropical glaciers by Sicart *et al.* [2011]. This model is based on equations of mass and energy conservation and, in theory, the physical interpretation of the parameters can be linked to measurable quantities [Beven, 1989]. However, because natural sub-grid heterogeneities can lead to a mismatch of scales between measured and modeled variables, the calibration process is often complicated [Klemes, 1983, Seibert, 1999]. Thus, to limit the risk of using a model calibration that makes the model subject to the three rules of classical tragedy: unity of place, unity of action and unity of time [De Marsily, 1994], a parsimonious approach was chosen to minimize the number of parameters with the aim of matching model complexity to the availability of field observations. In this way, we expect that even if the models represent more fiction than reality, the fiction nevertheless helps understand reality [Seibert, 1999].

We chose to apply the model to the tropical Zongo Glacier (16° S, Bolivia) because this glacier's present-day climate controls have already been well studied [e.g., Rabatel *et al.*, 2013, Autin *et al.*, 2022] and its LIA maximum reconstructed [Rabatel, 2005]. Run at an



hourly time step, this model makes it possible to investigate the effect of changes in seasonal precipitation patterns on the surface mass balance at the LIA maximum extent of the glacier.

For this purpose, we created scenarios that differ in precipitation amounts and seasonality but where the physical consistency between measured meteorological variables is retained (i.e. physically-coherent scenarios). The scenarios that resulted in an equilibrium state (surface mass balance close to zero) were considered plausible LIA scenarios.

## 2. Study site and background on the LIA in the tropical Andes

### 2.1. Study site and present climate

Zongo Glacier (16° 15' S, 68° 10' W, Cordillera Real, Bolivia) is a valley-type glacier extending from ~6000 m a.s.l. down to ~4950 m a.s.l. A glacier monitoring program began in 1991 through collaboration between a Bolivian university (*Universidad Mayor de San Andrés*) and IRD [French Research Institute for Development, Francou *et al.*, 1995, Ribstein *et al.*, 1995]. Rabatel *et al.* [2013] and Sicart *et al.* [2015] provide a detailed overview of the monitoring network, which includes three automatic weather stations (AWS), one on the glacier, and two off-glacier, along with a discharge gauging station, and several ablation stakes and snow pits (Figure 1).

Being in the outer tropics, the climate over the glacier is characterized by marked precipitation and cloud cover seasonality, with a single wet season in austral summer and a pronounced dry season in austral winter [Troll, 1941]. A number of authors have identified and described three seasons with distinct melt mechanisms over this tropical glacier where the hydrological year starts on the first of September [e.g., Sicart *et al.*, 2011, Rabatel *et al.*, 2012, Autin *et al.*, 2022].

The transition season (September to November) is characterized by increasing melt rates due to limited precipitation [around 20% of the annual amount, Ramallo, 2013] and increasing solar radiation as the top of atmosphere radiation approaches its maximum annual value reached at the austral summer solstice [Sicart *et al.*, 2011]. Autin *et al.* [2022] simulated nine years of distributed surface energy balance over Zongo Glacier using DEBAM

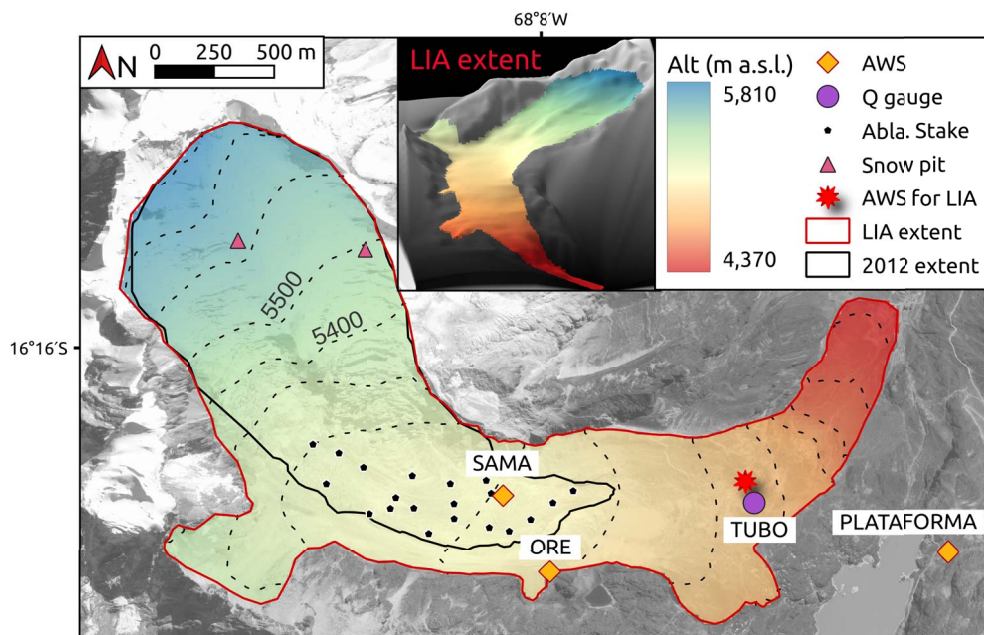
and showed that this period plays a key role in controlling annual mass losses and that the timing of precipitation events in the season is a key driver of melt.

The wet season (December to March) is linked to the mature phase of the South American monsoon system (SAMS). During the SAMS, a strong anticyclonic system in the upper troposphere [the Bolivian High, Lenters and Cook, 1997] coupled with the southward displacement of the South American low-level jet (SALLJ) result in strong easterly winds on the eastern slopes of the Cordillera Real. This leads to moisture transport over the Cordillera Real, which, coupled with solar heating of the surface, results in frequent afternoon/early evening convection events over Zongo [e.g., Garreaud *et al.*, 2003]. This season is characterized by decreasing melt rates due to the strong cloud cover and albedo feedback effect. Over this period, little ice mass loss is observed as most of the melt comes from large snowfall events [70% of the annual amounts, Ramallo, 2013]. In addition, it is the glacier accumulation season.

April–May is a second transition period, from the end of the wet season to the onset of the dry season (June to August). During this period, the precipitation amounts and distribution are highly variable due to the interannual variability of the demise of the wet season (based on 18 years of measurements, the amount of precipitation received over this period is  $50 \pm 27$  mm w.e.).

The dry season is characterized by low melt rates and rare precipitation events. The absence of clouds in the dry season creates an incoming longwave radiation deficit leaving little energy available for melt [Sicart *et al.*, 2011]. In addition, the dry and cold air coupled with high wind speeds favors sublimation over melt [Wagnon *et al.*, 1999, Sicart *et al.*, 2005], limiting mass loss through this period.

A number of studies [e.g., Sicart *et al.*, 2005, Wagnon *et al.*, 1999] identified the net all-wave radiation flux as the main driver of melt at seasonal and interannual time scales, more specifically, the incoming shortwave radiation flux. Sicart *et al.* [2005, 2016] showed that the marked cloud and precipitation seasonality is an important factor in controlling seasonal mass balance variability. Building on this, Autin *et al.* [2022] showed that melt energy is highest during the transition season and that the interannual variability of the surface mass balance is mostly driven by



**Figure 1.** Zongo Glacier (16° 15' S, 68° 10' W) showing its monitoring network. The red outline represents the maximum extent of the LIA and the corresponding digital elevation model, the black line shows the extent of the glacier in 2012. The map in the top inset is a 3D view of the extent of the glacier at its LIA maximum extent. SAMA (5050 m a.s.l.), ORE (5050 m a.s.l.) and PLATAFORMA (4750 m a.s.l.) are the three automatic weather stations (AWS) installed on and around the glacier, TUBO is a discharge gauging station located at 4830 m a.s.l. The red star shows the position at which the measurements are used to force the model in LIA simulations. Source of the background image: Pleiades satellite in 2013, © CNES—Airbus D&S. Digital Elevation Model: Rabatel [2005].

the temporal distribution of precipitation events and their amounts and, to a lesser extent, by the cloud cover throughout this season.

## 2.2. LIA climate reconstructions in the tropical Andes

Ice cores are paleoclimate proxies that provide valuable information at seasonal and annual scales over the last 1000 years as well as information at a coarser temporal scale over the last 20,000 years [Vimeux *et al.*, 2009, Hurley *et al.*, 2019]. Using an ice core drilled in the Quelccaya ice cap (QIC, southern Peru), Thompson *et al.* [1986, 2006] showed a significant decrease in both  $\delta^{18}\text{O}$  and dust concentration during the LIA (1500–1800 AD) were associated with a decrease in temperatures. More recently, Hurley *et al.* [2019] showed that changes in  $\delta^{18}\text{O}$  are primarily linked with ENSO activity which, in turn, is related

to changes in the activity of the South American monsoon system (SAMS) rather than temperature changes. Thus, in agreement with Hurley *et al.* [2019], we consider that the lower  $\delta^{18}\text{O}$  observed during the LIA is associated with increased precipitation rather than with colder temperatures.

Morales *et al.* [2012] analyzed *Polylepis Tarapacana* tree rings across the Altiplano to reconstruct the climate between 1300 and 2000 AD and found that sustained wet conditions prevailed between 1600 and 1750 AD leading to precipitation positive anomalies (up to 40% at its highest with respect to the average measurement for the period 1982–2000).

Speleothems [Campos *et al.*, 2019, Apaéstegui *et al.*, 2018, Vuille *et al.*, 2012] and sediments of diverse origin [Ledru *et al.*, 2013, Bird *et al.*, 2011, Sachs *et al.*, 2009, Haug *et al.*, 2001] provide information on changes in temperature and precipitation patterns over the past 2000 years. Their wide spatial availability across tropical South America enables a 2000-year reconstruction of the SAMS. These studies suggest that a southern shift of the Inter Tropical Convergence Zone (ITCZ) linked to decreased radiative forcing [Maunder minimum, 1645–1715 AD, Lean, 2000] caused colder North Atlantic Sea Surface Temperatures during the LIA, resulting in an intensification of SAMS that led to unprecedented moisture levels over the Eastern Cordillera.

Diaz and Vera [2018] assessed four Paleoclimate Model Intercomparison Project 3/Climate Model Intercomparison Project 5 [PMIP3/CMIP5, Braconnot *et al.*, 2011] model ensemble runs between 850 and 2005 AD over South America. These authors found that the models can correctly reproduce temperature variations over the past millennium (with respect to paleoclimate proxy information) but cannot represent the associated changes in precipitation. Similarly, we considered temperature and precipitation time series of two ensemble runs with paleoclimate proxy information from Michel [2020] and Tardif *et al.* [2019] thereby corroborating the findings of Diaz and Vera [2018].

Diaz and Vera [2018] also analyzed large-scale atmospheric conditions (250 hPa and 800 hPa zonal winds) of PMIP3/CMIP5 ensemble runs and showed an equatorial shift of the southern branch of the Hadley cell during the LIA that resulted in a wetter tropical South America [Lu and Vecchi, 2014]. However, because the simulated precipitation amounts do not match proxy information, these models cannot be used to compare past precipitation amounts and present conditions.

Glacier modeling can be used to assess climate conditions which lead to a known past glacier extent. Based on a synthesis of sensitivity studies of tropical South American glaciers, Kaser [1999] attributed two thirds of the glacier retreat in the Cordillera Blanca since the LIA to a decrease in vapor pressure, and the remaining third to an increase in air temperature. Malone *et al.* [2015] applied a coupled ice-flow mass balance model on the QIC to reconstruct annual climate conditions during the LIA and concluded that

the glacier is more sensitive to temperature than to precipitation. Finally, Rabatel *et al.* [2005, 2008] applied the Kaser [2001] model and Hastenrath [1984] approach to Bolivian tropical glaciers and found that for the glaciers to have reached their maximum extent during the LIA, a 20–30% increase in precipitation (that goes hand-to-hand with an increase in cloudiness) and cooler temperatures ( $\sim -1$  to  $-1.2$  K) with respect to today's climate conditions would have been required.

These different studies suggest that during the LIA, precipitation amounts were between 20 and 40% higher than today and that the temperature was 1 to 1.5 °C lower. However, uncertainties on how this extra precipitation was distributed within the year remain, which is just what our study aims to clarify, as using an energy balance model is the only way to investigate the effect of seasonal precipitation patterns on the annual surface mass balance.

### 3. Method

We used the distributed surface energy balance model [DEBAM, Hock and Holmgren, 2005] to simulate the glacier surface mass balance of Zongo Glacier over the reconstructed surface area of the glacier during the LIA. The model is forced by hourly time series of meteorological data based on measurements recorded by an automatic weather station for nine years distributed between 1999 and 2017. The observations are used to construct four climate scenarios that are then assessed based on their ability to yield a near-balanced mass budget. Further details are given below.

#### 3.1. Mass balance model

##### 3.1.1. Model description

DEBAM simulates the glacier surface mass balance and melt discharge at hourly intervals over the grid cells of a digital elevation model ( $20 \times 20$  m cells) using an energy-balance approach [Hock and Holmgren, 2005]. In this study, we use the multi-year calibration of Zongo by Autin *et al.* [2022]. The model requires seven input meteorological variables (air temperature and relative humidity, wind speed, precipitation, incoming short and longwave radiation and

outgoing longwave radiation) and solves the following surface energy balance equation:

$$Q_M = SW_{in}(1 - \alpha) + LW_{net} + H + LE + Q_G + R \quad (1)$$

where  $Q_M$  is the melt energy,  $SW_{in}$  is the incoming shortwave radiation,  $\alpha$  the albedo,  $LW_{net}$  the net longwave radiation balance,  $H$  and  $LE$  are the turbulent sensible and latent heat fluxes, respectively,  $Q_G$  is the subsurface heat flux and  $R$  the sensible heat supplied by rain [negligible over Zongo Glacier, Sicart *et al.*, 2011]. The model convention is such that energy fluxes directed towards the surface are positive and those away from the surface are negative.

Incoming shortwave radiation was extrapolated across the glacier by splitting it into its direct and diffuse components. The albedo was modeled using a version of Oerlemans and Knap [1998]'s albedo parameterization adapted to tropical glaciers. This parameterization accounts for the rapid alternation (a few days) of accumulation and melt in the wet season, and the impact of ice on the albedo over shallow snow depths [Sicart, 2002]. The incoming longwave radiation flux was assumed to be spatially constant.

Surface temperature was calculated at the AWS from the outgoing longwave radiation measurements and an altitudinal gradient of  $-0.1$  K/100 m was applied to extrapolate surface temperatures across the glacier (with an upper limit of 273.15 K). The subsurface snow module [Hock and Tijm-Reijmer, 2012] was not used to evaluate spatial changes in the surface temperature due to lack of data to constrain the model (i.e. subsurface temperature and density profiles). This implies that we neglected the sub-surface heat flux in our simulations.

The sensible and latent turbulent heat fluxes were calculated according to the aerodynamic profile method between the glacier surface and the measurement height (wind speed, air temperature and relative humidity) based on the Monin–Obukhov similarity theory [see Hock and Holmgren, 2005 for more details]. The stability functions were based on Beljaars and Holtslag [1991] for stable conditions and on the Businger–Dyer expressions from Paulson [1970] for unstable conditions. Relative humidity and wind speed were considered spatially constant across the glacier.

Discharge at the glacier terminus was simulated using a runoff module that routes glacier melt and rain water through the glacier using three linear

reservoirs representing firn, snow and ice [storage constants of 350, 30 and 16 h, respectively, taken from Sicart *et al.*, 2011]. In addition, for non-glacierized surfaces (30% of the catchment in the recent climate), a runoff coefficient of 0.8 was considered [based on Ribstein *et al.*, 1995].

### 3.1.2. Model calibration and validation

We used the model parameter values from Autin *et al.* [2022], who calibrated and validated the model over Zongo Glacier over nine years distributed between 1999 and 2017. This calibration, summarized on Table 1, was a trade-off between years with high and low mass losses where particular emphasis was placed on accurately representing the transition and the wet season, as they play a key role both in the annual surface mass balance and in year-to-year variability. The transition season is the period when the glacier loses on average half its annual mass in the ablation zone, while the wet season is the accumulation period. Because their calibration did not include the multi-layer snow model, melt is overestimated during the dry season (JJA) which contributes to most of the average annual net mass loss overestimation [based on the 9 simulated years, the mass loss overestimation is 0.18 m w.e., when the average mass loss is  $-0.88$  m w.e., Autin *et al.*, 2022].

Analysis of 19 years of daily discharge and precipitation data provided by Ramallo [2013] showed that only 9% of the annual volume passes through the weir at TUBO (Figure 1) during the dry season. Not all of this discharge is a result of glacier mass loss, as an average of 4.3% of the annual precipitation occurred over the same period. Nonetheless, most of the discharge observed at TUBO comes from the glacier meltwater as its catchment is 70% glacierized and the snow on the moraines melts completely in a few days [e.g. Lejeune, 2009]. This low dry season discharge highlights the limited mass loss that occurs at this time of the year [see also Autin *et al.*, 2022, Sicart *et al.*, 2011].

Therefore, the dry season is generally a period of low ablation whereas the high variability of the mass balance in the transition and wet seasons mostly controls the interannual ablation and accumulation variability. As model uncertainties are high during the dry season (importance of poorly known turbulent fluxes plus we disregard the subsurface heat flux), we end the simulations on May 31.

**Table 1.** Summary of the main parameter values used for the DEBAM simulations

Parameter	Value
Fresh snow albedo	$\alpha_{\text{fresh-snow}} = 0.85$
Firn albedo	$\alpha_{\text{firn}} = 0.6$
Ice albedo	$\alpha_{\text{ice}} = 0.3$
Clear sky transmissivity for solar radiation	$\tau_{\text{clear-sky}} = 0.8$
Roughness height of momentum over ice	$z_{0\text{ice}} = 0.026 \text{ m}$
Roughness heights of temperature, humidity and momentum over snow	$z_{0\text{snow}} = z_{0T} = z_{0q} = 0.0026 \text{ m}$
Precipitation gradient from 4370 to 5200 m asl	$\frac{\partial P}{\partial z} = +10\%/100 \text{ m}$
Glacier surface temperature gradient	$\frac{\partial T_s}{\partial z} = -0.1 \text{ K}/100 \text{ m}$

To achieve a 1 K cooling representative of the climate in the late 17th century, the AWS measurements were applied to a grid cell at 4850 m a.s.l. instead of at 5050 m a.s.l. (the method is detailed in Section 3.2.1). To avoid unrealistically high precipitation amounts at the top of the glacier caused by lowering the elevation of the forcing data, we applied the precipitation gradient of +10%/100 m used by Autin *et al.* [2022] from the glacier front up to 5200 m a.s.l. [instead of up to 5400 m a.s.l. in Autin *et al.*, 2022].

### 3.2. Input datasets

#### 3.2.1. Glacier extent at the LIA maximum extent and model initialization

Using lichenometry techniques, the LIA maximum extent of the glacier was dated to  $1680 \pm 28 \text{ AD}$  [Rabatel *et al.*, 2008]. A digital elevation model (DEM) of Zongo Glacier was reconstructed by redrawing contour lines so that the glacier fills the moraines [Rabatel, 2005]. Between the LIA maximum extent and 2012, the glacier retreated 1.5 km, the terminus elevation rose by 550 m and the glacier lost 43% of its surface area and  $0.16 \pm 0.01 \text{ km}^3$  of ice (see the red and black lines outlining the glacier on Figure 1).

We hypothesize that since tropical glaciers react rapidly to climate perturbations [Kaser, 1999, Rabatel, 2005], a hydrological year for which the glacier is close to equilibrium ( $B \approx 0 \text{ m w.e.}$ ) in the current climate or during the LIA is representative of their respective steady state climate.

To obtain information on potential LIA glacier equilibrium characteristics (to serve as comparison for the simulation results), we evaluated the LIA steady state equilibrium-line altitude (ELA<sub>0</sub>) considering the frequently used constant area accumulation ratio (AAR) method [e.g. Rabatel, 2005, Jomelli *et al.*, 2009]. We evaluated the current climate ELA<sub>0</sub> (5233 m a.s.l.) by linearly interpolating 25 observed annual ELAs as a function of surface mass balance. We then defined the corresponding AAR (0.68) based on the mean AAR value obtained from the existing glacier digital elevation models [DEM, Autin *et al.*, 2022]. Considering a constant AAR, we derived the LIA ELA<sub>0</sub> (ELA<sub>0,LIA</sub> = 5078 m a.s.l., i.e. 155 m lower).

The equilibrium snow line altitude (SLA<sub>0</sub>) for the LIA maximum extent was considered to have shifted by the same elevation as the ELA<sub>0</sub> putting it at an altitude of 4883 m a.s.l. considering the recent climate SLA<sub>0</sub> = 5037 m a.s.l. [based on data from Rabatel *et al.*, 2012]. The SLA<sub>0</sub> defined this way was considered as the limit of the firn for model initialization.

#### 3.2.2. LIA climate scenarios

We created four LIA climate scenarios based on nine years of available meteorological data between 1999 and 2017 [Autin *et al.*, 2022]. All four scenarios assume a 1 K reduction in air temperature compared to the current climate. Three scenarios assume a 19–22% increase in annual precipitation based on paleoclimate proxy information and glacier climate studies of the past (Section 2.2), but differ in the way

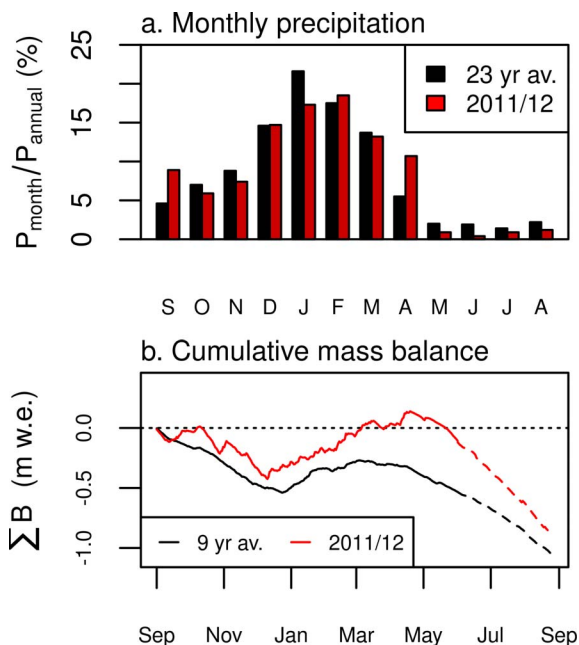
the additional precipitation is distributed across the year:

- Reference scenario: no difference in precipitation compared to the current climate
- Scenario 1: an earlier onset and longer wet season
- Scenario 2: a wetter wet season (December to March)
- Scenario 3: a wetter transition season (September–November) and April–May (the second transition period).

The reference scenario (hydrological year 2011/12) was selected from the nine years of glacier observations so that the precipitation seasonality most closely matched the average monthly variations observed in the 23 years of precipitation from PLATAFORMA, which we assumed to represent current climate precipitation seasonality (Figures 2a, 3a). We use precipitation data from this AWS, rather than the on-glacier SAMA AWS to define average seasonality due to the much longer record (the annual amounts are well correlated:  $R^2 = 0.74$ ). In addition, when the model is applied to the 2012 geometry, the surface mass balance up to June for 2011/12 is close to 0 m w.e. making it an ideal candidate to assess the impact of cooling on the glacier LIA extent mass balance (Figure 2b).

Scenarios 1 to 3 were obtained by replacing selected months in 2011/12 by the same month's data from other years in the 9-year data set, chosen to yield an increase of  $\sim 20\%$  in annual precipitation compared to the long term average. We chose this procedure in preference to artificially adjusting precipitation amounts to preserve consistency between all meteorological variables for each month. In detail, we constructed the three precipitation scenarios as follows:

Scenario 1 (an earlier and longer wet season) was obtained by replacing the 2011/12 October and November data by the data from the same month of another year chosen among the two wettest months in the dataset (Figure 3b). Scenario 2 (a more intense wet season) was obtained by replacing the months of December to March by those in the dataset that had the highest number of events  $\geq 2 \text{ mm}\cdot\text{d}^{-1}$  (Figure 3c). Finally, scenario 3 (wetter September, October, November, April and May), was obtained by selecting the months in the transition seasons that had above

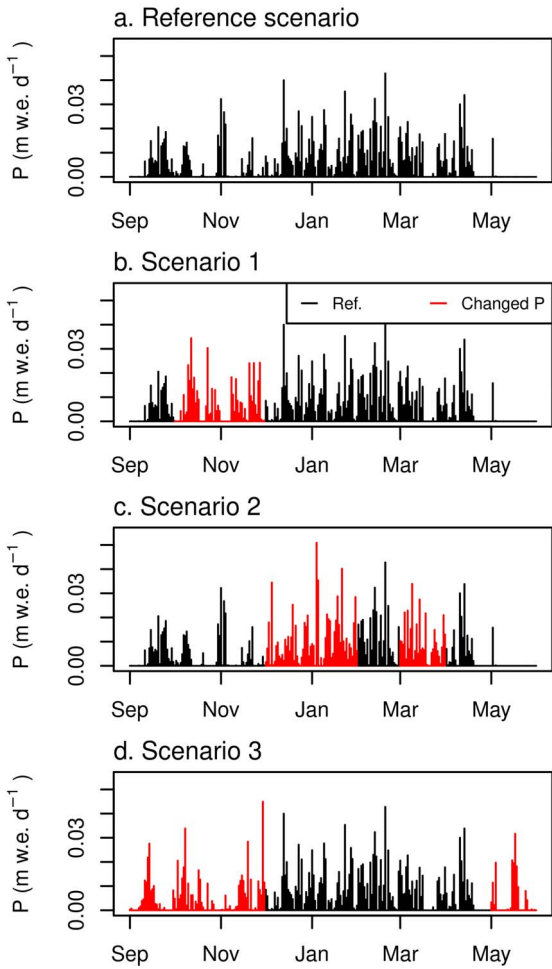


**Figure 2.** (a) Proportion of annual precipitation for each month at the on-glacier AWS SAMA (2011/12) and off-glacier AWS PLATAFORMA (average of 23 years between 1991 and 2015). (b) Simulated cumulative surface mass balance for the 9-year average (1999–2017) and 2011/2012 (reference scenario) over the current glacier geometry. The dashed part of the cumulative surface mass balance lines represent the period for which the model systematically overestimates melt since the subsurface heat flux is neglected.

average precipitation compared to the long term on-glacier average (Figure 3d). For all the scenarios, we used a trial-and-error procedure to determine which year's monthly data to use to replace corresponding data in the 2011/2012 dataset, to achieve the targeted  $\sim 20\%$  increase in annual precipitation amounts.

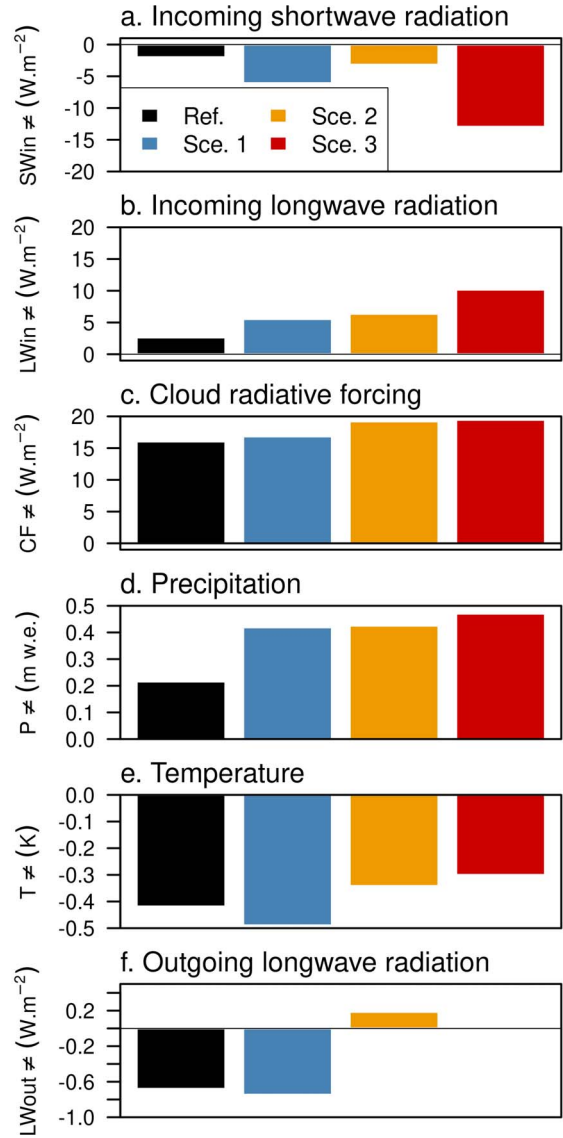
The reference scenario has 1.70 m w.e. of annual precipitation while scenarios 1 to 3 have 1.92, 1.92 and 1.97 m w.e., respectively, which represents a 6% increase in precipitation for the reference scenario, of 19% for scenarios 1 and 2, and a 22% increase for scenario 3, compared to the average of the nine years of the current climate (1.61 m w.e.).

Comparing the scenarios with the 9-year average (up to May 31) revealed that all the scenarios have less incoming shortwave radiation (Figure 4a) and in-



**Figure 3.** Daily precipitation in the four LIA climate scenarios analyzed in this study. The modifications applied to the reference scenario (2011/2012) are in red. The precipitation distribution after the end of May is not accounted for in the simulations.

creased longwave radiation (Figure 4b). The latter is due to increased cloud radiative forcing that summarizes the impact of the clouds on the incoming radiation fluxes (Figure 4c). This is a direct result of selecting/creating years with above-average precipitation (Figure 4d). All scenarios are colder than the measured 9-year average (Figure 4e), a result of the increased cloudiness and associated reduced incoming radiation. Finally, all the scenarios have annual scale glacier surface temperatures close to the nine-



**Figure 4.** Annual scale differences between the scenarios and the 9-year average ( $X_{Sce} - X_{RT}$ ) of (a) incoming shortwave and (b) longwave radiation, (c) cloud radiative forcing, (d) precipitation, (e) air temperature and (f) outgoing longwave radiation used to derive the glacier surface temperature.

year average (illustrated by the mean outgoing longwave radiation in Figure 4f).

## 4. Results and discussion

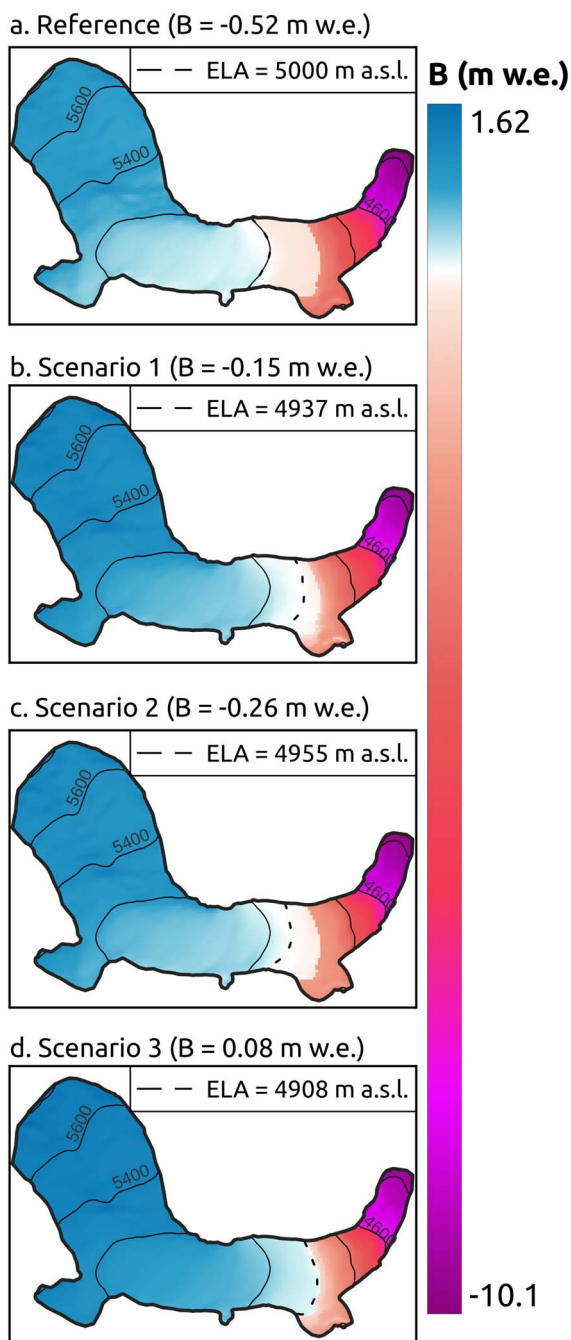
### 4.1. Surface mass balance for the LIA glacier geometry

Figure 5 shows the impact of the scenarios on the cumulative surface mass balance between September 1 and May 31. At the glacier scale, the reference scenario yields a highly negative surface mass balance (Figure 5a,  $B = -0.52$  m w.e.), scenario 1 results in close to equilibrium conditions (Figure 5b,  $B = -0.15$  m w.e.). Scenario 2 results in a slightly more negative mass balance (Figure 5c,  $B = -0.26$  m w.e.). Finally, scenario 3 results in a slightly positive surface mass balance (Figure 5d,  $B = 0.08$  m w.e.).

In scenario 3, the surface mass balance reaches 1.62 m w.e. in the upper reaches of the glacier (Figure 5d). Considering that melt can be disregarded over this part of the glacier (at about 5800 m a.s.l.), it is a reasonable value because the annual precipitation amounts received at the AWS (close to 2 m w.e. for all scenarios) which, via the precipitation gradient applied in the simulations (+10%/100 m up to 5200 m a.s.l.), implies that over 2.56 m w.e. reaches the upper part of the glacier.

At the glacier terminus, the cumulative mass loss, which ranges between 9.0 and 10.1 m w.e. depending on the scenario (Figure 5), is significantly higher than that observed in recent times, when the glacier wide mass balance has been close to zero (for example, measurements over the year 2000/2001 show  $B_{\text{Glacier-wide}} = 0.49$  m w.e. and  $B_{\text{Glacier terminus}} = -6.24$  m w.e.). This difference can be explained by the larger accumulation zone in the LIA configuration (2.50 km<sup>2</sup> versus 1.55 km<sup>2</sup> in 2000/2001) and similar annual precipitation amounts (2000/2001 is one of the wettest year in the current climate dataset) leading to a larger accumulation volume (an extra  $1.8 \times 10^6$  m<sup>3</sup>). As a result, the down-glacier ice flux is strengthened during the LIA, allowing the glacier front to reach low altitudes where melt rates are high (especially since in the lower reaches of the LIA glacier extent, the slope of the bedrock is steep and the width of the tongue is constrained by the topography, top inset in Figure 1).

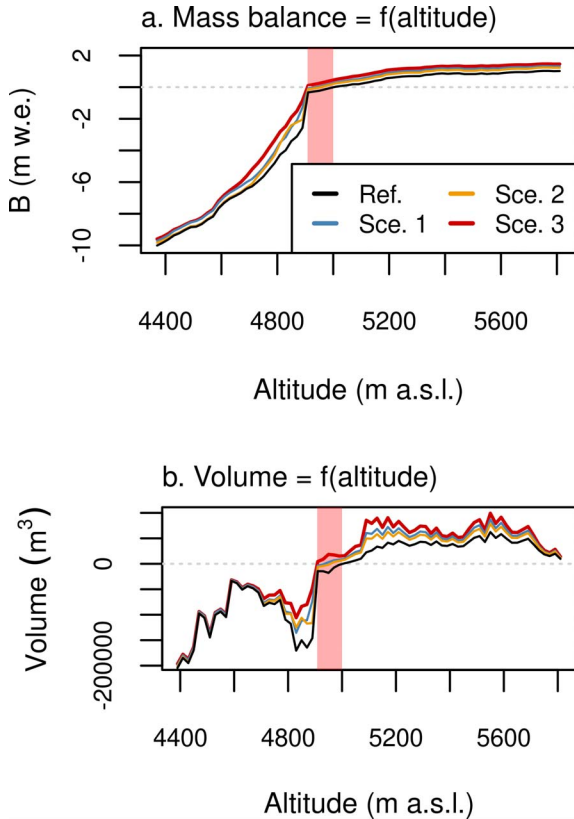
The ablation area of tropical glaciers is characterized by steep altitudinal mass balance gradients due to marked variations in albedo [Sicart *et al.*, 2011]. The simulations give similar mass balance gradients during the LIA as under the current climate (around



**Figure 5.** Annual surface mass balance maps for the four LIA scenarios. The dashed lines represent the simulated ELA.

2 m w.e./100 m, Figure 6a) probably because this gradient is mostly controlled by atmospheric variables and the LIA scenarios were built using shuffled





**Figure 6.** Surface mass balance as a function of altitude (a) and corresponding changes in mass (b) in the four scenarios. The shaded area highlights the altitudinal range for the simulated equilibrium-line altitudes (4908–5000 m a.s.l.).

present-day meteorological forcing data.

For the LIA scenarios, the specific mass turnover  $\tau$  [Equation (2), e.g., Ohmura *et al.*, 1992] was larger than for recent climatic conditions ( $\tau_{LIA}$ , all scenarios  $> 2$  m w.e. whereas  $\tau_{2000/01} = 0.98$  m w.e., Table 2) emphasizing the role of the ice flux in compensating for the melt rates in the ablation zone.

$$\tau = \frac{c + |a|}{2} \quad (2)$$

where  $\tau$  is the specific mass turnover,  $c$  the specific mass gain (in the accumulation zone) and  $a$  is the specific mass loss (in the ablation zone).

Considering a constant AAR method to define the  $ELA_0$  is not valid here as this method yields a value that is much higher than our simulations show ( $ELA_{0,LIA} = 5078$  m a.s.l. with the AAR method versus 4908 to 5000 m a.s.l. depending on the scenario). Indeed, the AAR obtained with the LIA simulations is  $\sim 0.8$ , whereas the  $AAR_0$  for the current climate is 0.68. Based on the analysis of the specific mass turnover and altitudinal mass balance gradients, this method is not valid primarily because of the topographic conditions (narrow and steep but because of the long ablation zone with a small surface area). This topographical setting coupled with the strong down-glacier ice flux enabled the glacier terminus to reach  $\sim 500$  m below the  $ELA_0$  during the LIA, as opposed to  $\sim 250$  m in current climate conditions.

Other methods exist to reconstruct past  $ELA_0$ , for example, based on the altitude of lateral moraines, Rabatel [2005] estimated the  $ELA_0$  during the LIA to be at 5090 m a.s.l. Based on our results, this altitude is overestimated by  $\sim 150$  m, suggesting that the lateral moraine method can only give a maximum ELA due to uncertainties linked to moraine accretion from the glacier fluctuations between the LIA maximum extent and now, or by moraine erosion since the glacier retreat.

Mass balance (Figure 6a) and changes in volume (Figure 6b) by altitude make it possible to assess the variability of the spatial mass balance induced by the LIA scenarios. The mass balance profiles show high variability in the ablation zone. Analysis of changes in simulated snow heights, albedo and mass balance in the ablation zone (4400–4850 m a.s.l.) showed that this variability is driven by albedo feedback effects (not shown). At the annual scale, the increased variation in albedo with altitude is due to the increased influence of snow temporarily covering the ice (from  $\pm 0.002$  at 4400 m a.s.l. to  $\pm 0.04$  at 4850 m a.s.l.). Hence, via an albedo feedback effect, the timing and intensity of the precipitation play an important role in controlling melt rates in the upper reaches of the ablation zone.

Above  $\sim 4600$  m a.s.l., scenarios 1 and 3 (Figure 6a, blue and red lines, respectively) diverge, with a less negative mass balance gradient in scenario 1 because in scenario 3, the precipitation events are more

**Table 2.** Summary of simulated glacier-wide surface mass balance  $B$ , specific ablation  $a$ , accumulation  $c$ , and mass turnover  $\tau$  in the four LIA scenarios along with the differences in  $a$  and  $c$  between the reference scenario and scenarios 1–3

Scenarios	$B$ (m w.e.)	$a$ (m w.e.)	$c$ (m w.e.)	$\tau$ (m w.e.)	$\Delta a$ (m w.e.)	$\Delta c$ (m w.e.)
Reference	−0.52	−3.31	0.96	2.13	0	0
1	−0.15	−4.28	0.90	2.59	0.97	0.06
2	−0.26	−3.78	0.93	2.35	0.47	0.03
3	0.08	−4.69	1.04	2.86	1.38	−0.07

evenly distributed, resulting in higher albedo values. Similarly, above  $\sim 4700$  m a.s.l., scenario 2 shows a shallower mass balance profile than the reference scenario (orange and black lines, respectively), because intense precipitation events throughout the wet season enable a thicker snowpack, thereby increasing the albedo and reducing annual mass loss.

Table 2 shows that the more positive the glacier-wide surface mass balance, the more negative the specific ablation and, in some cases, the lower the specific accumulation (observed in scenarios 1 and 2). Considering the mean differences between the reference scenario and in the three scenarios with changes in precipitation patterns, the variability of the specific ablation rate is higher than that of the accumulation rate (whereas  $\Delta \bar{a} = 0.94 \pm 0.46$  m w.e.  $\Delta \bar{c} = 0.01 \pm 0.07$  m w.e., Table 2) because of the effect of albedo feedback on the melt rate.

Considering the two most extreme scenarios in terms of surface mass balance value (the reference scenario and scenario 3), despite higher variability of the surface mass balance profile in the lower reaches of the glacier (below the ELA, Figure 6a), the biggest differences in volume occur in the upper reaches (above the ELA, Figure 6b), suggesting that accumulation processes play an important role in the glacier-wide surface mass balance. Compared to the reference scenario, scenario 3 generates an excess in volume of  $0.55 \times 10^6$  m<sup>3</sup> ( $1.1 \times 10^6$  m<sup>3</sup>) in the lower (upper) reaches of the glacier.

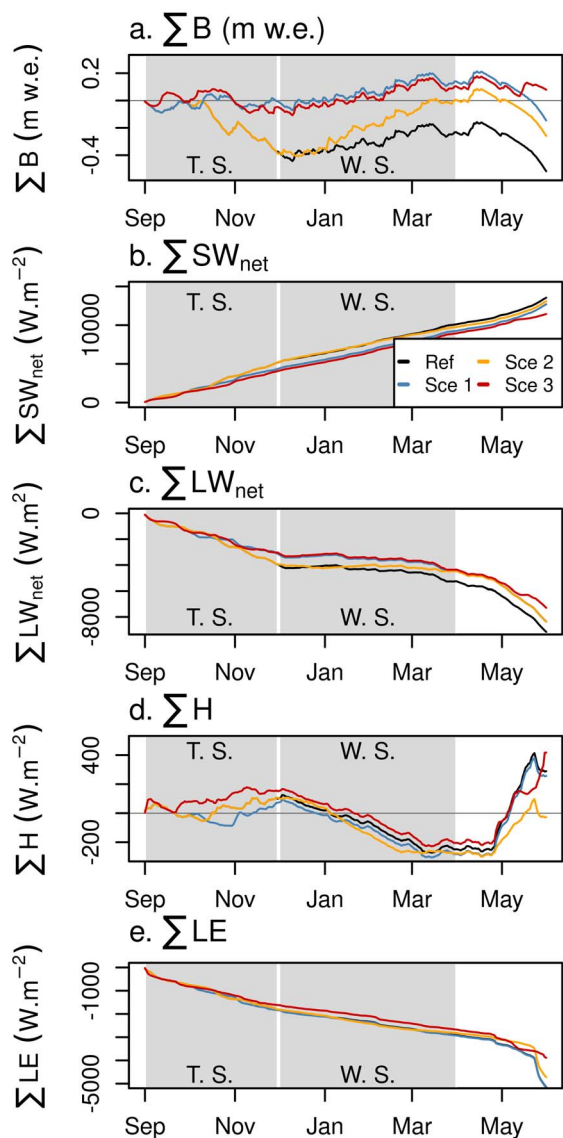
#### 4.2. Impact of the climate scenarios on the surface energy fluxes

Scenarios 1 and 3 yield close to equilibrium annual glacier-wide surface mass balances whereas the reference scenario and scenario 2 result in negative

mass balance conditions. The main differences in melt rates between these two groups occur between September and November (Figure 7a). As mentioned in Section 2.1, the transition season is characterized by increased solar radiation at the top of atmosphere and a low surface albedo, leading to large net shortwave energy fluxes that are responsible for the highest melt rates of the year, meaning this period plays a key role in controlling the annual surface mass balance [Sicart *et al.*, 2011]. As a result, the two scenarios that are the most efficient at limiting mass loss are those with increased precipitation amounts and frequencies during the transition season (scenarios 1 and 3). Indeed, clouds reduce the incoming shortwave radiation fluxes and snowfall increases the glacier albedo, thereby reducing melt energy compared to the other two scenarios (Figure 7b).

During the wet season, scenario 2 differs from scenarios 1 and 3 because it was constructed to maximize the number of large-scale precipitation events (Section 3.2.2). It reduces the amount of melt energy because the combined effect of clouds and precipitation reduces the net shortwave energy fluxes more than the clouds increase the net longwave energy budget (Figures 7b and c, respectively). This reduction, combined with the large precipitation amounts in this scenario explain the rapid mass gains observed in Figure 7a.

An early and longer wet season (scenario 1) is more efficient at reducing mass loss than a more intense wet season (scenario 2) because it reduces the incoming shortwave radiation (cloud cover) and increases the albedo (snowfall events) during the period of maximum potential solar radiation (summer solstice), which limits the mass loss over the transition season. During the core wet season, increased precipitation intensity and frequency (scenario 2) in-



**Figure 7.** Cumulative glacier-wide surface mass balance (a), net shortwave (b) and longwave (c) energy budgets along with the sensible (d) and latent (e) heat fluxes in the four LIA scenarios. The gray areas represent the transition (T.S.) and wet (W.S.) seasons.

creases the mass gain but does not compensate for the losses during the transition season (Figure 7a).

During the core wet season, the glacier-wide averaged cumulative sensible heat flux is negative in all the scenarios, showing that the temperature of the glacier surface remains higher than the temper-

ature of the air (Figure 7d). The higher temperature in the simulations is probably because in the model, the air temperature is extrapolated using a stronger altitudinal gradient than the surface temperatures ( $-0.55$  K/100 m versus  $-0.1$  K/100 m), resulting in a large part of the glacier where the simulated surface temperature is warmer than the air. However, considering the small cumulative sums of these fluxes over the year ( $-200$  and  $200$   $W \cdot m^{-2}$ ), the impact on the melt rate is limited. Finally, in all the scenarios, the latent heat fluxes act as an energy sink (sublimation) whose annual amount is about half the sum of the net longwave energy flux (Figure 7e). Because more energy is lost *via* the latent heat flux than the energy provided by the sensible heat flux, the sum of the turbulent fluxes is negative (energy sink), a feature observed under the current climate [e.g., Sicart et al., 2011].

#### 4.3. Sources of modeling errors

An important source of error is that we neglect the subsurface heat flux in the simulations, implying errors in the surface temperature. Although under the current climate, the subsurface heat flux only plays a significant role in reducing melt rates via nighttime cooling over the glacier’s accumulation zone and over the whole surface of the glacier in the dry season [e.g., Sicart et al., 2011], this heat flux may be significant during a large part of the year in the cooler LIA climate, particularly over the upper reaches of the glacier. Errors in the simulated surface temperature are partly offset by the surface temperature altitudinal gradient applied in the simulations. As the accumulation area is larger during the LIA than under the current climate ( $AAR = 0.8$  versus  $0.68$ ), the model may overestimate the glacier-wide melt rates.

The model may also underestimate the turbulent fluxes. This is due to a number of reasons including the fact that a uniform wind speed is considered across the glacier, or the fact that errors in the simulated albedo impact the type of surface simulated (snow/firn/ice) via imprecise melt rates. Overestimation of the simulated surface temperatures in the accumulation zone also influences the turbulent fluxes via a reduced surface-air temperature gradient. Another source of error is that the model considers constant stability functions throughout the year whereas

in reality, these vary with the meteorological conditions [Sicart *et al.*, 2011]. As a result, the glacier wide latent heat flux is underestimated making the turbulent fluxes a smaller energy sink than they should be, thereby contributing to overestimating the melt rates.

Other sources of errors that impact the simulations are the air temperature and precipitation gradients which are poorly known under current climate conditions (and may have been different in the LIA climate). Considering these sources of errors in the modeling, the most important ones being that the subsurface heat flux is neglected and potential erroneous albedo simulations, it is likely that the model overestimates the melt rates and hence, the annual mass losses.

## 5. Concluding remarks

This study assessed the seasonal climate changes that occurred in tropical South America during the Little Ice Age (late 17th century) using a distributed surface energy balance applied to the tropical Zongo glacier (16° S, Bolivia) and scenarios that maintain physical coherence between the measured meteorological variables.

Using information provided by paleoclimate proxies, global climate model simulations and glacier-climate sensitivity studies of past glacier extents over tropical South America, we constructed three scenarios that focus on the distribution of precipitation over time. One scenario considered a wetter year (above average precipitation in September, October, November, April and May) and the other two others tested different wet season configurations: one considered an early onset of the wet season and the other a more intensive wet season. To ensure physical coherence between the measured variables, the scenarios were obtained by shuffling months of measurements of the current climate based on 9 years of continuous hourly measurements. The three scenarios were compared to a reference scenario in which the monthly precipitation is representative of the current climate and a 1.1 K decrease in temperature. The climate scenarios were assessed based on the resulting surface mass balance of the glacier's LIA maximum extent. They were considered plausible if close to equilibrium conditions were evaluated ( $B \approx 0$  m w.e.).

The reference scenario was the one that resulted in the highest mass loss ( $B = -0.52$  m w.e.) suggesting that, by itself, 1.1 K cooling cannot explain the glacier's maximum extent during the LIA. Spreading precipitation out over the year, and more specifically, including more frequent precipitation events during the transition periods, via an albedo feedback effect, is key to reproducing a glacier that is close to equilibrium at its maximum extent during the LIA. In addition, a longer wet season (scenario 1,  $B = -0.15$  m w.e.) is more efficient in reducing mass loss than a more intense one (scenario 2,  $-0.26$  m w.e.), as the former shortens the transition period and hence limits the otherwise large mass losses that occur in the wet season. This confirms recent findings concerning the important role of the distribution of precipitation over the transition season in controlling the annual mass loss [e.g., Autin *et al.*, 2022].

Thus, the simulations indicate that the LIA climate was likely 1.1 K cooler with ~20% extra precipitation at the annual scale, in rough agreement with Rabatel *et al.* [2008]. However, our method also allowed us to investigate the effects of changes in the seasonal distribution of precipitation on the mass balance. Our results suggest that the 20% annual increase in precipitation is in agreement with the LIA glacier extent only if increased precipitation occurs in the transition season (period of highest melt rates) that precede the core wet season (December–March).

The simulations also reveal a significant change in the AAR of the glacier from around 0.68 for the current climate versus 0.80 for the LIA climate, mostly due to topographic conditions (the ablation area was steep and narrow during the LIA). This suggests that methods based on constant AAR to reconstruct the LIA climate result in overestimation of the past ELA<sub>0</sub> of about 120 m. Indeed, we hypothesize that this topographic setting coupled with the large accumulation volumes favored a strong down-glacier ice flux that allowed the glacier to extend farther below the ELA during the LIA than what is observed under the current climate.

Applying the model to other glaciers in the tropical Andes with different morpho-topographic variables would provide further insights into which LIA scenarios are most plausible and an in-depth understanding of the limits of the AAR method. Another possible further study includes coupling DEBAM to a

model of glacier dynamics to better quantify the role of the ice flux and mass turnover in reaching glacier steady-state conditions.

## Acknowledgements

This study was conducted in the framework of the International Joint Laboratory LMI GREAT-ICE, the International Research Network Andes-C2H (IRD and Andean partners) and the French Service National d'Observation GLACIOCLIM ([www.glacioclim.osug.fr](http://www.glacioclim.osug.fr), UGA-OSUG, CNRS-INSU, IRD, IPEV, INRAE). GLACIOCLIM is coordinated in La Paz by Dr. Alvaro Soruco (UMSA/IGEMA) who organizes the follow-up of the glaciological measurements on Zongo Glacier. The authors are grateful to everyone involved in the long-term *in-situ* monitoring of Zongo Glacier and acknowledge the support of LabEx OSUG@2020 (*Investissements d'Avenir*—ANR-10-LABX-0056).

## References

- Apaéstegui, J., Cruz, F. W., Vuille, M., Fohlmeister, J., Espinoza, J. C., Sifeddine, A., Strikis, N., Guyot, J. L., Ventura, R., Cheng, H., and Edwards, R. L. (2018). Precipitation changes over the eastern Bolivian Andes inferred from speleothem ( $\delta^{18}\text{O}$ ) records for the last 1400 years. *Earth Planet. Sci. Lett.*, 494, 124–134.
- Autin, P., Sicart, J. E., Rabatel, A., Soruco, A., and Hock, R. (2022). Climate controls on the inter-seasonal and interannual variability of the surface mass and energy balances of a tropical glacier (Zongo Glacier, Bolivia, 16° S): new insights from the multi-year application of a distributed energy balance. *J. Geophys. Res. Atmos.*, 127(7), article no. e2021JD035410.
- Beljaars, A. and Holtslag, A. (1991). Flux parameterization over land surface for atmospheric models. *J. Appl. Meteorol.*, 30(3), 327–341.
- Beven, K. J. (1989). Changing ideas in hydrology: the case of physically based models. *J. Hydrol.*, 105, 157–172.
- Bird, B. W., Abbott, M. B., Vuille, M., Rodbell, D. T., Stansell, N. D., and Rosenmeier, M. F. (2011). A 2300-year-long annually resolved record of the South American summer monsoon from the Peruvian Andes. *Proc. Natl Acad. Sci. USA*, 108(21), 8583–8588.
- Braconnot, P., Harrison, S. P., Otto-Bliesner, B. L., Abe-Ouchi, A., Jungclaus, J., and Peterschmitt, J.-Y. (2011). The paleoclimate modeling intercomparison project contribution to CMIP5. *CLIVAR Exchanges*, 16(56), 15–19.
- Campos, J. L., Cruz, F. W., Ambrizzi, T., Deininger, M., Vuille, M., Novello, V. F., and Strikis, N. M. (2019). Coherent South American monsoon variability during the last millennium revealed through high-resolution proxy records. *Geophys. Res. Lett.*, 46(14), 8261–8270.
- Campos, N., Alcalá-Reygosa, J., Watson, S., Koukoulos, I., Quesada Román, A., and Grima, N. (2021). Modeling the retreat of the Aneto Glacier (Spanish Pyrenees) since the Little Ice Age, and its accelerated shrinkage over recent decades. *Holocene*, 31, 1315–1326.
- De Marsily, G. (1994). Quelques réflexions sur l'utilisation des modèles en hydrologie. [Tribune libre]. *Rev. des sci. de l'eau/J. Water Sci.*, 7(3), 219–234.
- Diaz, L. B. and Vera, C. S. (2018). South American precipitation changes simulated by PMIP3/CMIP5 models during the Little Ice Age and the recent global warming period. *Int. J. Climatol.*, 38(6), 2638–2650.
- Francou, B., Ribstein, P., Saravia, R., and Tiriau, E. (1995). Monthly balance and water discharge of an inter-tropical glacier: Zongo Glacier, Cordillera Real, Bolivia, 16° S. *J. Glaciol.*, 41(137), 61–67.
- Garreaud, R., Vuille, M., and Clement, A. C. (2003). The climate of the Altiplano: observed current conditions and mechanisms of past changes. *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, 194(1–3), 5–22.
- Grove, J. M. (1988). *The Little Ice Age*. Routledge, Oxfordshire.
- Hastenrath, S. (1984). *The Glaciers of Equatorial East Africa*. D. Reidel Publishing Co., Dordrecht. Guilders 195.
- Hastenrath, S. and Ames, A. (1995). Diagnosing the imbalance of Yanamarey Glacier in the Cordillera Blanca of Peru. *J. Geophys. Res.*, 100(D3), 5105–5112.
- Haug, G. H., Hughen, K. A., Sigman, D. M., Peterson, L. C., and Röhl, U. (2001). Southward migration of the intertropical convergence zone through the Holocene. *Science*, 293(5533), 1304–1308.
- Hock, R. and Holmgren, B. (2005). A distributed surface energy-balance model for complex topogra-

- phy and its applications to Storglaciaren. *J. Glaciol.*, 51(172), 25–36.
- Hock, R. and Tijm-Reijmer, C. (2012). A mass-balance, glacier runoff and multi-layer snow model DEBAM and DETIM distributed energy balance and distributed enhanced temperature index model users manual. <http://www2.gi.alaska.edu/~regine/modelmanual.pdf>.
- Hurley, J. V., Vuille, M., and Hardy, D. R. (2019). On the interpretation of the ENSO signal embedded in the stable isotopic composition of Quelccaya Ice Cap, Peru. *J. Geophys. Res. Atmos.*, 124(1), 131–145.
- Jomelli, V., Favier, V., Rabatel, A., Brunstein, D., Hoffmann, G., and Francou, B. (2009). Fluctuations of Andean tropical glaciers since the last millennium and palaeoclimatic implication: a review. *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, 281, 269–282.
- Jomelli, V., Khodri, M., Favier, V., Brunstein, D., Ledru, M. P., Wagnon, P., Blard, P. H., Sicart, J. E., Braucher, R., Grancher, D., Bourlès, D. L., Braconnot, P., and Vuille, M. (2011). Irregular tropical glacier retreat over the Holocene epoch driven by progressive warming. *Nature*, 474(7350), 196–199.
- Kaser, G. (1999). A review of the modern fluctuations of tropical glaciers. *Glob. Planet. Change*, 22(1–4), 93–103.
- Kaser, G. (2001). Glacier-climate interaction at low latitudes. *J. Glaciol.*, 47(157), 195–204.
- Klemes, V. (1983). Conceptualization and scale in hydrology. *J. Hydrol.*, 65, 1–23.
- Lean, J. (2000). Evolution of the Sun's spectral irradiance since the Maunder Minimum. *Geophys. Res. Lett.*, 27, 2425–2428.
- Ledru, M.-P., Jomelli, V., Samaniego, P., Vuille, M., Hidalgo, S., Herrera, M., and Ceron, C. (2013). The Medieval Climate Anomaly and the Little Ice Age in the eastern Ecuadorian Andes. *Clim. Past*, 9(1), 307–321.
- Lejeune, Y. (2009). *Apports des modèles de neige CROCUS et de sol ISBA à l'étude du bilan glaciologique d'un glacier tropical et du bilan hydrologique de son bassin versant*. PhD thesis, UJF, Grenoble.
- Lenters, J. D. and Cook, K. H. (1997). On the origin of the Bolivian high and related circulation features of the South American climate. *J. Atmos. Sci.*, 54(5), 656–677.
- Lu, J. and Vecchi, G. (2014). *Hadley Circulation in Encyclopedia of the Atmospheric Sciences*. Elsevier, Amsterdam, 2nd edition.
- Malone, A. G. O., Pierrehumbert, R. T., Lowell, T. V., Kelly, M. A., and Stroup, J. S. (2015). Constraints on southern hemisphere tropical climate change during the Little Ice Age and Younger Dryas based on glacier modeling of the Quelccaya Ice Cap, Peru. *Quat. Sci. Rev.*, 125, 106–116.
- Michel, S. (2020). *Caractérisation des variations climatiques du dernier millénaire à l'aide de nouvelles méthodes statistiques: reconstruction paléoclimatiques et intégration dans un modèle de circulation globale du climat*. PhD thesis, Univ. Bordeaux.
- Morales, M. S., Christie, D. A., Villalba, R., Argollo, J., Pacajes, J., Silva, J. S., Alvarez, C. A., Llanabure, J. C., and Gamboa, C. C. (2012). Precipitation changes in the South American Altiplano since 1300 AD reconstructed by tree-rings. *Clim. Past*, 8(2), 653–666.
- Neukom, R., Gergis, J., Karoly, D. J., Wanner, H., Curran, M., Elbert, J., González-Rouco, F., Linsley, B. K., Moy, A. D., Mundo, I., Raible, C. C., Steig, E. J., Van Ommen, T., Vance, T., Villalba, R., Zinke, J., and Frank, D. (2014). Inter-hemispheric temperature variability over the past millennium. *Nat. Clim. Change*, 4(5), 362–367.
- Oerlemans, J. and Knap, W. H. (1998). A 1 year record of global radiation and albedo in the ablation zone of Morteratschgletscher, Switzerland. *J. Glaciol.*, 44(147), 231–238.
- Ohmura, A., Kasser, P., and Funk, M. (1992). Climate at the equilibrium line of glaciers. *J. Glaciol.*, 38(130), 397–411.
- Paul, F. and Bolch, T. (2019). *Glacier Changes Since the Little Ice Age*. Springer, Cham.
- Paulson, C. A. (1970). The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer. *J. Appl. Meteorol.*, 9, 857–861.
- Rabatel, A. (2005). *Chronologie et interprétation paléoclimatique des fluctuations des glaciers dans les Andes de Bolivie (16° S) depuis le maximum du Petit Âge Glaciaire (17ème siècle)*. PhD thesis, UJF, Grenoble 1.
- Rabatel, A., Bermejo, A., Loarte, E., Soruco, A., Gomez, J., Leonardini, G., Vincent, C., and Sicart, J. E. (2012). Can the snowline be used as an indicator of the equilibrium line and mass balance for glaciers in the outer tropics? *J. Glaciol.*, 58(212), 1027–1036.
- Rabatel, A., Francou, B., Jomelli, V., Naveau, P., and

- Grancher, D. (2008). A chronology of the Little Ice Age in the tropical Andes of Bolivia (16° S) and its implications for climate reconstruction. *Quat. Res.*, 70(2), 198–212.
- Rabatel, A., Francou, B., Soruco, A., Gomez, J., Cáceres, B., Ceballos, J. L., Basantes, R., Vuille, M., Sicart, J.-E., Huggel, C., Scheel, M., Lejeune, Y., Arnaud, Y., Collet, M., Condom, T., Consoli, G., Favier, V., Jomelli, V., Galarraga, R., Ginot, P., Maisincho, L., Mendoza, J., Ménégos, M., Ramirez, E., Ribstein, P., Suarez, W., Villacis, M., and Wagnon, P. (2013). Current state of glaciers in the tropical Andes: a multi-century perspective on glacier evolution and climate change. *Cryosphere*, 7(1), 81–102.
- Rabatel, A., Jomelli, V., Naveau, P., Francou, B., and Grancher, D. (2005). Dating of Little Ice Age glacier fluctuations in the tropical Andes: Charquini glaciers, Bolivia, 16 S. *C. R. Géosci.*, 337(15), 1311–1322.
- Ramallo, C. (2013). *Caractérisation du régime pluviométrique et sa relation à la fonte du Glacier Zongo (Cordillère Royale)*. PhD thesis, UGA, Grenoble.
- Ribstein, P., Tiriau, E., Francou, B., and Saravia, R. (1995). Tropical climate and glacier hydrology: a case study in Bolivia. *J. Hydrol.*, 165(1–4), 221–234.
- Sachs, J. P., Sachse, D., Smittenberg, R. H., Zhang, Z., Battisti, D. S., and Golubic, S. (2009). Southward movement of the Pacific intertropical convergence zone AD 1400–1850. *Nat. Geosci.*, 2(7), 519–525.
- Seibert, J. (1999). *Conceptual Runoff Models: Fiction or Representation of Reality?* Uppsala University Press, Uppsala, Sweden.
- Sicart, J. E. (2002). *Contribution à l'étude Des Flux d'énergie, Du Bilan de Masse et Du Débit de Fonte d'un Glacier Tropical: Le Zongo, Bolivie*. PhD thesis, Université Paris IV Pierre et Marie Curie.
- Sicart, J. E., Espinoza, J. C., Quéno, L., and Medina, M. (2016). Radiative properties of clouds over a tropical Bolivian glacier: seasonal variations and relationship with regional atmospheric circulation. *Int. J. Climatol.*, 36(8), 3116–3128.
- Sicart, J. E., Hock, R., Ribstein, P., Litt, M., and Ramirez, E. (2011). Analysis of seasonal variations in mass balance and meltwater discharge of the tropical Zongo Glacier by application of a distributed energy balance model. *J. Geophys. Res.*, 119(D13), 1–18.
- Sicart, J. E., Villacís, M., Condom, T., and Rabatel, A. (2015). GREAT ICE monitors glaciers in the tropical Andes. *Eos*, 96.
- Sicart, J. E., Wagnon, P., and Ribstein, P. (2005). Atmospheric controls of the heat balance of Zongo Glacier (16° S, Bolivia). *J. Geophys. Res. Atmos.*, 110(12), 1–17.
- Solomina, O., Bradley, R. S., Hodgson, D. A., Ivy-Ochs, S., Jomelli, V., Mackintosh, A. N., Nesje, A., Owen, L. A., Wanner, H., Wiles, G. C., and Young, N. E. (2015). Holocene glacier fluctuations. *Quat. Sci. Rev.*, 111, 34.
- Solomina, O., Bradley, R. S., Jomelli, V., Geirsdottir, A., Kaufman, D., Koch, J., Masiokas, M., Miller, G., Nesje, A., Nicolussi, K., Owen, L., Wanner, H., Wiles, G., and Yang, B. (2016). Glacier fluctuations in the last 2000 years. *Quat. Sci. Rev.*, 149, 61–90.
- Soruco, A., Vincent, C., Rabatel, A., Francou, B., Thibert, E., Sicart, J. E., and Condom, T. (2015). Contribution of glacier runoff to water resources of La Paz City, Bolivia (16° S). *Ann. Glaciol.*, 56(70), 147–154.
- Tardif, R., Hakim, G., Perkins, W., Horlick, K., Erb, M., Emile-Geay, J., Anderson, D., Steig, E., and Noone, D. (2019). Last Millennium Reanalysis with an expanded proxy database and seasonal proxy modeling. *Clim. Past*, 15, 1251–1273.
- Thompson, L. G., Mosley-Thompson, E., Brecher, H., Davis, M., León, B., Les, D., Lin, P. N., Mashiotto, T., and Mountain, K. (2006). Abrupt tropical climate change: past and present. *Proc. Natl Acad. Sci. USA*, 103(28), 10536–10543.
- Thompson, L. G., Mosley-Thompson, E., Dansgaard, W., and Grootes, P. M. (1986). The Little Ice Age as recorded in the stratigraphy of the tropical Quelccaya ice cap. *Science*, 234(4774), 361–364.
- Troll, C. (1941). *Studien zur vergleichenden Geographie der Hochgebirge der Erde, Bonner Mitt. 21*. Univ. Bonn, Bonn, Germany.
- Vimeux, F., Ginot, P., Schwikowski, M., Vuille, M., Hoffmann, G., Thompson, L. G., and Schotterer, U. (2009). Climate variability during the last 1000 years inferred from Andean ice cores: a review of methodology and recent results. *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, 281(3–4), 229–241.
- Vincent, C., Kappenberger, G., Valla, F., Bauder, A., Funk, M., and Le Meur, E. (2004). Ice ablation as evidence of climate change in the Alps over the 20th century. *J. Geophys. Res.*, 109(D10), 1–9.
- Vincent, C., Meur, E. L., Six, D., and Funk, M. (2005). Solving the paradox of the end of the Little Ice Age

- in the Alps. *J. Geophys. Res.*, 32(9), 1–4.
- Vuille, M., Burns, S. J., Taylor, B. L., Cruz, F. W., Bird, B. W., Abbott, M. B., Kanner, L. C., Cheng, H., and Novello, V. F. (2012). A review of the South American monsoon history as recorded in stable isotopic proxies over the past two millennia. *Clim. Past*, 8(4), 1309–1321.
- Wagnon, P., Ribstein, P., Francou, B., and Pouyaud, B. (1999). Annual cycle of energy balance of Zongo Glacier, Cordillera Real, Bolivia. *J. Geophys. Res.*, 104(D4), 3907–3923.





Research article

## Geo-hydrological Data & Models

# How to assess water quality change in temperate headwater catchments of western Europe under climate change: examples and perspectives

Chantal Gascuel-Oudou<sup>Ⓢ,\*</sup>, Ophélie Fovet<sup>Ⓢ,<sup>a</sup></sup>, Mikael Faucheux<sup>a</sup>, Jordy Salmon-Monviola<sup>Ⓢ,<sup>a</sup></sup> and Laurent Strohmenger<sup>Ⓢ,<sup>b</sup></sup>

<sup>a</sup> INRAE, Institut Agro, UMR SAS, 35000 Rennes, France

<sup>b</sup> INRAE, UMR HYCAR, 92160 Antony, France

*E-mails:* chantal.gascuel@inrae.fr (C. Gascuel-Oudou), ophelie.fovot@inrae.fr (O. Fovet), mikael.faucheux@inrae.fr (M. Faucheux), jordy.salmon-monviola@inrae.fr (J. Salmon-Monviola), laurent.strohmenger@inrae.fr (L. Strohmenger)

**Abstract.** Climate change effects on water quality are related directly and indirectly to the water cycle and human activities. We present examples of these effects using a retrospective analysis across European catchments according to three objectives: (i) identification of extreme or anomalous values in climatic and chemical variables at multiple time scales, (ii) assessment of variability in seasonal and inter-annual chemical cycles, and (iii) identification of a general water chemistry response to the North Atlantic Oscillation. From these examples, we highlight four challenges for research on the relationships between climate and water quality: (i) developing functional typologies of chemical elements, (ii) performing multi-temporal and multi-spatial analysis by aggregating data into water-cycle periods, (iii) decoupling effects of climate conditions and human activities by testing hypotheses using parsimonious models, and (iv) incorporating water quality and aquatic ecosystem health into integrated models. Water quality integrates current and past (legacy) conditions, flow pathways, and biogeochemical reactivity, which themselves depend on the climate. As our results highlight, water quantity and quality need to be studied together.

**Keywords.** Water chemical concentrations, Nutrients, Climate variability, Catchment, Water quality, Climate change, Long term observatory.

*Manuscript received 23 April 2022, revised 20 July 2022, accepted 21 July 2022.*

## 1. Introduction

In an overview of problems that the world's water resources would face in 2050, De Marsily [2007] wrote, "The perceived potential risks were linked to climate change consequences for water resources availability, to the increase of the World's population, to the impact of society on water ecosystems, to potentially

deteriorating drinking water quality,<sup>1</sup> and to the increasing number of mega-cities (with more than ten million inhabitants) mostly in the developing world".

Climate change influences all fluxes of the water cycle: rainfall, evapotranspiration, and discharge regimes via the frequency of extreme events, such as floods and droughts. In addition to these hydrologic impacts, surface water quality also responds to climate change [Delpla et al., 2009, Eekhout and

\* Corresponding author.

<sup>1</sup> Underlined in the original text.

de Vente, 2022, Michalak, 2016, Ryberg and Chanut, 2021, Whitehead *et al.*, 2009]. Indeed, water pollution is related directly to urban, industrial and agricultural activities of human beings, all of which are modified by, and contribute to, climate change. For example, in response to increased water and temperature stress over the centuries, human beings have continuously adapted their agricultural practices, which have subsequently led to profound changes in water quality [Merot *et al.*, 2014]. Water quality degradation is a global and growing issue that influences not only economic and social welfare but also the sustainability of ecosystems and biodiversity [Michalak, 2016]. However, there remains large uncertainty linking water quality changes, their causes, and their severity to changes in climate.

Changes in water quality in response to climate change are more challenging to address than changes in water quantity for three main reasons. First, data on water quality, chemical concentrations, and chemical fluxes are usually less common, with shorter time-series and more parameters than typical hydrologic flux and storage measurements. Second, we lack detailed understanding about feedbacks between climate and water quality [Lintern *et al.*, 2021], whereas feedbacks between climate and hydrology are better understood from global and regional hydrologic modelling. Third, it is difficult to parse the confounding effects of climate change on water quality with those of changes in human activities, particularly agricultural activities, which makes it difficult to distinguish true climate effects [Lungarska and Chakir, 2018].

These challenges in linking water quality to climate change are particularly difficult in populated temperate zones, where downstream waters are influenced by both point-source wastewater-treatment plants, industrialised agriculture and its associated diffuse inputs of nutrients and pesticides. In these regions, the intensification of agricultural activities that began in the 1960s (e.g., increased production and simplification of agricultural systems, increased energy demands, grassland and hedgerow destruction) has recently abated in response to environmental regulations after the 1980s. Despite these regulatory efforts, researchers observed persistent retention of chemical elements (especially nutrients) in the soil and groundwater and the need for long response times to mitigate these stocks in the environ-

ment [Dupas *et al.*, 2020]. These storage processes may lead to unsatisfying lag times in water quality response, especially as climate change effects on hydrology can reduce or exacerbate transport at multiple time scales (e.g. long-term trends, inter-annual, seasonal, event-scale) [Watts *et al.*, 2015a].

There are several linked effects among climate change, hydrology, and water quality. For example, reduced water availability induced by climate change is likely to decrease dilution effects during low-flow periods, thereby increasing concentrations of potentially harmful solutes. For example, modifications of soil and water temperature [Ducharne, 2008, Seyedhashemi *et al.*, 2022], soil water content, and drying and wetting cycles of catchments directly modify primary production, respiration, and decomposition, and indirectly modify water quality. In the short term, floods, which are increasing in frequency and extent under climate change [Hirabayashi *et al.*, 2013, Arnell and Gosling, 2016], can rapidly bring huge quantities of nutrients, especially in particulate form, which connect areas of the agricultural or urban landscape mosaic to hydrographic networks. On the other hand, drought, which is also increasing in frequency and duration [Prudhomme *et al.*, 2014, Spinoni *et al.*, 2018], can disrupt the biogeochemical processing functions of stream networks and even modify microbial populations and their functional abilities [Mosley, 2015]. In the long term, the increase in the atmospheric carbon dioxide concentration increases plant growth, which is an important component of nutrient balances and the water cycle (e.g. interception, transpiration) in agricultural catchments, and thus of potential losses. Indirect effects caused by progressive adaptations of agriculture can contribute to changes in water quantity and quality [Salmon-Monviola *et al.*, 2013].

Along with these basic processes, many complex interactions among climate, nutrient cycles, and stream water flows can be highlighted, even at the local scale in agricultural headwater catchments, and antagonistic effects can occur at the same time. For example, biotransformation can be accelerated by higher temperature but slowed by lower soil water content. C, N, and P exports can be accelerated in winter and slowed in summer. Climate change can modify hydrological connectivity (i.e. flow pathways, velocities, and transport capacities) in catchments, whether between fields due to land use and ecolo-

gical infrastructure, or between hydrological compartments (e.g. groundwater, riparian wetlands) [Ehrhardt *et al.*, 2021], and therefore can also modify concentration dynamics [Aubert *et al.*, 2013a, Strohmenger *et al.*, 2020]. However, modifications of the water cycle can also influence the resident time of water in these compartments, biogeochemical reactivity, and regulation of pollutant export from agricultural land. These changes in water quality can have differing effects on dissolved and particulate forms and on concentrations of a variety of chemical species, nutrients, pesticides, or emerging pollutants.

Literature on climate change effects on water quality is moderately abundant, having increased significantly since 2008. When combining queries for “water quality” and “climate change” in Web of Science, most of the articles we found were focused on aquatic ecosystem health (e.g. habitats; eutrophication; ecological effects, especially on emblematic organisms). Fewer studies focused on physical or chemical water quality. Literature reviews in this subject area [Delpla *et al.*, 2009, Giri *et al.*, 2021, Watts *et al.*, 2015b, Whitehead *et al.*, 2009] highlight the inherent complexity of interactions between climate change and water quality and the fact that they are often studied at different spatial scales [Michalak, 2016]. While climate change is studied at global and regional scales, water quality is usually studied locally. Effects of climate on water quality are analysed using three main approaches [Delahaye and Gascuel-Oudoux, 2013]: (i) experiments under controlled climate conditions [Forber *et al.*, 2017, Gu *et al.*, 2018], (ii) simulations of human activities and climate scenarios using integrated agro-hydrological models [Oni *et al.*, 2012], and (iii) retrospective analysis of long-term data time-series from research and operational observatories [Worrall *et al.*, 2015, Zhang *et al.*, 2013].

The experimental approach is well suited for soil processes but not for the catchment scale. Consequently, effects of climate change on the water cycle and river regimes have been analysed mainly by simulation modelling studies that predicted water flows under future climate-change scenarios at multiple time scales. Theoretically, this modelling approach is the most relevant, and some models have been developed for large catchments [Kaushal *et al.*, 2014, Garnier *et al.*, 2018, Rozmeijer *et al.*, 2021, Dallison *et al.*, 2022], but simulations are subject to

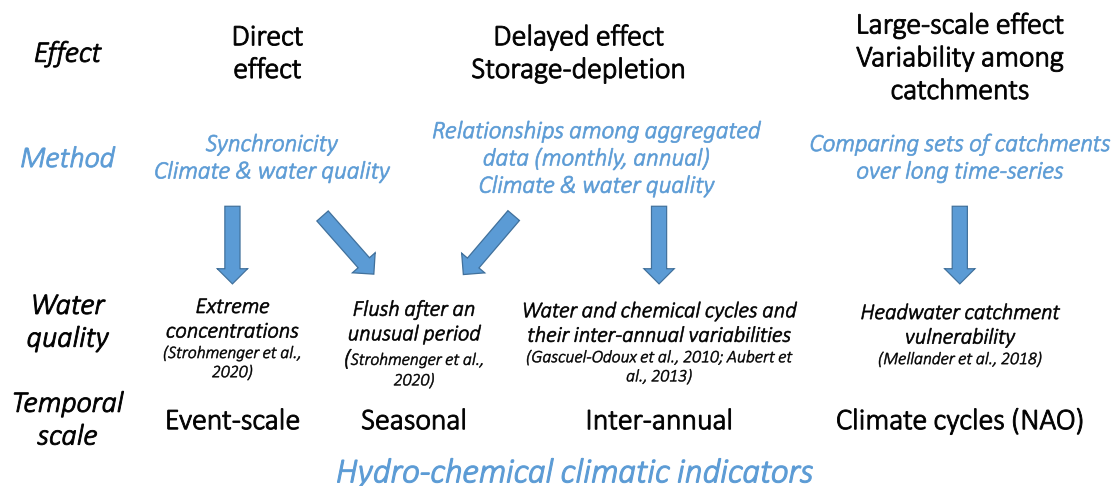
multiple sources of uncertainty [Whitehead *et al.*, 2009]: human-activity scenarios, emission scenarios used in global climate models (GCM); the GCM models themselves (e.g. due to uncertainty in understanding of processes, parameter values), downscaling uncertainty, and uncertainty in agro-hydrological models, as they are used to simulate climate conditions beyond those for which they were calibrated [Refsgaard *et al.*, 2014]. This suggests that the overall uncertainty of the modelling approach will be so large as to render its predictions questionable. For example, Salmon-Monviola *et al.* [2013] predicted that climate change would likely have the net effect of decreasing nitrate concentrations, but that the effect depended strongly on the denitrification function used in the model, and that climate change could modify this function by influencing microbial activities. Moreover, models cannot easily analyse the climate variability component of climate change due to effects that cascade from climate scenarios to water quality, water quantity and agriculture that cannot be known or represented in models. Therefore, despite limited availability of long-term data, the retrospective approach ultimately appears to be the most relevant, and the present article focuses on synthesising retrospective analyses.

The objective of this article is to illustrate effects of climate change on water quality in headwater catchments in temperate zones of western Europe, using examples of retrospective analysis of long-term data time-series from long-term research or operational observatories (Figure 1), and then to identify and highlight research perspectives for studying these relations.

## **2. Retrospective data analysis of effects of the climate on water quality**

### *2.1. Searching for synchronous climatic and chemical anomalies at multiple temporal scales*

The two following examples were based on a daily dataset from 2000–2020 from the AgrHyS French Critical Zone Observatory [Fovet *et al.*, 2018, Gascuel-Oudoux *et al.*, 2018], in the Kervidy-Naizin headwater catchment (Brittany, France), a 4.9-km<sup>2</sup> agricultural



**Figure 1.** Conceptual framework, where the examples used to illustrate effects of climate change on water quality are classified by the method, specific objective, and temporal scale, using retrospective analysis of long-term data time-series of headwater catchments in temperate zones of western Europe.

area that drains a second-Strahler-order intermittent stream with mean annual precipitation and discharge of  $810 \pm 180$  and  $296 \pm 150$  mm for 2002–2017, respectively. The catchment contains intensive mixed crop-livestock farming, with 91% of its area under agriculture, mainly crops used for animal feed, such as maize (silage or grain) and grasslands.

#### 2.1.1. Research

- At the event scale, Strohmenger *et al.* [2020] analysed effects of climate on the frequency of solute-concentration classes (C, N, and P) to capture changes in low and high concentrations over time. They hypothesised that effects of climate, particularly its highest variability, could be more apparent on extreme values. Focusing on extreme values is an original way to assess relations between climate and chemical signals in water, which is particularly relevant for storm events.
- At the seasonal scale, Strohmenger *et al.* [2020] analysed effects of climate on chemical flushing, when flow starts in autumn, as a function of the length of the dry phase in an intermittent stream. Relations between climate conditions and water quality can be revealed at the seasonal scale, particularly when climate conditions influence the hydrological regime.

#### 2.1.2. Conclusions

These two examples analysed specific temporal patterns of water quality, extreme concentrations, and hydro-climatic drivers, as well as concentrations during flushing after the summer dry phase, their timing during the water cycle, and their duration and intensity. Because the water quality signal was weak and dominated by moderate values, focusing on extreme concentrations or on atypical/extreme periods in catchment was a way to simplify it by focusing on the concentrations or periods that were influenced the most.

#### 2.2. Annual and inter-annual variabilities: cycles controlled by climate, internal storage, and depletion of this storage in a catchment

The first example is again in Kervidy-Naizin catchment, while the other two focus on a set of coastal rivers of an agricultural region (Brittany, France) with intensive mixed crop-livestock farming and crops used mainly for animal feed. Nutrient and pesticide pressure in this region has been high for decades.

##### 2.2.1. Research

- At the annual scale, Aubert *et al.* [2013b] analysed effects of climate conditions on seasonal patterns of stream-water chemistry by

plotting monthly mean in-stream concentrations of five solutes (i.e. nitrate, chloride, sulphate, and dissolved organic and inorganic C) vs. monthly mean air temperature or discharge. All solutes displayed hysteresis loops due to the time needed to rewet catchment soils, and the direction of the loop depended on the spatial control of each element. Comparison of chemical concentrations during wet vs. dry years and cold vs. hot years showed the influence of catchment drying–wetting cycles on chemical export. This method, used here to study inter-annual climate variability, could also be used to capture effects of climate change over a longer time-series.

- At the inter-annual scale, Gascuel-Oudoux *et al.* [2010] showed that annual cyclic patterns of concentrations depended on those of previous years, by plotting annual mean nitrate concentrations vs. annual mean nitrate fluxes for several rivers in Brittany (western France). Dupas *et al.* [2018] performed a similar analysis for five rivers whose data time-series were longer but had a lower temporal resolution, by plotting N loads vs. N surplus over time. In the medium term (ca. 10 years), climate variability was the main driver of concentration dynamics due to the strong legacy effect in these rivers' catchments.

### 2.2.2. Conclusions

Due to noisy daily or weekly signals in concentrations, annual and inter-annual patterns and their relations to hydro-climatic variables cannot be identified without aggregating the data (here, monthly and annual, respectively). In these two examples, the hysteresis relations observed indicate that climate influences chemical export by mobilising and depleting chemical storage in the catchment, which indicates an indirect climate effect via water flow.

### 2.3. Aggregation over time and space: effects of the North Atlantic Oscillation in north-western Europe

The previous examples were at the scales of a small catchment (Kervidy-Naizin) or region (i.e. a set of

neighbouring catchments in Brittany). Analysis of climate change at the scale of a larger region of Europe requires integrated approaches. Here, we present the development of “climate-chemical” indicators of diffuse pollution based on a large set of highly monitored agricultural catchments in western Europe, with a high diversity of agricultural intensification and environmental contexts, considered as representative of the response of water quality dynamics to hydro-climatic dynamics [Mellander *et al.*, 2018].

#### 2.3.1. Research

Continental-scale climate patterns, characterized by the intensity of the North Atlantic Oscillation (NAOi), were found to influence nutrient export to water differently among rivers in agricultural catchments in north-western Europe (i.e. Ireland, Norway and north-western France). A sharply increasing trend in NAOi, which has amplified annual weather patterns in recent years, may override positive benefits of local management depending on the catchment and year. Because these effects depend on a catchment's characteristics, agricultural history (inputs), and flow pathways, the catchment's response is specific and cannot be predicted easily.

#### 2.3.2. Conclusions

The use of the ocean-scale NAOi, whose prediction still includes many uncertainties at the local scale, and local-scale water quality data is an example of using a climate-chemical indicator in the framework of the science of integrated systems. Developing climate-chemical indicators of diffuse pollution in highly monitored catchments in western Europe is useful for increasing the predictability of their vulnerability.

## 3. Perspectives for identifying effects of climate change on water quality

Some conclusions and research perspectives can be drawn from these examples.

### 3.1. Functional typology of chemical elements

Observation, modelling, and retrospective analysis cannot be performed for all chemical parameters related to water quality. A typology needs to be created

that considers (i) the operational feasibility (e.g. cost, technical analysis) of long-term monitoring, if possible at a high temporal resolution, which is necessary to detect changes at all temporal scales; (ii) characteristics of chemical elements, particularly their affinity for soil and their sources (e.g. human inputs), and thus their flow pathways in a catchment; and (iii) research sites, to monitor specific key compartments related to hydro-climatic processes (e.g. soil, wetlands, groundwater). Because nutrients (C, N, and P) generally meet these three considerations, they could be considered as model elements for assessing effects of climate change on water quality. Nevertheless, it remains difficult to include all three of these elements (and their chemical species) simultaneously in long time-series; thus, doing so could be made a priority. We recommend monitoring suspended matter as well because it carries many elements adsorbed on soil particles and reacts strongly to storm events [Piazza *et al.*, 2018]. Pesticides, and more generally organic compounds, would be important to consider, but their data time-series likely do not allow for retrospective analysis of effects of climate variability.

### 3.2. *Multi-temporal and multi-spatial analysis: aggregating data in space and time to capture trends*

Interactions between climate and water quality must be considered across a cascading range of multiple temporal and spatial scales [Aubert *et al.*, 2015, Michalak, 2016]. The examples presented have adequately illustrated key temporal scales (e.g. storm-event, seasonal, annual, and inter-annual variabilities; long-term trends). We stress the utility of considering their timing within time-series as well. For example, water quality during storm events depends on when storms occur during the annual cycle [Aubert *et al.*, 2013c, Dupas *et al.*, 2015, Piazza *et al.*, 2018]. Water quality during a given year depends on previous inter-annual cycles. We also emphasize the utility of considering time not only on an absolute scale but on a contextual scale, by considering past dynamics from the viewpoints of climate (e.g. inter-annual phenomena such as NAO), hydrology (e.g. connectivity of flow pathways, catchment wetness), hydrochemistry (e.g. internal storage related to the history of human activity; changes in

microbial biotransformation functions in soils, particularly in wetlands) and agricultural activities. For example, the dry year of 1976, which is often considered the beginning of increasing nitrate concentrations in north-western Europe, combined a hot and dry summer with low plant growth, which caused soil N surpluses that leached during the following rainy winter and contributed greatly to internal storage of N in the groundwater and N export [Gascuel-Odoux *et al.*, 2010, Hong *et al.*, 2005, Jones and Smart, 2005].

Effects of climate change as a function of spatial scale have been illustrated less often. Mellander *et al.* [2018] clearly observed large differences in the responses of headwater catchments to climate conditions and the many drivers involved. This scale of approach highlights the utility of developing simple indicators that can be applied easily to many rivers, such as the mean nitrate concentration per month or per water year [Ebeling *et al.*, 2021]. Effects of the climate on river continuums are rarely analysed [Abbott *et al.*, 2018, Minaudo *et al.*, 2021]. From upstream to downstream, a river's internal functioning, deposits, remobilisation of deposits, consumption, and biological production can smooth the variability upstream, and climate can have a dominant influence on water quality. Moreover, the footprint and complexity of types of feedback of climate change depend on the spatial scale. Transit times in hillslopes and rivers, as well as flow pathways between soil, groundwater, and rivers, are a function of catchment size and stream order [Fovet *et al.*, 2015]. It is therefore essential to assess effects of climate change on water-quantity/quality interactions, acquire observations at multiple scales, and develop equations to describe scale effects [Pinay *et al.*, 2015].

### 3.3. *Decoupling effects of the climate and human activities: from analysis of multiple data time-series to testing of climatic scenarios with new models*

The future of water quality, especially for elements emitted by human activities depends on changes in these activities. Regardless of the changes, whether abrupt (i.e. over a few years) or slow (i.e. over a few decades), responses of water quality to them are always smoothed over time, as identified by combining data bases on these changes, usually in response to regulations (e.g. European Union's Water Framework

Directive, Common Agricultural Policy, or Strategy on Green and blue Infrastructure). When effects of human activities cannot be identified, effects of the climate are often hypothesised, whether they have been demonstrated or not. Analysis of multiple chemical elements endogenous to a catchment, such as dissolved organic C, whose dynamics depend mainly on climate conditions [Evans *et al.*, 2005], can also help identify climate effects on other chemical elements, as can analysis of chemical dynamics in key hydrological compartments. Supporting these hypotheses, however, often requires combining approaches, including modelling [e.g., Birkel *et al.*, 2014].

The models that can be used are diverse, such as simplified physical-based one-dimensional models of hillslope storage [Guillaumot *et al.*, 2021, Marcais *et al.*, 2017], multiple-box models possibly supplemented with [Strohmeinger *et al.*, 2021], and semi-distributed models (e.g. INCA, HYPE, SWAT). However, using these models beyond their calibration domains (due to climate change) often raises questions [Nicolle *et al.*, 2021, Refsgaard *et al.*, 2014]. Few models have considered changes in the atmospheric carbon dioxide concentration, which has a strong influence on agroecosystems and thus chemical export (by vegetation). Thus, to test hypotheses, attention must be paid to the processes simulated, calibration/evaluation strategies in changing contexts [Wagener *et al.*, 2022], and the flexibility of the model itself, which influences the ability to simplify it or make it more complex. To improve consideration of scales in changing contexts, physical realism of models is important, without being over-parameterized; the most important characteristic is to remain consistent with agronomic, climatic, and hydrological hypotheses [Kirchner, 2006].

### 3.4. *Incorporating water quality and aquatic ecosystem health: analysing changes in stoichiometry, habitat, and abiotic and biotic conditions*

Changes in water quality are often analysed based on one or a few chemical elements but not as a whole, which could lead to negative consequences on habitats, aquatic organisms, and ecosystems. The French scientific expertise on eutrophication [Pinay *et al.*, 2017], after reviewing the literature, highlighted an increasing risk of eutrophication as one effect of

climate change, due to cascading effects. Changes in stoichiometric C:N:P ratios have been rarely studied [Pinay *et al.*, 2017, Le Moal *et al.*, 2019]. New methods are needed to analyse element concentrations and patterns, compare them, and analyse loads and element ratios, which could have effects on organisms. Many studies have focused on climate effects on organisms without explicitly relating them to abiotic parameters (i.e. physical and chemical water quality) or their impacts on habitats.

## 4. Conclusion

Despite difficulties in identifying effects of climate change on water quality at the multiple temporal and spatial scales at which they operate, retrospective studies have highlighted that these effects can be numerous, as demonstrated by observation of long-term time-series. These effects on water quality vary greatly among storm-event patterns (e.g. extreme values, especially for elements associated with particles; concentration distributions), seasonal patterns (e.g. change in seasonal timing, range of variation in summer and winter, flushing after intermittent flow), annual cycles (C/Q or C/T annual hysteresis), and inter-annual cycles (C/Q or C/T inter-annual hysteresis due to hydro-climatic phenomena). These effects can be characterised by hydro-chemical indicators.

These effects need to be investigated in more detail to characterize them and their potential interactions with mitigation measures to restore water quality. Time must be considered explicitly due to chemical legacies in catchments. Long-term observation at higher temporal resolution to capture as many extreme events as possible is one recommendation for research observatories and for river monitoring for the Water Framework Directive. Due to the noise and diversity of water quality data, temporal aggregation over functional periods and spatial aggregation from multiple headwater catchments to river continuums is necessary to demonstrate certain effects of climate on water quality. Modelling is facing new challenges, in which the objective is not to represent current activities but to represent and test the chemical legacies in catchments and their mobilization due to climate conditions, variabilities, and trends.

Society increasingly focuses on the quantity of water (e.g. groundwater, lake and river water, reservoirs) with regard to demographics (e.g. urbanization, migration, drinking water), agriculture, and food security. It is essential, however, to combine the analysis of water quantity and quality. Doing so would require strengthening relations among scientific communities in integrated ecosystem approaches and investigating wider topics (e.g. the water-energy-agriculture nexus, food), particularly when addressing water quality in the context of the ecological transition of societies and adaptation to climate change. Transforming food systems, water and energy consumption in agriculture, and agricultural systems according to agroecological principles, particularly by decreasing agricultural inputs, can help mitigate degradation of water quality, now and under the future climate.

### Conflicts of interest

Authors have no conflict of interest to declare.

### Acknowledgments

We thank all those who help monitor water quality in the headwater catchments mentioned. We also thank Michelle and Michael Corson for English copyediting and proofreading of the article.

### References

- Abbott, B. W., Gruau, G., Zarnetske, J. P., Moatar, F., Barbe, L., Thomas, Z., Fovet, O., Kolbe, T., Gu, S., Pierson-Wickmann, A. C., Davy, P., and Pinay, G. (2018). Unexpected spatial stability of water chemistry in headwater stream networks. *Ecol. Lett.*, 21(2), 296–308.
- Arnell, N. W. and Gosling, S. N. (2016). The impacts of climate change on river flood risk at the global scale. *Clim. Chang.*, 134(3), 387–401.
- Aubert, A. H., Gascuel-Odoux, C., Gruau, G., Akkal, N., Faucheux, M., Fauvel, Y., Grimaldi, C., Hamon, Y., Jaffrézic, A., Lecoq-Boutnik, M., Molénat, J., Petitjean, P., Ruiz, L., and Merot, P. (2013a). Solute transport dynamics in small, shallow groundwater-dominated agricultural catchments: insights from a high-frequency, multisolute 10 yr-long monitoring study. *Hydrol. Earth Syst. Sci.*, 17(4), 1379–1391.
- Aubert, A. H., Gascuel-Odoux, C., and Merot, P. (2013b). Annual hysteresis of water quality: A method to analyse the effect of intra- and inter-annual climatic conditions. *J. Hydrol.*, 478, 29–39.
- Aubert, A. H., Merot, P., and Gascuel-Odoux, C. (2015). Temporal patterns of water quality: decadal high-frequency data-driven analysis from an hydrological observatory under agricultural land-use. *Houille Blanche-Revue Internationale de l'Eau*, 6, 5–11.
- Aubert, A. H., Tavenard, R., Emonet, R., de Lavenne, A., Malinowski, S., Guyet, T., Quiniou, R., Odobez, J. M., Merot, P., and Gascuel-Odoux, C. (2013c). Clustering flood events from water quality time series using Latent Dirichlet Allocation model. *Water Resour. Res.*, 49(12), 8187–8199.
- Birkel, C., Soulsby, C., and Tetzlaff, D. (2014). Integrating parsimonious models of hydrological connectivity and soil biogeochemistry to simulate stream DOC dynamics. *J. Geophys. Res. Biogeosci.*, 119, 1030–1047.
- Dallison, R. J. H., Williams, A. P., Harris, I. M., and Patil, S. D. (2022). Modelling the impact of future climate change on streamflow and water quality in Wales, UK. *Hydrol. Sci. J.*, 67(6), 939–962.
- De Marsily, G. (2007). An overview of the world's water resources problems in 2050. *Ecohydrol. Hydrobiol.*, 7(2), 147–155.
- Delahaye, D. and Gascuel-Odoux, C. (2013). Écosystèmes continentaux aquatiques et terrestres. In Merot, P., Dubreuil, V., et al., editors, *Changement climatique dans l'Ouest. Evaluation, Impacts, Perceptions*, pages 179–182. Presses Universitaires de Rennes, Rennes.
- Delpla, I., Jung, A. V., Thomas, O., Baures, E., Clement, M., and Thomas, O. (2009). Impacts of climate change on surface water quality in relation to drinking water production. *Environ. Int.*, 35(8), 1225–1233.
- Ducharne, A. (2008). Importance of stream temperature to climate change impact on water quality. *Hydrol. Earth Syst. Sci.*, 12, 797–810.
- Dupas, R., Ehrhardt, S., Musolff, A., Fovet, O., and Durand, P. (2020). Long-term nitrogen retention and transit time distribution in agricultural catchments in western France. *Environ. Res. Lett.*, 15(11), article no. 115011.
- Dupas, R., Minaudo, C., Gruau, G., Ruiz, L., and



- Gascuel-Oudoux, C. (2018). Multidecadal trajectory of riverine nitrogen and phosphorus dynamics in rural catchments. *Water Resour. Res.*, 54(8), 5327–5340.
- Dupas, R., Tavenard, R., Fovet, O., Gilliet, N., Grimaldi, C., and Gascuel-Oudoux, C. (2015). Identifying seasonal patterns of phosphorus storm dynamics with dynamic time warping. *Water Resour. Res.*, 51(11), 8868–8882.
- Ebeling, P., Dupas, R., Abbott, B., Kumar, R., Ehrhardt, S., Fleckenstein, J. H., and Musolff, A. (2021). Long-term nitrate trajectories vary by season in western European catchments. *Global Biogeochem. Cycles*, 35(9), article no. e2021GB007050.
- Eekhout, J. P. C. and de Vente, J. (2022). Global impact of climate change on soil erosion and potential for adaptation through soil conservation. *Earth Sci. Rev.*, 226, article no. 103921.
- Ehrhardt, S., Ebeling, P., Dupas, R., Kumar, R., Fleckenstein, J. H., and Musolff, A. (2021). Nitrate transport and retention in western European catchments are shaped by hydroclimate and subsurface properties. *Water Resour. Res.*, 57(10), article no. e2020WR029469.
- Evans, C. D., Monteith, D. T., and Cooper, D. M. (2005). Long-term increases in surface water dissolved organic carbon: Observations, possible causes and environmental impacts. *Environ. Pollut.*, 37(1), 55–71.
- Forber, K. J., Ockenden, M. C., Wearing, C., Hollaway, M. J., Falloon, P. D., Kahana, R., Villamizar, M. L., Zhou, J. G., Withers, P. J. A., Beven, K. J., Collins, A. L., Evans, R., Hiscock, K. M., Macleod, C. J. A., and Haygarth, P. M. (2017). Determining the effect of drying time on phosphorus solubilization from three agricultural soils under climate change scenarios. *J. Environ. Qual.*, 46(5), 1131–1136.
- Fovet, O., Ruiz, L., Faucheux, M., Molenat, J., Sekhar, M., Vertes, F., Aquilina, L., Gascuel-Oudoux, C., and Durand, P. (2015). Using long time series of agricultural-derived nitrates for estimating catchment transit times. *J. Hydrol.*, 522, 603–617.
- Fovet, O., Ruiz, L., Gruau, G., Akkal, N., Aquilina, L., Busnot, S., Dupas, R., Durand, P., Faucheux, M., Fauvel, Y., Flechard, C., Gilliet, N., Grimaldi, C., Hamon, Y., Jaffrezic, A., Jeanneau, L., Labasque, T., Le Henaff, G., Merot, P., Molenat, J., Petitjean, P., Pierson-Wickmann, A. C., Squividant, H., Vi-  
aud, V., Walter, C., and Gascuel-Oudoux, C. (2018). AgrHyS: an observatory of response times in agro-hydro systems. *Vadose Zone J.*, 17(1), 1–16.
- Garnier, J., Ramarson, A., Billen, G., Thiéry, S., Thiéry, D., Thieu, V., Minaudo, C., and Moatar, F. (2018). Nutrient inputs and hydrology together determine biogeochemical status of the Loire River (France): Current situation and possible future scenarios. *Sci. Total Environ.*, 637, 609–624.
- Gascuel-Oudoux, C., Auroousseau, P., Durand, P., Ruiz, L., and Molenat, J. (2010). The role of climate on inter-annual variation in stream nitrate fluxes and concentrations. *Sci. Total Environ.*, 408, 5657–5666.
- Gascuel-Oudoux, C., Fovet, O., Gruau, G., Ruiz, L., and Merot, P. (2018). Evolution of scientific questions over 50 years in the Kervidy-Naizin catchment: From catchment hydrology to integrated studies of biogeochemical cycles and agroecosystems in a rural landscape. *Cuad. Invest. Geogr.*, 44(2), 535–555.
- Giri, S., Mishra, A., Zhang, Z., Lathrop, R. G., and Al-nahit, A. O. (2021). Meteorological and hydrological drought analysis and its impact on water quality and stream integrity. *Sustainability*, 13(15), article no. 8175.
- Gu, S., Gruau, G., Malique, F., Dupas, R., Petitjean, P., and Gascuel-Oudoux, C. (2018). Drying/rewetting cycles stimulate release of colloidal-bound phosphorus in riparian soils. *Geoderma*, 321, 32–41.
- Guillaumot, L., Marçais, J., Vautier, C., Guillou, A., Vergnaud, V., Bouchez, C., Dupas, R., Durand, P., de Dreuzy, J. R., and Aquilina, L. (2021). A hillslope-scale aquifer-model to determine past agricultural legacy and future nitrate concentrations in rivers. *Sci. Total Environ.*, 800, article no. 149216.
- Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H., and Kanae, S. (2013). Global flood risk under climate change. *Nat. Clim. Chang.*, 3(9), 816–821.
- Hong, B. G., Swaney, D. P., Woodbury, P. B., and Weinstein, D. A. (2005). Long-term nitrate export pattern from Hubbard Brook watershed 6 driven by climatic variation. *Water Air Soil Pollut.*, 160, 293–326.
- Jones, A. L. and Smart, P. L. (2005). Spatial and temporal changes in the structure of groundwater nitrate concentration time series (1935–1999) as demonstrated by autogressive modelling. *J. Hy-*

- drol.*, 310, 201–215.
- Kaushal, S. S., Mayer, P. M., Vidon, P. G., Smith, R. M., Pennino, M. J., Newcomer, T. A., Shuiwang, D., Welty, C., and Belt, K. T. (2014). Land use and climate variability amplify carbon, nutrient, and contaminant pulses: A review with management implications. *J. Am. Water Resour. Assoc.*, 50(3), 585–614.
- Kirchner, J. W. (2006). Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resour. Res.*, 42(3), article no. W03S04.
- Le Moal, M., Gascuel-Odoux, C., Menesguen, A., Souchon, Y., Etrillard, C., Levain, A., Moatar, F., Pannard, A., Souchu, P., Lefebvre, A., and Pinay, G. (2019). Eutrophication: A new wine in an old bottle? *Sci. Total Environ.*, 651, 1–11.
- Lintern, A., Liu, S. C., Minaudo, C., Dupas, R., Guo, D. L., Bende-Michl, U., and Duvert, C. (2021). The influence of climate on water chemistry states and dynamics in rivers across Australia. *Hydrol. Process.*, 35(12), article no. e14423.
- Lungarska, A. and Chakir, R. (2018). Climate-induced land use change in France: Impacts of agricultural adaptation and climate change mitigation. *Ecol. Econ.*, 147, 134–154.
- Marcais, J., de Dreuzy, J. R., and Erhel, J. (2017). Dynamic coupling of subsurface and seepage flows solved within a regularized partition formulation. *Adv. Water Resour.*, 109, 94–105.
- Mellander, P. E., Jordan, P., Bechmann, M., Fovet, O., Shore, M. M., McDonald, N. T., and Gascuel-Odoux, C. (2018). Integrated climate-chemical indicators of diffuse pollution from land to water. *Sci. Rep.*, 8, article no. 944.
- Merot, P., Corgne, S., Delahaye, D., Desnos, P., Dubreuil, V., Gascuel, C., Giteau, J. L., Joannon, A., Quenol, H., and Narcy, J. B. (2014). Assessment, impact and perception of climate change in the western part of France: The CLIMASTER project. *Cah. Agric.*, 23(2), 96–107.
- Michalak, A. M. (2016). Study role of climate change in extreme threats to water quality. *Nature*, 535(7612), 349–350.
- Minaudo, C., Abonyi, A., Leitao, M., Lancon, A. M., Flourey, M., Descy, J. P., and Moatar, F. (2021). Long-term impacts of nutrient control, climate change, and invasive clams on phytoplankton and cyanobacteria biomass in a large temperate river. *Sci. Total Environ.*, 756, article no. 144074.
- Mosley, L. M. (2015). Drought impacts on the water quality of freshwater systems; review and integration. *Earth Sci. Rev.*, 140, 203–214.
- Nicolle, P., Andréassian, V., Royer-Gaspard, P., Perrin, C., Thirel, G., Coron, L., and Santos, L. (2021). RAT—a robustness assessment test for calibrated and uncalibrated hydrological models. *Hydrol. Earth Syst. Sci.*, 25(9), 5013–5027.
- Oni, S. K., Futter, M. N., Molot, L. A., and Dillon, P. J. (2012). Modelling the long term impact of climate change on the carbon budget of Lake Simcoe, Ontario using INCA-C. *Sci. Total Environ.*, 414, 387–403.
- Piazza, G. A., Dupas, R., Gascuel, C., Grimaldi, C., Pinheiro, A., and Kaufmann, V. (2018). Influence of hydroclimatic variations on solute concentration dynamics in nested subtropical catchments with heterogeneous landscapes. *Sci. Total Environ.*, 635, 1091–1101.
- Pinay, G., Gascuel, C., Ménesguen, A., Souchon, Y., Le Moal, M., Levain, A., Moatar, F., Pannard, A., and Souchu, P. (2017). *Eutrophication: Manifestations, Causes, Consequences and Predictability*. Synthesis of the Joint Scientific Appraisal. CNRS - Ifremer - INRA - Irstea, France.
- Pinay, G., Peiffer, S., De Dreuzy, J. R., Krause, S., Hannah, D. M., Fleckenstein, J. H., Sebilo, M., Bishop, K., and Hubert-Moy, L. (2015). Upscaling nitrogen removal capacity from local hotspots to low stream orders' drainage basins. *Ecosystems*, 18(6), 1101–1120.
- Prudhomme, C., Giuntoli, I., Robinson, E. L., Clark, D. B., Arnell, N. W., Dankers, R., Fekete, B. M., Franssen, W., Gerten, D., Gosling, S. N., Hagemann, S., Hannah, D. M., Kim, H., Masaki, Y., Satoh, Y., Stacke, T., Wada, Y., and Wisser, D. (2014). Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment. *Proc. Natl. Acad. Sci. USA*, 111(9), 3262–3267.
- Refsgaard, J. C., Madsen, H., Andréassian, V., Arnbjerg-Nielsen, K., Davidson, T. A., Drews, M., Hamilton, D. P., Jeppesen, E., Kjellstrom, E., Olesen, J. E., Sonnenborg, T. O., Trolle, D., Willems, P., and Christensen, J. H. (2014). A framework for testing the ability of models to project climate change and its impacts. *Clim. Chang.*, 122, 271–282.

- Rozmeijer, J., Noordhuis, R., Ouwerkerk, K., Pires, M. D., Blauw, A., Hooijboerb, A., and Oldenborgh, G. J. (2021). Climate variability effects on eutrophication of groundwater, lakes, rivers, and coastal waters in the Netherlands. *Sci. Total Environ.*, 771, article no. 145366.
- Ryberg, K. R. and Chanat, J. G. (2021). Climate extremes as drivers of surface-water-quality trends in the United States. *Sci. Total Environ.*, 809, article no. 152165.
- Salmon-Monviola, J., Moreau, P., Benhamou, C., Durand, P., Merot, P., Oehler, F., and Gascuel-Odoux, C. (2013). Effect of climate change and increased atmospheric CO<sub>2</sub> on hydrological and nitrogen cycling in an intensive agricultural headwater catchment in western France. *Clim. Chang.*, 120(1–2), 433–447.
- Seyedhashemi, H., Vidal, J.-P., Diamond, J. S., Thiéry, D., Monteil, C., Hendrickx, F., Maire, A., and Moatar, F. (2022). Regional, multi-decadal analysis on the Loire River basin reveals that stream temperature increases faster than air temperature. *Hydrol. Earth Syst. Sci.*, 26(9), 2583–2603.
- Spinoni, J., Vogt, J. V., Naumann, G., Barbosa, P., and Dosio, A. (2018). Will drought events become more frequent and severe in Europe? *Int. J. Climatol.*, 38(4), 1718–1736.
- Strohmenger, L., Fovet, O., Akkal-Corfini, N., Dupas, R., Durand, P., Faucheux, M., Gruau, G., Hamon, Y., Jaffrezic, A., Minaudo, C., Petitjean, P., and Gascuel-Odoux, C. (2020). Multitemporal relationships between the hydroclimate and exports of carbon, nitrogen, and phosphorus in a small agricultural watershed. *Water Resour. Res.*, 56(7), article no. e2019WR026323.
- Strohmenger, L., Fovet, O., Hrachowitz, M., Salmon-Monviola, J., and Gascuel-Odoux, C. (2021). Is a simple model based on two mixing reservoirs able to reproduce the intra-annual dynamics of DOC and NO<sub>3</sub> stream concentrations in an agricultural headwater catchment? *Sci. Total Environ.*, 794, article no. 148715.
- Wagener, T., Reinecke, R., and Pianosi, F. (2022). On the evaluation of climate change impact models. *WIREs Clim. Change*, 13(3), article no. e772.
- Watts, G., Battarbee, R. W., Bloomfield, J. P., Crossman, J., Daccache, A., Durance, I., Garner, G., Hanford, J., Hannah, D. M., Hess, T., Jackson, C. R., Kernan, M., Knox, J., Mackay, J., Monteith, D. T., Ormerod, S. J., Rance, J., Wade, A. J., Wade, S. D., Weatherhead, K., Whitehead, P. G., and Wilby, R. L. (2015a). Climate change and water in the UK - past changes and future prospects. *Prog. Phys. Geogr.*, 39(1), 6–28.
- Watts, G., Hannah, D. M., and Watkinson, A. (2015b). Introduction to the special issue on the impact of climate change on water in the UK. *Prog. Phys. Geogr.*, 39(1), 3–5.
- Whitehead, P. G., Wilby, R. L., Battarbee, R. W., Kernan, M., and Wade, A. J. (2009). A review of the potential impacts of climate change on surface water quality. *Hydrol. Sci. J.*, 54(1), 101–123.
- Worrall, F., Howden, N. J. K., and Burt, T. P. (2015). Time series analysis of the world's longest fluvial nitrate record: evidence for changing states of catchment saturation. *Hydrol. Process.*, 29(3), 434–444.
- Zhang, Q., Brady, D. C., and Ball, B. P. (2013). Long-term seasonal trends of nitrogen, phosphorus, and suspended sediment load from the non-tidal Susquehanna River Basin to Chesapeake Bay. *Sci. Total Environ.*, 452–453, 208–221.





Research article

## Geo-hydrological Data & Models

# Future climate or land use? Attribution of changes in surface runoff in a typical Sahelian landscape

Roland Yonaba<sup>\*,a</sup>, Lawani Adjadi Mounirou<sup>\*,a</sup>, Fowé Tazen<sup>\*,a</sup>, Mahamadou Koïta<sup>\*,a</sup>,  
Angelbert Chabi Biauou<sup>\*,a</sup>, Cheick Oumar Zouré<sup>\*,a</sup>, Pierre Queloz<sup>\*,b</sup>, Harouna  
Karambiri<sup>\*,a</sup> and Hamma Yacouba<sup>a</sup>

<sup>a</sup> Laboratoire Eau, Hydro-Systèmes et Agriculture (LEHSA), Institut International d'Ingénierie de l'Eau et de l'Environnement (2iE), 01 BP 594 Ouagadougou 01, Burkina Faso

<sup>b</sup> Institute of Territorial Engineering (INSIT), School of Management and Engineering Vaud (HEIG-VD), Yverdon-les-Bains, Switzerland

*E-mails:* ousmane.yonaba@2ie-edu.org (R. Yonaba), adjadi.mounirou@2ie-edu.org (L. A. Mounirou), tazen.fowe@2ie-edu.org (F. Tazen), mahamadou.koita@2ie-edu.org (M. Koïta), angelbert.biauou@2ie-edu.org (A. C. Biauou), cheickoumar.zoure@gmail.com (C. O. Zouré), pierre.queloz@gmail.com (P. Queloz), harouna.karambiri@2ie-edu.org (H. Karambiri), hamma.yacouba@2ie-edu.org (H. Yacouba)

**Abstract.** In this study, the Soil and Water Assessment Tool (SWAT) model is used to assess changes in surface runoff between the baseline (1995–2014) and future (2031–2050) periods in the Tougou watershed (37 km<sup>2</sup>) in Burkina Faso. The study uses a combination of land use maps (for current and future periods) and a bias-corrected ensemble of 9 CMIP6 climate models, under two warming scenarios. An increase in rainfall (13.7% to 18.8%) is projected, which is the major contributor to the increase in surface runoff (24.2% to 34.3%). The land use change narrative (i.e. conversion of bare areas to croplands) is expected to decrease in surface runoff, albeit minor in comparison to the effect of future climate change. Similar findings are observed for annual maximum surface runoff. This study sheds light on the need to consider simultaneously future climate and land use in framing water management policies.

**Keywords.** Climate change, Hydrological modelling, Land use/land cover change, Sahel, Surface runoff, SWAT, Tougou watershed.

*Manuscript received 29 March 2022, revised 12 August 2022 and 30 September 2022, accepted 25 October 2022.*

## 1. Introduction

In the hydrological cycle, the factors driving catchment-scale processes are mainly climate and land use [de Marsily, 2007]. Here, “*land use*” both includes land cover, land use and land management [Yin et al., 2017]. Climate change and variability are

primarily driven by large scale conditions and result in global and regional substantial changes [Bhagat et al., 2022, Connors et al., 2022, IPCC, 2022]. Locally mitigating the harming impacts of climate change on the water cycle could be fostered by informed land use and water management at the community level [Zipper et al., 2018].

The issue is more acute to the West African Sahel, which is known to be an ecoclimatic context

\* Corresponding author.

strongly affected by intense droughts, severe rainfall deficits and increased rainfall variability driven by abrupt climate oscillations [Biasutti, 2019]. The region also features the highest populations exerting strong pressure on natural resources (land clearing for cropping, ecosystem services). Besides, the capacity of populations living in these regions to adapt to climate change and variability impacts remains low [Serdeczny *et al.*, 2017]. In such drought-prone contexts, improving local resilience through integrated water management should both consider land use decisions, water use and allocation [Belemtougri *et al.*, 2021, Fovet *et al.*, 2021, Kafando *et al.*, 2021, Zipper *et al.*, 2018]. This, in turn, requires, at the core, a deeper understanding of the individual and combined contributions of climate and land use changes on the hydrological cycle, which is generally not well assessed, especially for such West African watersheds [Aich *et al.*, 2015, Yira *et al.*, 2017, Yonaba, 2020, Gbohoui *et al.*, 2021, 2022].

The complexity of processes in the water cycle in Sahelian watersheds remains a major challenge to hydrologists to date, for various reasons: first, rainfall conditions in this context are specifically defined by high intensities rainfall of short durations [Biasutti, 2019], which quickly onsets a highly non-linear Hortonian surface runoff response [Mounirou *et al.*, 2020]. Second, surface runoff generation mechanisms are typically dependent on both antecedent rainfall but also soil surface conditions, the latter being prone to surface crusting and sealing; which, in turn, affects other key processes such as infiltration, and evaporation [Valentin, 2018, Zouré *et al.*, 2019]. Third, long-term and spatially complete sets of rainfall-runoff observations are scarce and rarely gap-free, since monitoring networks are lacking [Mahé and Paturel, 2009]. These conditions severely hamper hydrological modelling applications in these contexts, which are nevertheless the preferred approach to assess the combined and individual contributions of climate and land use dynamics on the water cycle [Yin *et al.*, 2017].

To the best of our knowledge, few modelling studies in the Sahel have been dedicated to this exercise still [Aich *et al.*, 2015, Akoko *et al.*, 2021, Angelina *et al.*, 2015, Dembélé *et al.*, 2022, Gal *et al.*, 2017, Grippa *et al.*, 2017, Karambiri *et al.*, 2011, Séguis *et al.*, 2004]. Among the reasons often put forward, some authors mention the difficulty in representing

accurately Sahelian hydrological processes in most of the available models [Cornelissen *et al.*, 2013, Karambiri *et al.*, 2003], or the data scarcity [Mahé and Paturel, 2009]. Another difficulty in achieving accuracy in simulating the water balance in hydrological models resides in a full acquaintance of land use/land cover changes (LULCC) during long-term runs in modelling experiments. Usually, a single and static LULC map of the landscape at a given time point is used, which might lead to failure in picturing the spatial and temporal patterns of evolution of hydrological processes [Wagner *et al.*, 2016, Yonaba *et al.*, 2021a]. Some hydrological models, such as the Soil and Water Assessment Tool [Arnold *et al.*, 1998] provide mechanisms to integrate a dynamical update of LULC conditions during simulations, through a dedicated land use update (LUP) module, which reportedly helped in achieving significant higher model performance [Aghsaei *et al.*, 2020, Yonaba *et al.*, 2021a].

The question of which of the two factors between climate and land use changes is mainly responsible for the alteration of hydrological processes is yet to be answered in the case of Sahelian watersheds. The most often occurring landscape trajectory in this region is the loss of natural vegetation, which has been primarily associated with the increase in surface runoff [Amogu *et al.*, 2015, Gal *et al.*, 2017, Yonaba, 2020]. Yet, some studies highlight that the attribution of the observed changes in surface runoff remains unclear, probably without a single definitive answer across all contexts [Aich *et al.*, 2015, Descroix *et al.*, 2018]. Regarding climate change, most future climate projections converge for warming in average temperatures and an increase in potential evapotranspiration over the Sahel [Diedhiou *et al.*, 2018], whereas for rainfall, forecasts remain mixed [Dosio *et al.*, 2020]. Such perspectives are leaving large uncertainties about the overall impact of climate or LULC change on the hydrological cycle [Stanzel *et al.*, 2018, Sylla *et al.*, 2018, Todzo *et al.*, 2020].

To meet this challenge, in this study, we introduce a methodological framework to assess the combined and relative contribution of future climate and LULC changes on the hydrological cycle, with a focus on surface runoff, in a typical Sahelian landscape under semi-arid climate. The Tougou watershed (37 km<sup>2</sup>), located in northern Burkina Faso, has been an observatory for the study of interrela-

tions between climate, environment and population dynamics for more than 5 decades [Yonaba, 2020] and is selected as a support for this research. In a previous study, LULC maps of the Tougou watershed have been produced using remote sensing for the years 1999, 2009 and 2017; the landscape evolution has also been projected to 2030, 2040 and 2050, through land change modelling, under the hypothesis of business as usual trends [Yonaba *et al.*, 2021b]. Also, the SWAT model has been calibrated based on rainfall-runoff observations between 2000 and 2017, using dynamics LULCs maps available for this period [Yonaba *et al.*, 2021a]. Building upon these results, the aim of this study is twofold: (i) to forecast the projected changes in the hydrological balance using future climate and land use; (ii) to quantify the combined and individual contribution of climate and land use to average annual and daily extreme surface runoff.

## 2. Materials and methods

### 2.1. Study area

#### 2.1.1. Location and physical setting of Tougou watershed

The study is carried out in the Tougou watershed (13.65° N; 2.26° E), which lies in north-eastern Burkina Faso, in the upper Nakanbe River Basin, within the Yatenga province. The location of the watershed, covering an area of 37 km<sup>2</sup>, is shown in Figure 1a. The climate on the watershed is semi-arid, with a single rainy season from July to October. The major soils types (Figure 1b) are slightly evolved soils (25% of the area), crude mineral soils (35% of the area, prone to physical degradation into glaucis) and hydromorphic soils (40% of the area) [IGB, 2002]. Elevation on the watershed varies between 318 to 338 meters, with slopes between 0–17% (Figure 1c).

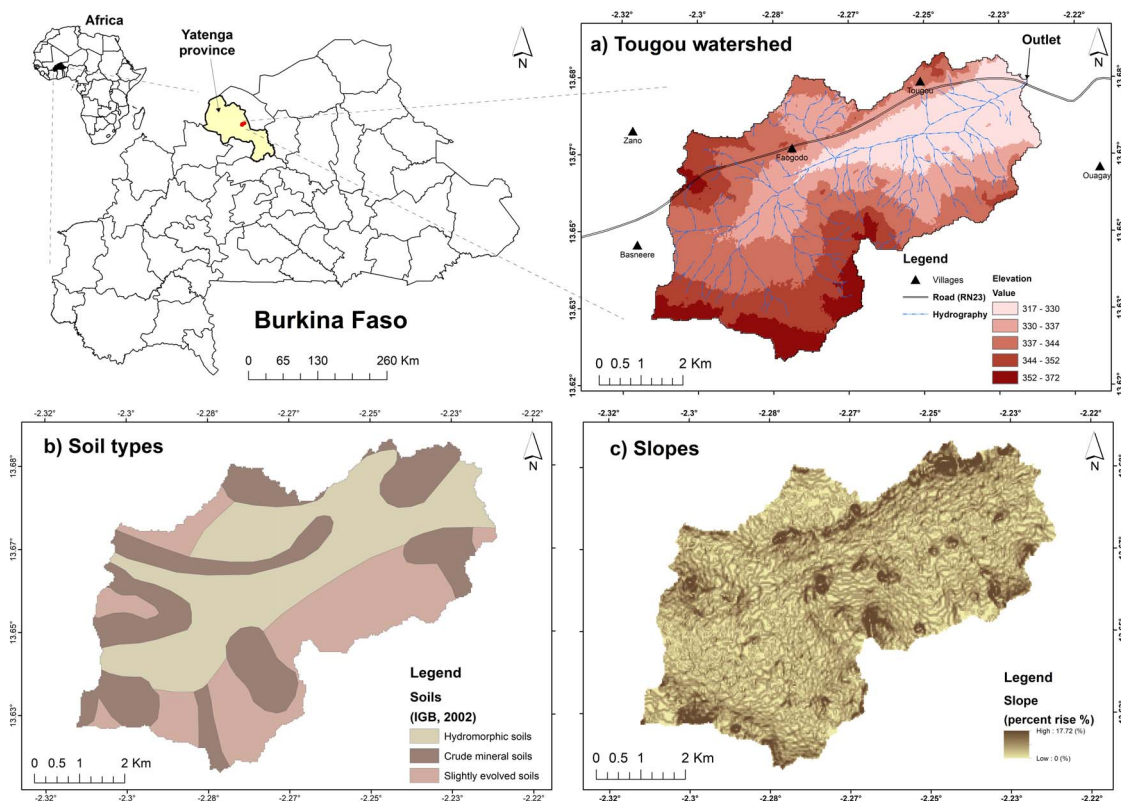
According to the meteorological records provided by the National Meteorology Agency (ANAM-BF), over the reference period 1995–2014, the average annual rainfall is 700 mm. Daily temperatures range from 12 to 45 °C. The vegetation cover is sparse on the watershed and mostly made of savannah, shrubs, steppes and agrarian parklands. The drainage network consists of a dendritic network of ephemeral

streams, collecting overland runoff to a main and intermittent river, long of 8 km [Mounirou, 2012, Yonaba, 2020].

#### 2.1.2. LULC evolution in Tougou watershed

Land use maps of the Tougou watershed for the years 1999, 2009 and 2017 (Figure 2a–c) have been previously produced in Yonaba *et al.* [2021b] through the analysis of Landsat satellite images. A land change model built around a Multilayer Perceptron neural network has been used, in conjunction with population growth trends, to further forecast the land use maps of the watershed in 2030, 2040 and 2050 (Figure 2d–f), under a business-as-usual scenario [Yonaba *et al.*, 2021b]. The land use categories considered in these maps are natural vegetation, bare lands and croplands. Natural vegetation areas are under permanent vegetation all the year-round, covered with herbaceous plants, shrub or trees; they are mostly found along the main river channel in the watershed. Bare lands exhibit a typically very low vegetation cover (less than 5%), made of residual clumps and herbaceous, with crusts and sealing developing at the surface, therefore limiting infiltration. Croplands are growing cereals (millet, sorghum) in the wet season (early June to mid-October), during which traditional farming practices are carried out on these soils to sustain crop production [Zouré *et al.*, 2019, Yonaba *et al.*, 2021b]. In the watershed, populations are relying of traditional wells (of 15–30 m depth) for water consumption; however, the amounts of such withdrawals are negligible [Rusagara *et al.*, 2022] and were therefore not considered in this study.

In total, from 1999 to 2050, croplands are projected to increase from 48.6% to 86.6%, bare lands to decrease from 44.8% to 10.7% whereas natural vegetation is expected to decrease from 6.6% to 2.7%. The large increase in croplands is mostly driven by the population growth trends. Conversely, the land use type targeted for conversion to croplands is bare land, thanks to the wide array of farming techniques and soil conservation measures introduced in the region since the 1990s [Lèye *et al.*, 2021, Nyamekye *et al.*, 2018, Zouré *et al.*, 2019].



**Figure 1.** Location of the Tougo watershed (northern Burkina Faso). (a) Elevation and hydrography. (b) Soil types [IGB, 2002]. (c) Slope map, derived from elevation map.

2.2. Overview of the methodology used in this study

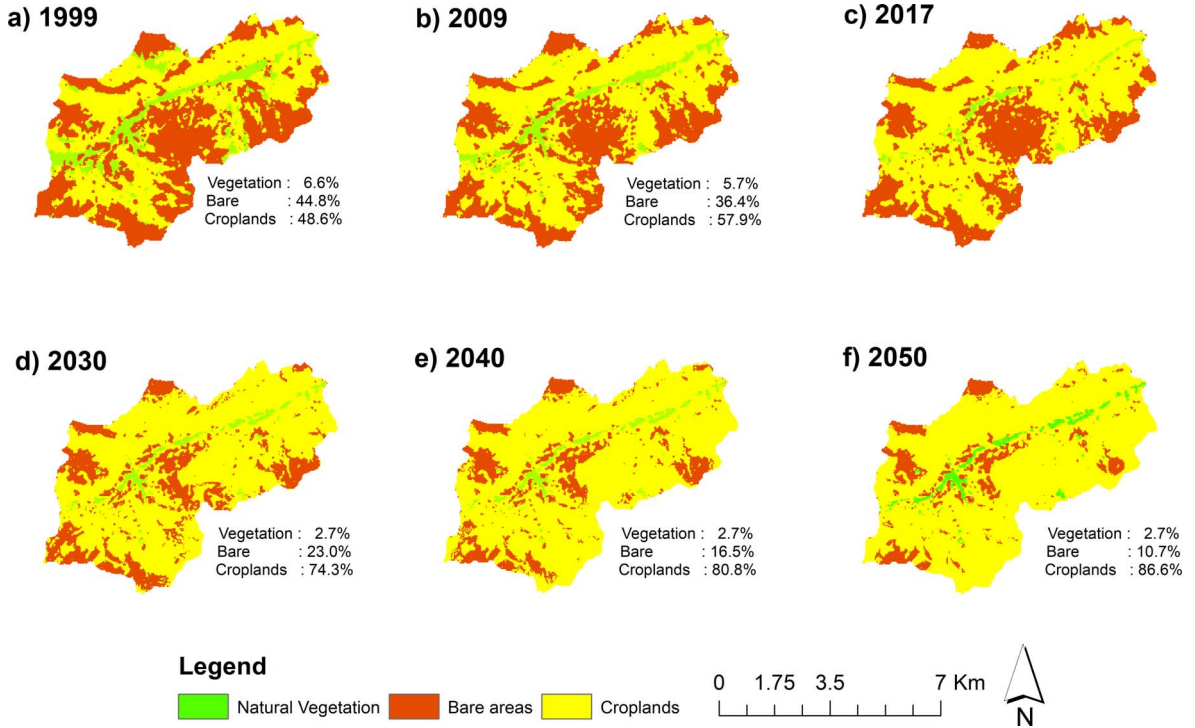
The general methodology adopted in this study is presented in Figure 3. The major steps consisted in selecting CMIP6 Global Circulation Models (GCM), then carrying out a bias correction of these models using observations during the baseline period 1995–2014. The SWAT model is then used to simulate hydrological processes in the baseline period (using LULC maps in 1999, 2009 and 2017), and the future period (using future LULC maps in 2030, 2040 and 2050). The comparison of changes in surface runoff under baseline and future conditions is used to assess the relative contributions of climate and land use changes.

2.3. Estimation of daily potential evapotranspiration (PET)

The observed climate variables available for the baseline period (1995–2014) are rainfall ( $pr$ ), average daily temperature ( $tas$ ), minimum daily temperature ( $tasmin$ ), maximum daily temperature ( $tasmax$ ), both provided by the National Meteorology Agency (ANAM-BF), at the daily timestep. Therefore, the use of the reference FAO-56 Penman–Monteith (PM) method is not applicable. To estimate daily PET, the Hargreaves–Samani (HS) equation [Hargreaves and Samani, 1985] is therefore selected as an alternative for its parsimonious estimation of PET solely relying on temperature data, which is available for the considered baseline period. The HS equation is given as in Equation (1) [Raziei and Pereira, 2013]:

$$PET = 0.0135k_{rs} \frac{R_a}{\lambda} (T_x - T_n)^n (T_m + b) \quad (1)$$





**Figure 2.** Land use maps of Tougo watershed for the period 1999–2050. Land use maps in 1999, 2009 and 2017 were produced through remote sensing images analysis. Land use maps in 2030, 2040 and 2050 were produced through land change modelling. Further details are given in Yonaba et al. [2021b].

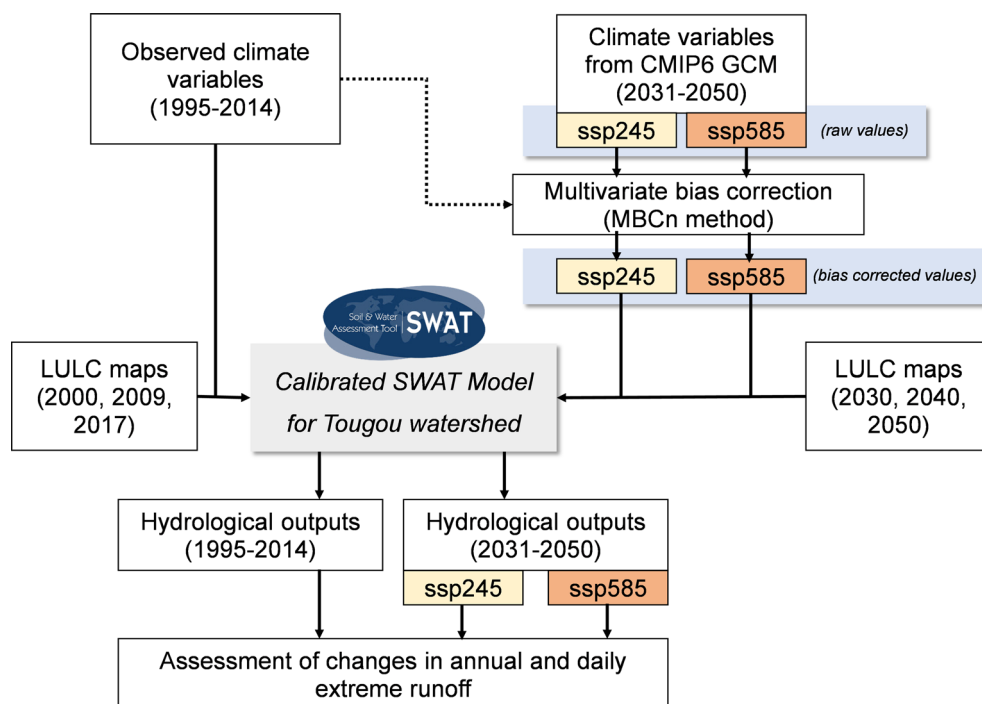
where  $R_a$  is the extraterrestrial radiation ( $\text{MJ}\cdot\text{m}^{-2}\cdot\text{d}^{-1}$ ),  $\lambda$  the latent heat of vaporization ( $2.45 \text{ MJ}\cdot\text{kg}^{-1}$ ),  $T_x$ ,  $T_n$  and  $T_m$  are respectively the daily maximum, minimum and average temperatures ( $^\circ\text{C}$ ). In Equation (1),  $k_{rs}$ ,  $n$  and  $b$  are coefficients, originally defined as 0.17, 0.5 and 17.8 respectively [Raziei and Pereira, 2013].

Previous studies reported the reliability of HS equation for PET estimation in Burkina Faso, yet highlighted significant biases often occurring in daily estimates, which can be further reduced by calibrating the empirical coefficients in the formula against reference data [Ibrahim, 2002, Ndiaye et al., 2017, Zouré, 2019, Yonaba, 2020]. In this study, these coefficients were calibrated for each month in the year, using PET data calculated from Penman–Monteith (PM) reference model over the previous period 1985–1994 (for which complete climate data were available). The calibration procedure is carried out using an adaptive non-linear least squares algorithm. The improvement of the calibration

step is evaluated through graphical methods (including scatterplots and box-plots) and also statistical evaluation (coefficient of determination— $R^2$ , Mann–Whitney  $U$ -test for the mean at  $\alpha = 5\%$  significance level, root mean square error—RMSE, and percent bias—PBIAS) Some of these metrics are calculated as in Equation (2):

$$\left\{ \begin{array}{l} R^2 = 1 - \frac{\sum_{i=1}^n (x_i^{\text{PM}} - x_i^{\text{HS}})^2}{\sum_{i=1}^n (x_i^{\text{PM}} - \overline{x_i^{\text{PM}}})^2} \\ \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i^{\text{HS}} - x_i^{\text{PM}})^2} \\ \text{PBIAS} (\%) = 100 \times \frac{1}{n} \sum_{i=1}^n \left( \frac{x_i^{\text{PM}} - x_i^{\text{HS}}}{x_i^{\text{PM}}} \right) \end{array} \right. \quad (2)$$

where  $x_i^{\text{HS}}$  and  $x_i^{\text{PM}}$  are daily PET calculated with HS and PM equations respectively,  $\overline{x_i^{\text{PM}}}$  is the average over  $x_i^{\text{PM}}$  values and  $n$  the length of the dataset.



**Figure 3.** Flowchart of the study methodology.

## 2.4. Processing future climate projections

### 2.4.1. Selection of GCM models

In this study, CMIP6 climate projections [Eyring et al., 2016] are used to project the future hydrological response. The data were retrieved from Copernicus Climate Data Store (CDS, <https://cds.climate.copernicus.eu/cdsapp#!/dataset/projections-cmip6>) for the baseline period 1995–2014 and the future period 2031–2050. Two Shared Socio-Economic Pathways (SSP) are considered: SSP2-4.5, a modest mitigation case scenario, standing as a most-plausible near future outcome ( $\sim 2.5$  °C global warming by 2100, relative to pre-industrial levels); and SSP5-8.5, a worst-case scenario defined by business-as-usual fossil-fuel intensive use, stringent of climate mitigation ( $\sim 5$  °C warming by 2100). The choice of these scenarios is also influenced by the fact that they remain comparable to the previous Representative Concentration Pathways (RCP) 4.5 and 8.5 scenarios, issued from CMIP5 and broadly used in past studies [Ayugi et al., 2022]. It is noteworthy to outline that as compared to the land use narratives in the Shared Socio-Economic Pathways

(SSPs) scenarios, the LULC futures in Tougou watershed used in this study are more in line with the “middle of the road” SSP2-4.5 scenario [Popp et al., 2017].

The initial set of models available within the CDS archive was filtered to retain only models providing complete daily data for rainfall, temperature (minimum, maximum and average) and for both SSP2-4.5 and SSP5-8.5 pathways. An ensemble of 9 climate models, listed in Table 1, is finally considered for this study. These models all have a horizontal resolution of 250 km and are issued from the same ensemble member (*r1i1p1f1*).

### 2.4.2. Multivariate bias correction of future climate data

GCM outputs generally feature systematic biases over the historical baseline, which needs to be removed before their use in impact studies. To address this issue, the application of bias correction techniques is recommended [Dieng et al., 2022]. In this study, a multivariate trend preserving bias correction method, termed as *MBCn* [Cannon, 2018] is used. The general steps of application of this method

**Table 1.** List of CMIP6 models used in this study

In this study	Abbreviation	Model name
M1	ACCESS-CM2	Australian Community Climate and Earth System Simulator Model Version 2
M2	CNRM-CM6-1	Centre National de Recherches Meteorologiques Model Version 6.1
M3	CNRM-ESM2-1	Centre National de Recherches Meteorologiques—Earth System Model Version 2.1
M4	INM-CM4-8	Institute for Numerical Mathematics (INM) Model Version 4.8
M5	INM-CM5-0	Institute for Numerical Mathematics (INM) Model Version 5
M6	MIROC6	Model for Interdisciplinary Research on Climate Version 6
M7	MIROC-ES2L	Model for Interdisciplinary Research on Climate—Earth System Model
M8	MPI-ESM1-2-LR	Max-Planck-Institut für Meteorologie—Earth System Model
M9	MRI-ESM2-0	Meteorological Research Institute Earth System Model Version 2.0

are as follows: (1) the data for each variable of interest are pooled by individual months in the baseline (1995–2014) and the future (2031–2050) periods; (2) for each climate model and for each time series (one for each month), model-projected quantiles are detrended, then adjusted through quantile mapping constructed from a sampling of the baseline observations. In this study, a sampling of 50% of the observations is used to construct the optimal targeted quantile distributions; then, the initial trends are re-introduced. These steps form the so-called *Quantile Delta Mapping* (QDM) bias-correction, which is further detailed in Cannon et al. [2015]. The *MBCn* algorithm further comes as a multivariate generalization of the QDM method for simultaneous bias-correction of multiple variables, to preserve joint dependence [Cannon, 2018].

The bias-corrected variables in this study are daily rainfall and temperatures (average, minimum and maximum). Future daily PET is further calculated with the calibrated HS equation (presented in Section 2.3). This strategy is adopted here since correcting biases in dependent variables (here, temperature) has the potential of significantly reduce biases in PET forecasts [Yang et al., 2021].

## 2.5. Hydrological simulation

The SWAT model is used in this study for the simulation of hydrological processes in the Tougou watershed. SWAT is a physically-based and semi-distributed hydrological model, which simulates hydrological processes at the scale of Hydrological Response Units (HRUs) [Arnold et al., 1998]. The water

balance equation as represented by the SWAT model is given by Equation (3):

$$SW_t = SW_0 + \sum_{i=1}^t (P_{\text{day}} - Q_{sf,\text{day}} - ET_{\text{day}} - w_{\text{seep},\text{day}} - Q_{gw,\text{day}}) \quad (3)$$

where  $SW_t$  and  $SW_0$  are the final and initial soil water (respectively),  $t$  is the elapsed time (in days),  $P_{\text{day}}$  is the daily rainfall,  $Q_{sf,\text{day}}$  is the daily surface runoff,  $ET_{\text{day}}$  is the daily actual evapotranspiration,  $w_{\text{seep},\text{day}}$  is the daily seepage from the soil profile to the vadose zone and  $Q_{gw,\text{day}}$  is the daily return flow from the aquifer. The surface runoff in the SWAT model is calculated using the empirical Soil Conservation Service (SCS) Curve Number (CN) method, given as in Equation (4) [Neitsch et al., 2011]:

$$Q_{sf,\text{day}} = \frac{(P_{\text{day}} - I_a)^2}{P_{\text{day}} - I_a + S}; \quad S = 25.4 \left( \frac{1000}{\text{CN}} - 10 \right) \quad (4)$$

where  $I_a$  is the sum of initial abstractions (including soil surface storage, interception and infiltration prior to the onset of runoff) and  $S$  is a retention parameter, which is controlled by spatial variations in soils, land use, management and slope. The CN parameter is therefore tabulated for all these conditions [Neitsch et al., 2011].

The SWAT model for the Tougou watershed was built using the elevation data from ASTER GDEM Digital Elevation Model, the soil map from the National Geographic Institute in Burkina Faso [IGB, 2002] and the available land use maps from Yonaba et al. [2021b]. The Land use Update (LUP) module within the SWAT model [Moriassi et al., 2019, Pai and Saraswat, 2011] is used to activate dynamic changes in land use during the simulations. The simulation

during the baseline period used land use maps in 1999, 2009 and 2017, whereas the simulation during the future period used projected land use maps in 2030, 2040 and 2050.

The SWAT model has been previously calibrated for the Tougou watershed (over the period 2004–2018) with satisfactory performance (*Kling-Gupta efficiency*:  $KGE = 0.95/0.94$ ; *percent bias*:  $PBIAS = -2.30\%/2.90\%$  for calibration/validation periods). The Table 2 is adapted from Yonaba et al. [2021a] and shows the calibrated values (and uncertainty range) for the model parameters, which were used in this study for hydrological simulation. Among these parameters, CN2 (SCS Curve Number), SOL\_AWC (soil water content), ESCO (soil evaporation compensation factor), OV\_N (Manning's n roughness for overland flow), SOL\_K (saturated hydraulic conductivity) are the ones which are affected straight by land use [Amogu et al., 2015, Gal et al., 2017, Descroix et al., 2018, Yonaba et al., 2021a].

## 2.6. Extreme value analysis

Extreme values (EV) often constitute the rationale for designing hydraulic structures for flood control, or devise mitigation strategies for water management at the community level. Such extremes are likely to be affected by land use or climate change [Houknpè et al., 2019, Tazen et al., 2019]. In this study, different EV distributions (including Gumbel, GEV, Gamma, Weibull) were fitted over the daily annual maximum rainfall and annual maximum surface runoff values distributions over the baseline period 1995–2014. The quality of the fit was evaluated using the Kolmogorov–Smirnov test (at  $\alpha = 5\%$  significance level). The three-parameter Weibull distribution (W3) was finally selected as the optimally fitting EV distribution ( $p$ -value = 0.992 and 0.945 for rainfall and surface runoff extremes respectively) and is therefore used in this study, similarly as in Sawadogo and Barro [2022] in Burkina Faso, or in Li et al. [2015] and Olivera and Heard [2019]. The probability density function (PDF) and the quantile (inverse cumulative distribution function) of this distribution are given in Equation (5):

$$\begin{cases} f(x) = \frac{\beta}{\gamma} \left( \frac{x-\mu}{\gamma} \right)^{\beta-1} e^{-\left(\frac{x-\mu}{\gamma}\right)^\beta} \\ Q(p) = \gamma(-\ln(1-p))^{1/\beta} + \mu \end{cases} \quad (5)$$

where  $x > \mu$  is a given quantile,  $\beta$ ,  $\gamma$  and  $\mu$  are the shape, scale and location parameters of the Weibull distribution (respectively),  $p = 1 - (1/T)$  is the non-exceedance probability associated with the return level  $T$  in years. The quality of the W3 fit was assessed through the non-parametric Kolmogorov–Smirnov test (at  $\alpha = 5\%$  significance level).

The fitted W3 distributions are used to derive extreme daily rainfall and surface runoff for 2, 5, 10, 15, 20, 25 and 30-years return periods. These values are further used to assess the various contributions of climate and land use change on surface runoff extremes.

## 2.7. Evolution of ecohydrological signatures

To assess the change in ecohydrological signatures of the Tougou watershed between the baseline and the future periods, the water-energy budget of Tomer and Schilling [2009] is used in this study. This framework states that as the climate change signal affects the watershed response, a tendency in maximizing unused water (defined as  $P_{ex}$ ) and unused energy ( $E_{ex}$ ) occurs, as a depiction of the shift in ecohydrological signatures of the watershed. The direction of the change can therefore be explained in terms of: (i) increasing runoff or water stress (moving right or left along the  $P_{ex}$  axis, respectively); (ii) increased humidity or drier conditions (moving upward or downward along the  $E_{ex}$  axis, respectively). The excess water  $P_{ex}$  and excess energy  $E_{ex}$  are calculated as given by Equation (6):

$$\begin{cases} P_{ex} = \frac{P - ET_a}{P} \\ E_{ex} = \frac{PET - ET_a}{PET} \end{cases} \quad (6)$$

where, for each period,  $P$  is the average annual rainfall,  $ET_a$  is the average annual actual evapotranspiration and  $PET$  is the average annual potential evapotranspiration.

## 2.8. Contributions of climate and land use changes to surface runoff

The total relative change in average annual surface runoff signal between the baseline and the future (SSP2-4.5 and SSP5-8.5) is averaged over the climate model ensemble between the baseline period

**Table 2.** Calibrated SWAT model parameters in Tougou watershed

Parameter	Description (Unit)	Calibration method	Initial range	Fitted values (Uncertainty range)
CN2	SCS Curve number (-)	<i>r</i>	-0.3–0.3	-0.0787 (-0.249;0.057)
EPCO	Plant uptake compensation factor (-)	<i>v</i>	0–1	0.2036 (0.091; 0.402)
SOL_K	Saturated hydraulic conductivity (mm·h <sup>-1</sup> )	<i>r</i>	-0.3–0.3	-0.0232 (-0.157;0.272)
GW_DELAY	Groundwater delay time (days)	<i>v</i>	0–500	52.584 (25.203; 53.799)
RCHRG_DP	Deep aquifer percolation fraction (-)	<i>v</i>	0–1	0.3566 (0.078; 0.239)
GW_REVAP	Groundwater “revap” coefficient (-)	<i>v</i>	0.02–0.2	0.0948 (0.065; 0.156)
OV_N	Overland flow Manning roughness (s·m <sup>-1/3</sup> )	<i>r</i>	-0.3–0.3	-0.3038 (-0.595; 0.152)
CH_K2	Hydraulic conductivity in channels (mm·h <sup>-1</sup> )	<i>v</i>	0–300	50.173 (23.66; 114.30)
SOL_AWC	Available soil water content (mm)	<i>r</i>	0–0.5	0.3707 (0.326; 0.405)
REVAPMN	Threshold depth of water for “revap” (mm)	<i>v</i>	0–500	344.08 (247.69; 416.06)
CH_N2	Manning’s roughness for channels (s·m <sup>-1/3</sup> )	<i>v</i>	0.01–0.3	0.178 (0.175; 0.287)
ALPHA_BF	Baseflow recession constant (days)	<i>v</i>	0–1	0.3791 (0.367; 0.790)
ESCO	Soil evaporation compensation factor (-)	<i>v</i>	0–1	0.9879 (0.966; 1.000)

The parameters are ranked out by order of decreasing importance in the model. The units provided in the description stand for the absolute parameter value [Neitsch et al., 2011]. The calibration method refers to the strategy used for updating the model parameters during the calibration phase: *r* (for relative) is the relative deviation applied to the original parameter value, which therefore preserves its spatial variability; *v* (for replace) is the absolute and unique value given to the parameter in the model, regardless of its location [Abbaspour et al., 2007, Yonaba et al., 2021a].

and the future SSP scenario considered, as shown in Equation (7):

$$\Delta Q(\%) = \frac{1}{n} \sum_{i=M_1}^{M_n} \left( \frac{Q_{M_i,SSP_j} - Q_b}{Q_b} \times 100 \right) \quad (7)$$

where  $M_i$ ,  $i \in \{1, \dots, n\}$  refers to each model within the ensemble,  $n = 9$  is the total number of climate models used,  $Q_{M_i,SSP_j}$  refers to the average annual surface runoff in the future period under  $SSP_j$  ( $j \in \{2-4.5 \text{ or } 5-8.5\}$ ) using climate model  $M_i$ ,  $Q_b$  the average annual surface runoff during the baseline period.

As this total relative change is both affected by climate and land use change, the isolated contributions of each factor are evaluated using a “fixing-changing” method, where a single factor (either land

use or climate) is changing [Yonaba et al., 2021a]. This needs intermediate simulations summarized in Table 3. The separation method however assumes that land use changes are independent of climate change, and the land-use change affects only actual evapotranspiration (ET). The latter assumption holds since actual ET is mostly dependent on vegetation (cover fraction, growing stage and development) and land use management [Dey and Mishra, 2017, Gbohoui et al., 2021, Yonaba et al., 2021a].

The relative contributions of climate ( $\eta_{climate}$ ) and land use ( $\eta_{lulc}$ ) changes to the total changes in surface runoff are evaluated under the assumption that these relative contributions, along with climate-land use interaction effects [Gbohoui et al., 2021, Yon-

**Table 3.** Simulation runs to isolate land use and climate change effects on surface runoff

Climate period	LULC period	Response	Description	Effect
Baseline	Baseline	$Q_b$	Runoff in the baseline period	-
Baseline	Future	$Q_{lulc}$	Changing land use while holding climate constant	Land use change
Future (ssp2-4.5, ssp5-8.5)	Baseline	$Q_{climate}$	Changing climate while holding LULC constant	Climate change
Future (ssp2-4.5, ssp5-8.5)	Future	$Q_{future}$	Runoff in the future period	Land use + Climate change

**Table 4.** Monthly calibrated parameters of Hargreaves–Samani equation for daily PET estimation in Tougou watershed

Parameters	J	F	M	A	M	J	J	A	S	O	N	D
$k$ Estimate	0.190	0.190	0.190	0.190	0.190	0.190	0.190	0.165	0.172	0.190	0.190	0.190
$k$ Std. error	0.072	0.071	0.087	0.088	0.079	0.051	0.062	0.066	0.050	0.118	0.071	0.068
$k$ $p$ -value	0.009	0.008	0.030	0.031	0.017	0.000	0.002	0.014	0.001	0.108	0.008	0.006
$n$ Estimate	0.221	0.229	0.234	0.232	0.389	0.502	0.468	0.506	0.495	0.175	0.275	0.213
$n$ Std. Error	0.056	0.060	0.052	0.045	0.072	0.066	0.071	0.077	0.053	0.051	0.059	0.056
$n$ $p$ -value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	$<2 \times 10^{-16}$	0.001	0.000	0.000
$b$ Estimate	50.000	50.000	50.000	50.000	22.926	10.000	10.000	10.000	10.000	47.423	34.179	50.000
$b$ Std. Error	22.127	24.905	37.118	38.718	19.677	8.571	9.325	11.486	8.176	47.241	20.293	20.984
$b$ $p$ -value	0.025	0.046	0.179	0.198	0.245	0.245	0.285	0.385	0.222	0.316	0.093	0.018
Residual error	0.648	0.781	0.967	0.895	1.114	1.010	0.846	0.816	0.591	0.609	0.592	0.576

aba et al., 2021a] sum up to 100%, as given in Equation (8).

$$\left\{ \begin{array}{l} \eta_{climate}(\%) + \eta_{lulc}(\%) + \eta_{int}(\%) = 100\% \\ \eta_{climate}(\%) = \left( \frac{Q_{climate} - Q_b}{Q_b} \right) \times 100 \\ \eta_{lulc}(\%) = \left( \frac{Q_{lulc} - Q_b}{Q_b} \right) \times 100 \\ \eta_{int}(\%) = 100\% - \eta_{climate}(\%) - \eta_{lulc}(\%) \end{array} \right. \quad (8)$$

The same approach is used to disentangle climate and land use change contribution to changes in average annual surface runoff and also to changes in maximum daily surface runoff. In the latter case, however,  $Q_{climate}$  is defined as the ensemble median (instead of the average), since discrepancies across models within the ensemble become large with increasing return periods [Wallach et al., 2016].

### 3. Results

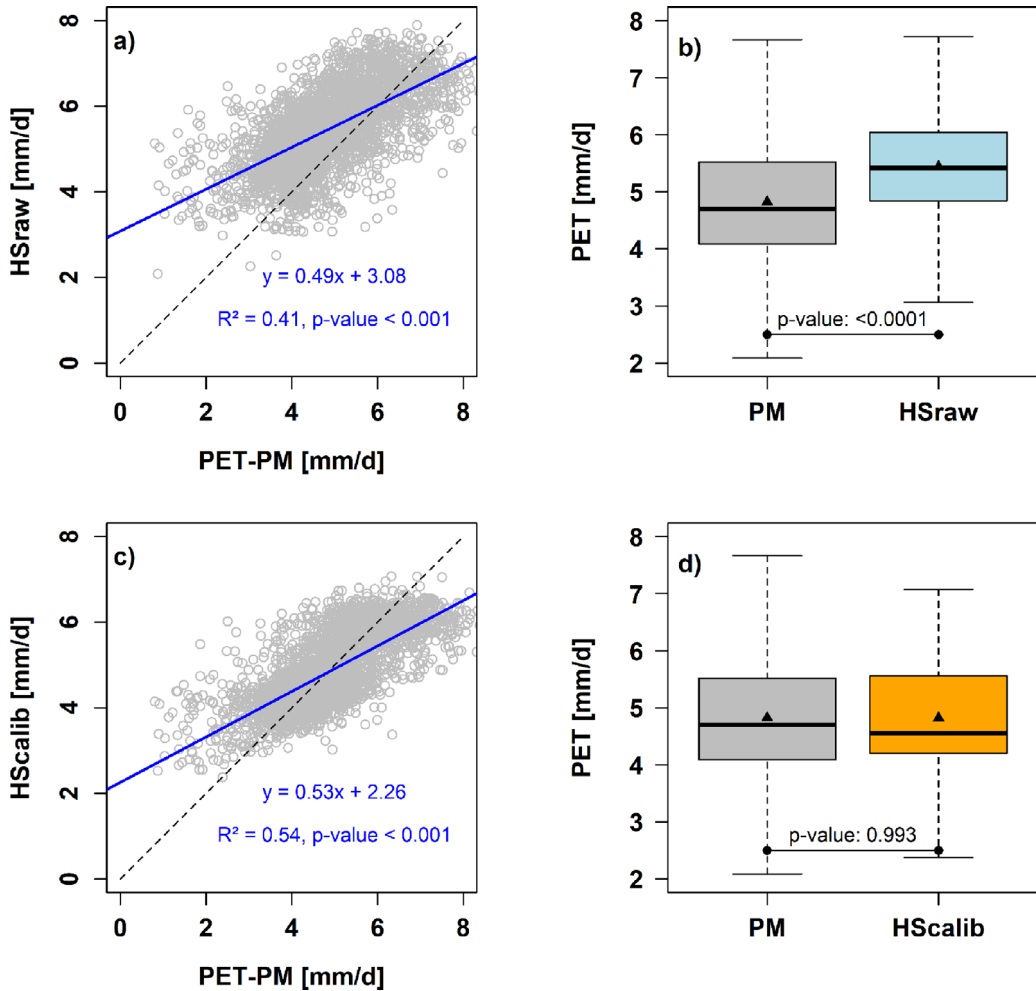
#### 3.1. Climate data processing

##### 3.1.1. Calibration of HS model for daily PET

The calibrated parameters of the Hargreaves–Samani equation for daily PET estimation is presented in Table 4, for each of the months in the year for which the formula is independently calibrated.

Figure 4 compares the distribution of daily PET values calculated with the reference Penman–Monteith (PM) model, the original HS model ( $HS_{raw}$ ) and the calibrated HS model ( $HS_{calib}$ ) over the period 1985–1994.

The  $HS_{raw}$  values overestimate daily PM (Mann–Whitney:  $p$ -value  $<0.0001$ ; RMSE = 1107 mm/d; PBIAS =  $-17.8\%$ ), whereas the calibrated  $HS_{calib}$  values are closer in distribution to PM values, with non-significant differences in the mean (Mann–Whitney:



**Figure 4.** Comparison between PET estimates from Penman–Monteith (PM) reference method and Hargreaves–Samani (HS) equation over the period 1985–1994. Panels (a) and (b) compare PM and uncalibrated HS formula. Panels (c) and (d) compare PET estimates from PM and calibrated HS formula.

$p$ -value = 0.993; RMSE = 0.801 mm/d; PBIAS = -3.7%). The calibrated HS model is therefore expected to provide reliable estimates of daily PET in the Tougou watershed.

### 3.1.2. Multivariate bias-correction of future climate data

Figure 5 shows the cumulative distribution functions (CDF) of the 9 climate models selected in this study, compared to those of observations over the baseline period 1995–2014, before and after the application of the multivariate bias-correction *MBCn*.

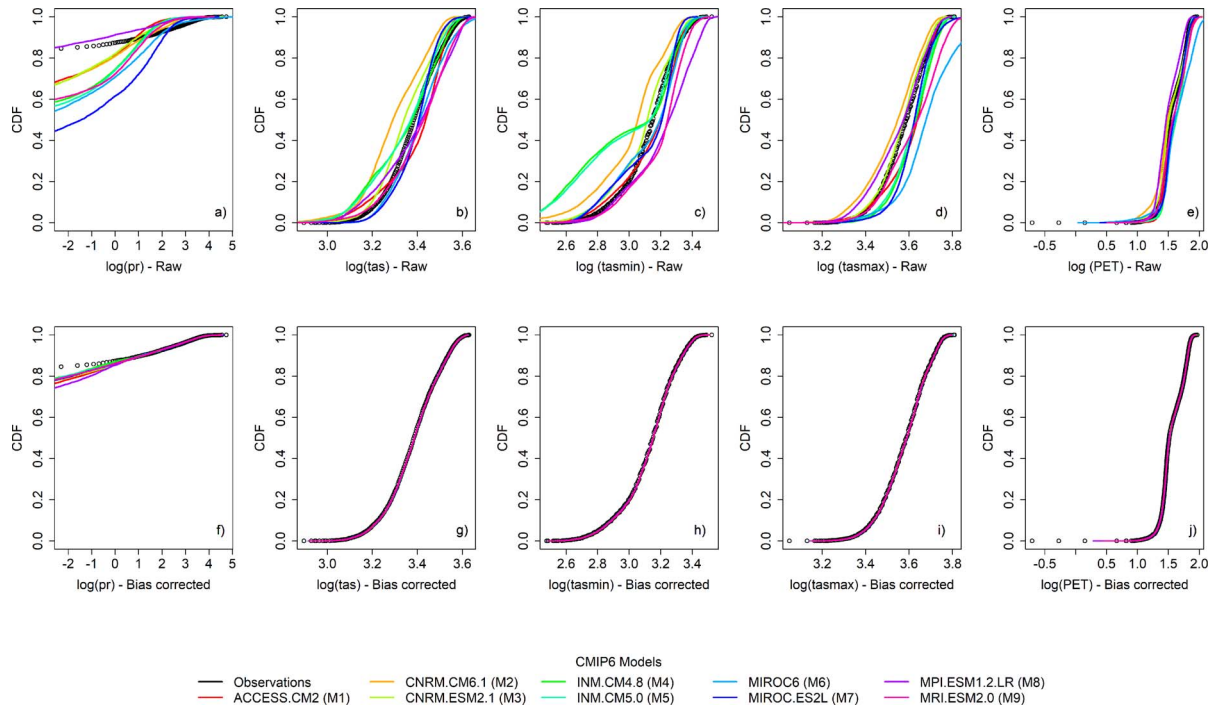
The moderate to large discrepancies observed in CDFs (especially in rainfall) initially become mar-

ginal after the bias-correction, hence demonstrating the performance of the bias-correction method in adjusting the distributions for all the considered climatic variables. The transfer functions used for adjusting the historical climate model outputs were applied to the future forecasts (SSP2-4.5 and SSP5-8.5), hence providing more skilful climate projections.

## 3.2. Analysis of the climate change signal

### 3.2.1. Projected changes in annual average rainfall and PET

The climate change signal in SSP2-4.5 and SSP5-8.5 relative to the base line period is presented



**Figure 5.** Cumulative distribution functions (CDFs) of climate variable before and after multivariate bias-correction over the baseline period 1995–2014. The  $x$ -axis scale for all panels were log-transformed to clearly visualize the differences in distributions (especially in rainfall).

in Figure 6a. The majority of the climate models within the ensemble project an increase in annual rainfall by 13.7% on average in SSP2-4.5 (−7.6% to 58.2%) and by 18.8% on average in SSP5-8.5 (−2.3% to 61.0%). Regarding annual PET, an increase is also expected by 1.3% on average in SSP2-4.5 (−0.2% to 2.5%) and by 1.5% on average in SSP5-8.5 (0.3% to 3.4%). The projected increase in PET is likely caused by the projected increase in temperatures.

An increase in monthly rainfall is expected (Figure 6b) for the months of June (17.6% and 14.3%), July (5.9% and 8.9%), August (13.7% and 20.8%) and September (15.8% and 26.5%), respectively under SSP2-4.5 and SSP5-8.5 scenarios. Similarly, for monthly PET, an increase is also projected, mostly for the warmest months of April (1.0% and 1.4%) and May (0.8% and 2.5%), but also for the rainy months of June (2.3% and 2.2%), July (1.7% and 2.0%) and August (1.7% and 0.9%).

### 3.2.2. Projected changes in daily rainfall extremes

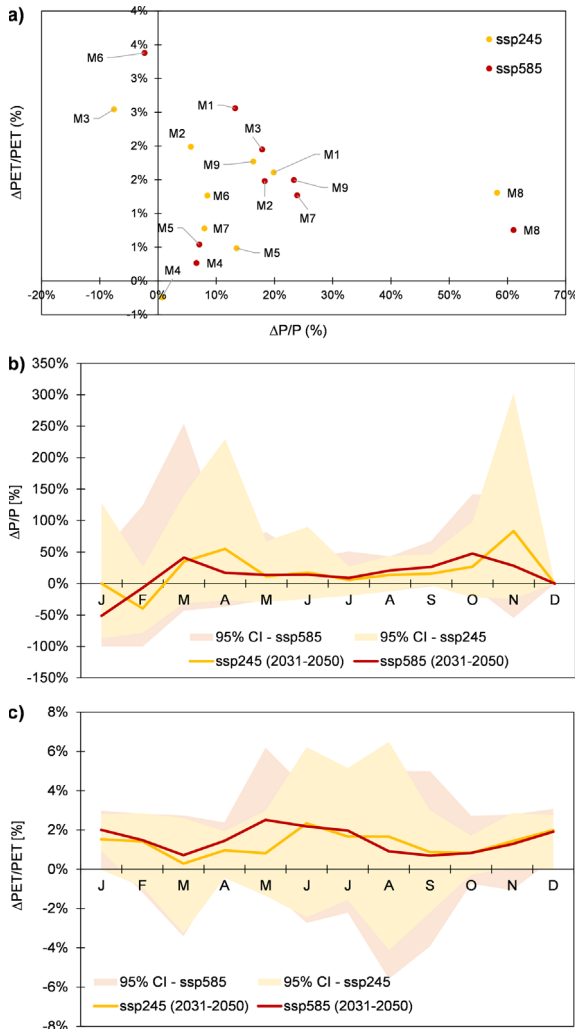
Figure 7 shows the evolution of selected rainfall quantile projected by SSPs scenarios in comparison to the baseline period, for return periods of 2, 5, 10, 15, 20, 25 and 30 years. The projected rainfall extreme considered in this study are defined as the median of the GCM ensemble. The projected increase in rainfall quantiles is significant and is estimated at 26.6% to 45.1% (under SSP2-4.5) and 35.5% to 77.2% (under SSP5-8.5) over the different return periods.

## 3.3. Projected changes on surface runoff

### 3.3.1. Annual changes

The analysis of hydrological simulation outputs for the baseline and the future period reveals significant changes in average annual components of the watershed hydrological balance, as shown in Figure 8. The increase in annual rainfall and PET is significant for both scenarios ( $p$ -value = 0.0032 and

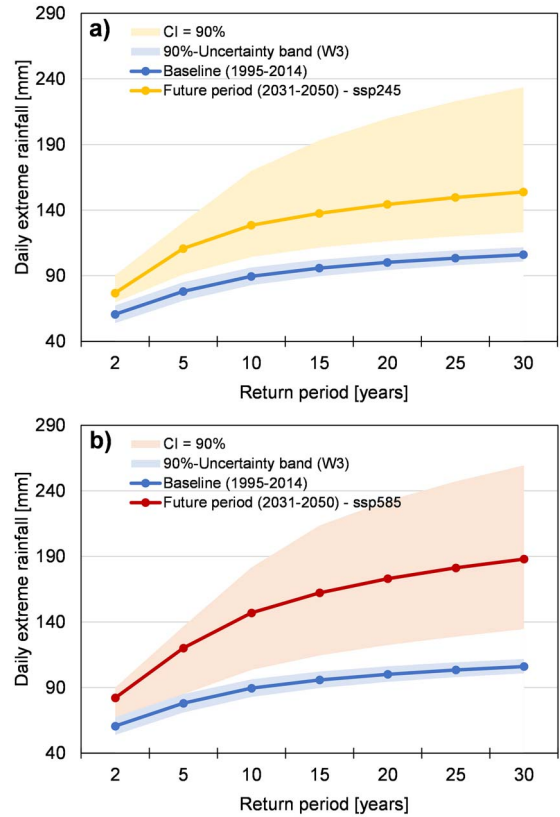




**Figure 6.** Annual and seasonal climate change signal in rainfall and PET relative to the baseline period 1995–2014. (a) Relative change in annual rainfall and PET. (b) Relative change in monthly rainfall. (c) Relative changes in monthly PET.

0.0003 respectively). Furthermore, significant increase in surface runoff (SURQ) by 24.2% (SSP2-4.5:  $p$ -value = 0.0018) and by 34.3% (SSP5-8.5:  $p$ -value = 0.0001) is projected. Likewise, the annual runoff coefficient (RC) is also projected to increase significantly by 11.7% under SSP5-8.5 scenario ( $p$ -value = 0.0095).

The relative changes in annual rainfall and surface runoff reported for each climate model in the en-



**Figure 7.** Projection of daily annual rainfall extremes in baseline (1995–2014) and future (2031–2050) periods. (a) Median projection under SSP2-4.5 scenario. (b) Median projection under SSP5-8.5 scenario. The shaded areas in light yellow and light orange represent the 90% uncertainty band around the median projections across the GCM ensemble. The blue shaded area refers to the 90% uncertainty band around Weibull-3 (W3) quantiles over the baseline period.

semble are presented in Figure 9. Under both SSPs, 8 climate models out of 9 (i.e. 89%) are projecting an increase in annual rainfall (Figure 9a) from 0.7% to 58.2% and from 6.6% to 61.0% (respectively under SSP2.4-5 and SSP5-8.5). Only 1 model (i.e. 11%) in the ensemble is projecting a decrease in rainfall of 7.6% (SSP2-4.5) and 2.3% (SSP5-8.5). Likewise, future average annual surface runoff (Figure 9b) is projected to increase under SSP2-4.5 (7 models, from 9.3% to 101.7%) and SSP5-8.5 scenarios (8 models,

from 10.9% to 110.8%).

### 3.3.2. Seasonal changes

The seasonal changes in surface runoff in the Tougou watershed are presented in Figure 10, superimposed with changes in monthly rainfall. Under the SSP2-4.5 scenario (Figure 10a), surface runoff is projected to increase especially in the months of June (15.7 mm) and August (23.6 mm), resulting in a bimodal distribution of the projected increase. Under the SSP5-8.5 scenario (Figure 10b), the projected increase is rather unimodal, with a peak projected in August (35.4 mm), gradually decreasing in September (23.9 mm). It also appears that the reported increase in surface runoff appear to be driven by the increase in rainfall, which shows that rainfall is likely the prominent factor driving seasonal changes in surface runoff, further altered to a lesser extent by other external factors.

### 3.4. Evolution of ecohydrological status

The evolution of the ecohydrological status of the Tougou watershed is presented in Figure 11. The majority of the climate projections (8 models under SSP2-4.5, 7 models under SSP5-8.5) shows a shift to the right in terms of excess rainfall ( $P_{ex}$ ), which translates as an increase in surface runoff under both SSP2-4.5 and SSP5-8.5 scenarios. However, in terms of aridity, less agreement is observed within the climate ensemble. A total of 4 models (out of 9) project a shift towards increased aridity or drier conditions (increase in excess energy  $E_{ex}$ ), whereas 5 models project a shift towards more humid conditions (decrease in excess energy  $E_{ex}$ ). The direction of the shift for the majority of the models in terms of ( $E_{ex}$ ) (in the positive  $P_{ex} - E_{ex}$  quadrant) suggests changes in land use such as deforestation, removal of perennials or the use of conservation tillage [Tomer and Schilling, 2009], which is in line with the future land use maps in the Tougou watershed (increase in agricultural areas, decrease of natural vegetation).

However, it should be critically assumed that such picture of the ecohydrological evolution of the Tougou watershed does not quantify streamflow change, but rather provide qualitative insights regarding whether the reported change is mostly driven by climate or land use [Dey and Mishra, 2017].

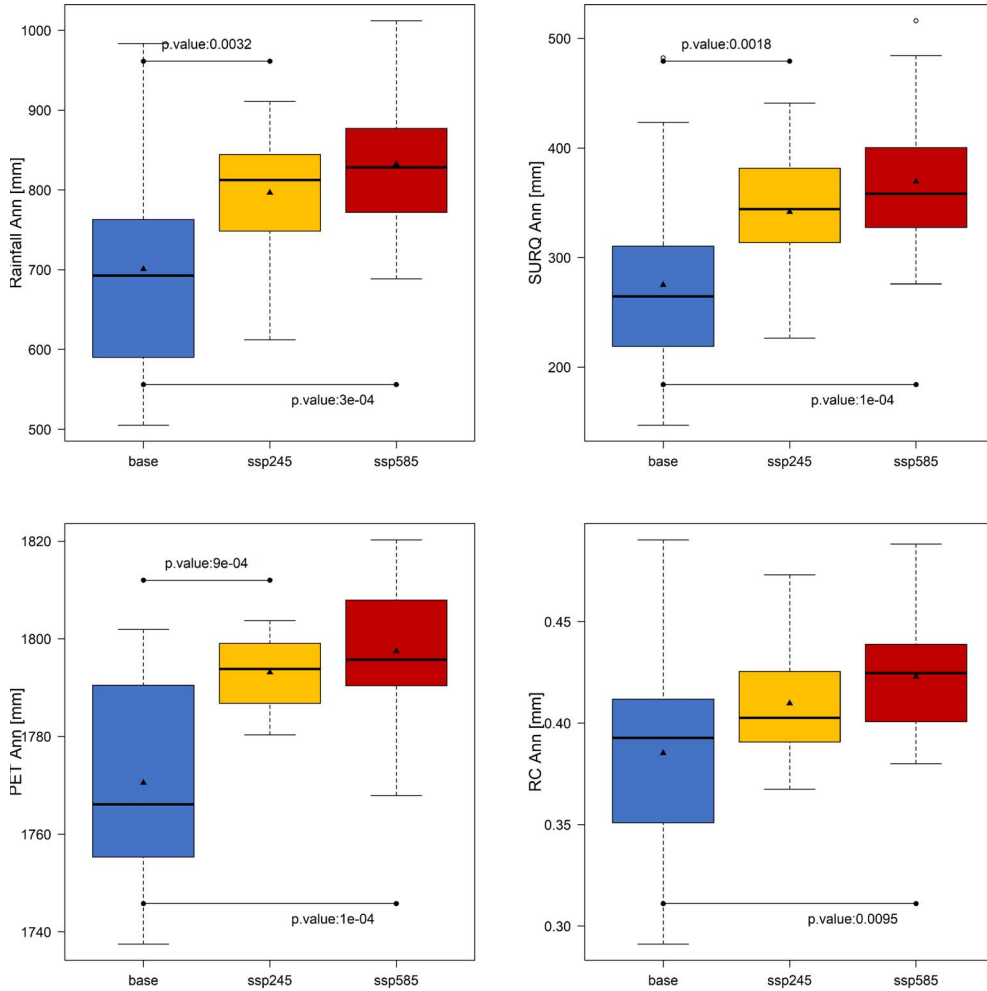
### 3.5. Evolution of surface runoff

#### 3.5.1. Sensitivities of surface runoff to rainfall and PET

The sensitivities of annual runoff change to rainfall and PET change under SSPs scenarios are presented in Figure 12 for all climate models in the ensemble used in this study. Such sensitivities, commonly referred to as streamflow elasticity, describes the change in streamflow related the changes in one climate variable [Andréassian *et al.*, 2016]. In this study, the analysis reveals that in the Tougou watershed, the change in surface runoff is not significantly affected by changes in annual PET (Figure 12a), as shown by the very small coefficients of determination (SSP2-4.5:  $R^2 = 0.0002$ ,  $p$ -value = 0.973; SSP5-8.5:  $R^2 = 0.1519$ ,  $p$ -value = 0.300). However, rainfall appears as the major factor driving changes in annual surface runoff (SSP2-4.5:  $R^2 = 0.9949$ ,  $p$ -value < 0.0001; SSP5-8.5:  $R^2 = 0.9928$ ,  $p$ -value < 0.0001), following a linear relationship (Figure 12b). Moreover, it can be inferred from Figure 12b that the projected average annual elasticity of streamflow to rainfall under SSP2-4.5 scenario is 0.69 (ratio of average annual increase of 66.6 mm in surface runoff out of 95.8 mm in rainfall). This elasticity increases to 0.72 under SSP5-8.5 scenario (ratio of average annual increase of 94.2 mm in surface runoff out of 131.6 mm in rainfall).

#### 3.5.2. Changes in daily Flow Duration Curve (FDC)

The changes in daily flow duration curves (FDC) during the rainy season (June to October) in the Tougou watershed are presented in Figure 13. The analysis shows that an increase in daily surface runoff is expected under both SSPs scenarios. For instance,  $Q_{10}$  (surface runoff equalled or exceeded 10% of the time) increases from 4.0 mm (baseline) to 4.2 mm (SSP2-4.5) or 4.6 mm (SSP5-8.5). Also,  $Q_{50}$  (surface runoff equalled or exceeded 50% of the time) increases from 1.2 mm (baseline) to 2 mm (under both SSPs). Likewise,  $Q_{80}$  (surface runoff equalled or exceeded 80% of the time) increases from 0.35 mm (baseline) to 0.70 mm (SSP2-4.5) or 0.80 (SSP5-8.5). The observed increase in FDC suggest and shift towards an intensification of flow regimes, likely driven by the intensification in the rainfall over the Tougou watershed.



**Figure 8.** Comparison of annual changes in hydrological processes in the Touyou watershed. PET: potential evapotranspiration. SURQ: surface runoff. RC: runoff coefficient, calculated as  $RC = SURQ/Rainfall$ .

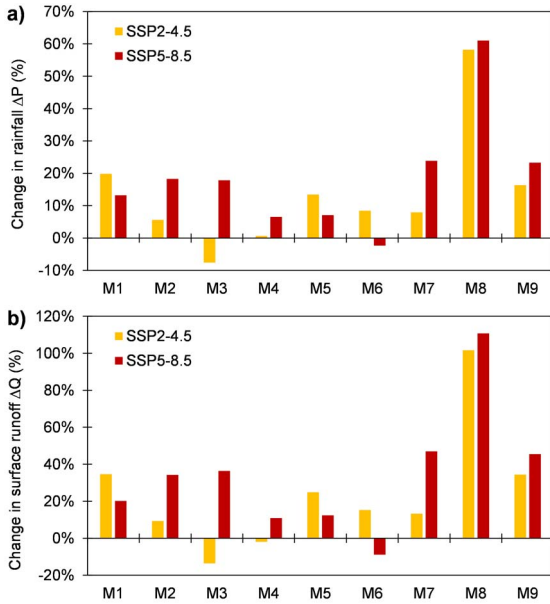
### 3.6. Effects of climate and land use changes on surface runoff

#### 3.6.1. Contribution of climate and land use to changes in annual surface runoff

The isolated relative contributions of climate and land use changes to annual surface runoff are presented in Figure 14, averaged over the period 2031–2050. Figures 14a and c shows the relative contribution of climate and land use for the specific case where surface runoff is projected to increase (respectively under SSP2-4.5 and SSP5-8.5 scenarios). At the opposite, Figures 14b and d shows the relative contribution of climate and land use for the specific case

where surface runoff is projected to decrease (respectively under SSP2-4.5 and SSP5-8.5 scenarios). For all panels in Figure 14, a positive (or relative) contribution translates that the projected changes in the factor (climate or land use) tends to cause an increase (or a decrease) in surface runoff.

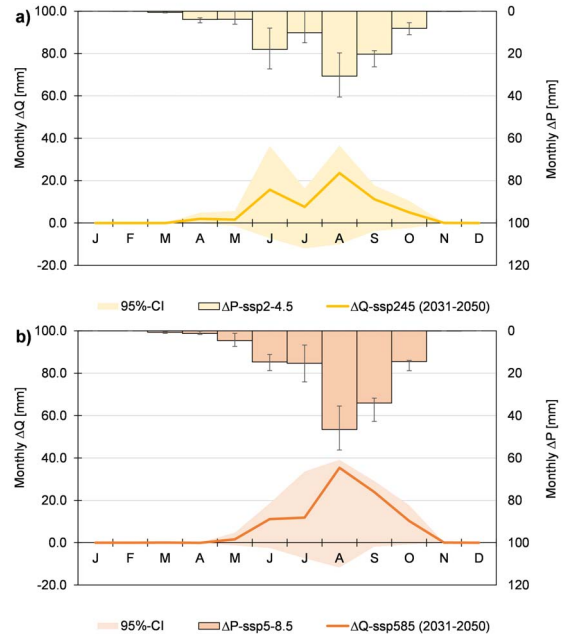
Under SSP2-4.5 scenario, surface runoff is projected to increase (between 25.7 mm and 279.6 mm) for 7 climate models (Figure 14a), which are the wetter ones (increase in average annual rainfall projected between 5.6% and 58.2%). In contrast (Figure 14b), for the 2 remaining drier climate models M3 and M4 (average annual rainfall change projected between  $-7.6\%$  and  $0.7\%$ ), surface runoff is projected



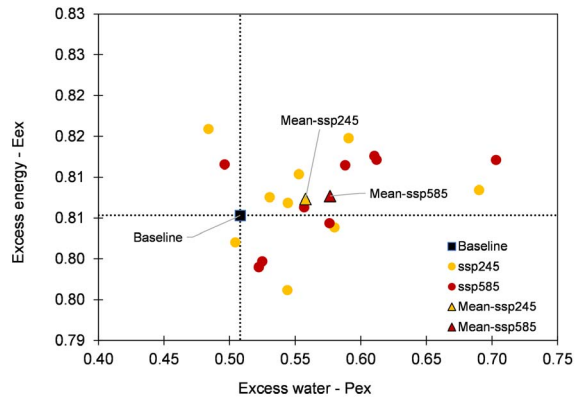
**Figure 9.** Relative average annual changes in rainfall (a) and surface runoff (b) between the baseline (1995–2014) and the future (2031–2050) period under SSP2-4.5 and SSP5-8.5 scenarios.

to decrease (between 2.3 and 37.5 mm). When surface runoff is projected to increase (Figure 14a), the relative contributions of climate and land use to such increase range from 109.0% to 170.7% and from -6.6% to -71.5% respectively. When surface runoff is projected to decrease (Figure 14b), these relative contributions now range from -246.9% to 55.3% and from 48.9% to 349.3%. In this case, the model M4 shows a positive contribution as it projects a small rainfall increase (0.7%), which tends to increase surface runoff, albeit this contribution appears minor in comparison to the land use effect (which consistently tends to decrease surface runoff).

Under SSP5-8.5 scenario, the projected figures are similar: for 8 climate models in the ensemble (Figure 14c), surface runoff is projected to increase (between 29.9 mm and 304.7 mm). These 8 climate models are mostly wet (average annual rainfall increase projected between 6.6% and 61.0%). The relative contributions of climate and land use to the projected increase in surface runoff range from 108.0% and 160.9% and from -6.0% to -61.3% respectively. For the remaining drier climate model, M6 (project-

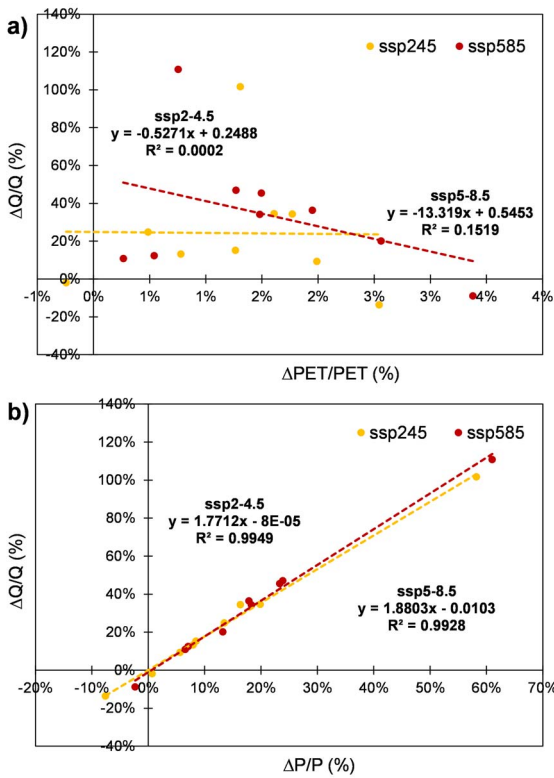


**Figure 10.** Monthly projected changes in surface runoff in the Tougoou watershed. (a) Under SSP2-4.5 scenario. (b) Under SSP5-8.5 scenario. The error bars on rainfall and the yellow and orange grey bands shows the 95% confidence interval around projected, calculated from the ensemble values.



**Figure 11.** Projected shifts in the ecohydrological status of the Tougoou watershed.

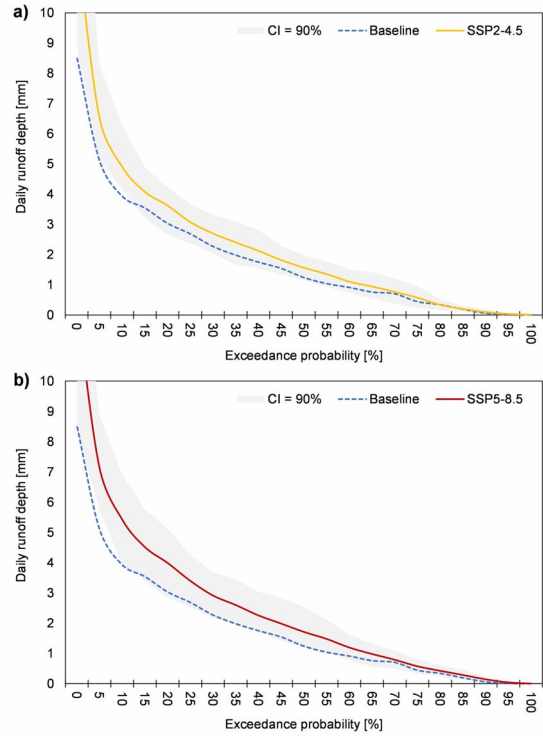
ing a decrease in annual rainfall of 2.3%), the annual surface runoff is projected to decrease by 24.4 mm: the relative contributions of climate and land use to this decrease are of 26.3% and 75.2% respectively



**Figure 12.** Sensitivities of annual surface runoff to rainfall and PET across all climate models in the ensemble of 9 models. Each dot shows the relative change in annual average for a given climate model between the baseline (1995–2014) and the future period (2031–2050).

(Figure 14d).

Overall, it appears that the relative contribution of climate is dominant under both SSP2-4.5 and SSP5-8.5 scenarios, as it is on absolute average 1.1 to 16.6 times higher (under SSP2-4.5) and 2.6 to 17.9 times higher (under SSP5-8.5) than that of land use. The increase in surface runoff could mainly be attributed to the increase in rainfall, further lessened by the effect of land use changes which typically tends to decrease surface runoff under all scenarios. The decrease in surface runoff can be explained by the decrease in bare areas between the baseline and the future period (from 44.8% to 10.7%). This decrease tends to produce less surface runoff. These bare areas are further replaced by cropland (which increased from 48.6% to 86.6%); these croplands tends to produce less surface runoff than bare areas [Yon-



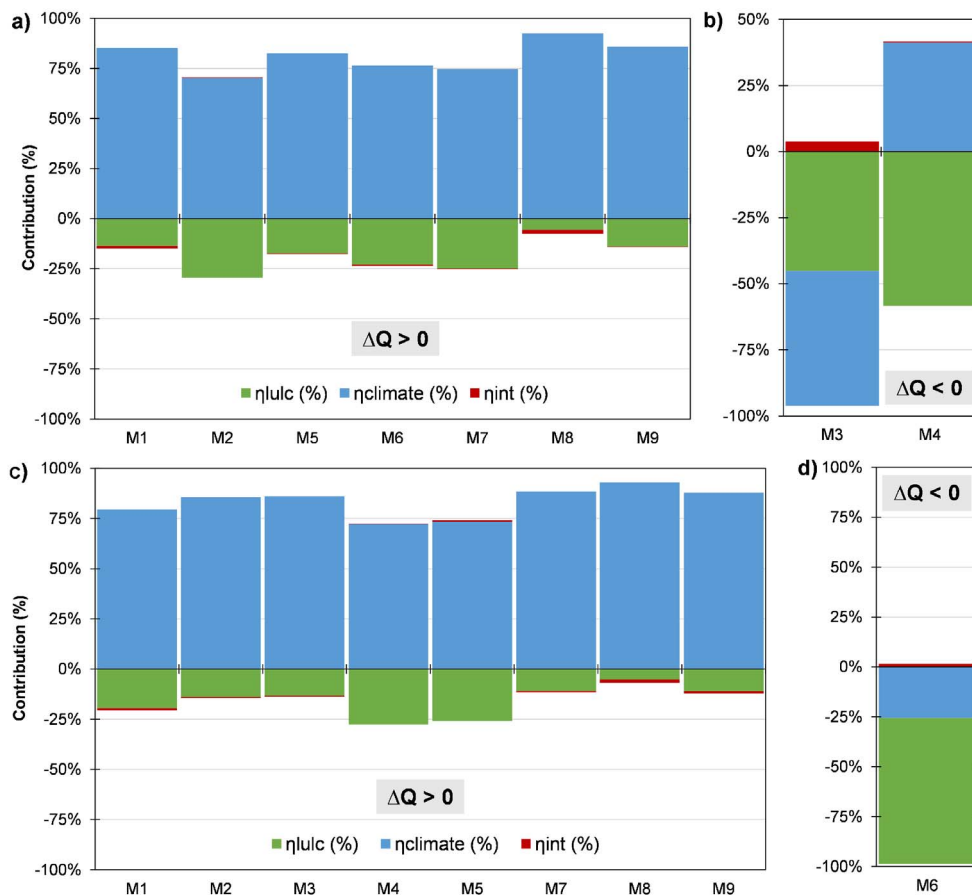
**Figure 13.** Projected changes in flow duration curve (FDC) in the Tougo watershed. (a) Changes in FDC under SSP2-4.5 scenario. (b) Changes in FDC under SSP5-8.5 scenario. The grey shaded area represents the 90%-confidence interval around the FDC curves for both scenarios.

aba et al., 2021a].

Smaller residual contributions, which range from –4.2% to 0.7% and from –2.0% to 1.6% under SSP2-4.5 and SSP5-8.5 respectively can be attributed to climate-land use interactions effects on surface runoff. Such contribution appears to be relatively small at the scale of the Tougo watershed. However, it should be acknowledged that these contributions convey various forms of uncertainties [Gbohoui et al., 2021, Onyutha et al., 2021, Yonaba et al., 2021a].

### 3.6.2. Contribution of climate and land use to changes in maximum daily surface runoff

The Figure 15 pictures the evolution of daily maximum surface runoff for return periods from 2 to 30 years, under various scenarios (climate change, land



**Figure 14.** Average relative contributions of climate ( $\eta_{\text{climate}}$ ) and land use ( $\eta_{\text{lulc}}$ ) to changes in average annual surface runoff between the baseline (1995–2014) and future (2031–2050) periods in Tougo watershed. (a) SSP2-4.5 scenario, surface runoff projected to increase. (b) SSP2-4.5 scenario, surface runoff projected to decrease. (c) SSP5-8.5 scenario, surface runoff projected to increase. (d) SSP5-8.5 scenario, surface runoff projected to decrease.  $\eta_{\text{int}}$  refers to the average relative contribution of interaction effects between climate and land use.

use change). The quantification of these changes is presented in Table 5. A total increase in maximum daily surface runoff (SSP2-4.5: 8.4 mm to 45.2 mm; SSP5-8.5: 17.1 mm to 70.8 mm) is projected. The analysis reveals that land use change causes a decrease in maximum daily surface runoff (SSP2-4.5:  $\eta_{\text{lulc}} = -22.8\%$  to  $-9.1\%$ ; SSP5-8.5:  $\eta_{\text{lulc}} = -11.1\%$  to  $-5.8\%$ ). Also, the absolute reduction induced by land use changes decreases for increasing return periods. At the opposite, climate change causes an increase in maximum daily surface runoff (SSP2-4.5:  $\eta_{\text{climate}} = 127.2\%$  to  $104.0\%$ ; SSP5-8.5:  $\eta_{\text{climate}} = 117.7\%$  to  $108.0\%$ ). Climate-land use interactions effects are

also quantified (SSP2-4.5:  $\eta_{\text{int}} = -4.4\%$  to  $5.1\%$ ; SSP5-8.5:  $\eta_{\text{int}} = -6.5\%$  to  $2.1\%$ ), although being relatively small in comparison to the isolated impact of climate and land-use changes. In definitive, the projected contributions of climate and land use on maximum daily surface runoff are similar to those observed on average annual values, albeit larger in magnitude.

**Table 5.** Isolated contributions of climate and land use to changes in daily surface runoff quantiles

Future scenario	Q <sub>2</sub>	Q <sub>5</sub>	Q <sub>10</sub>	Q <sub>15</sub>	Q <sub>20</sub>	Q <sub>25</sub>	Q <sub>30</sub>	
SSP2-4.5	ΔQ <sub>t</sub> (mm)	8.4	25.0	36.2	39.8	42.1	43.8	45.2
	ΔQ <sub>lulc</sub> (mm)	-1.9	-2.8	-3.4	-3.7	-3.9	-4.0	-4.1
	ΔQ <sub>climate</sub> (mm)	10.7	28.9	38.6	42.8	44.7	46.0	47.1
	η <sub>lulc</sub> (%)	-22.8%	-11.3%	-9.3%	-9.2%	-9.2%	-9.2%	-9.1%
	η <sub>climate</sub> (%)	127.2%	115.8%	106.7%	107.6%	106.1%	105.0%	104.0%
	η <sub>int</sub> (%)	-4.4%	-4.5%	2.7%	1.7%	3.1%	4.2%	5.1%
SSP5-8.5	ΔQ <sub>t</sub> (mm)	17.1	35.4	49.0	57.1	62.8	67.2	70.8
	ΔQ <sub>lulc</sub> (mm)	-1.9	-2.8	-3.4	-3.7	-3.9	-4.0	-4.1
	ΔQ <sub>climate</sub> (mm)	20.2	38.9	53.4	61.9	67.9	72.6	76.5
	η <sub>lulc</sub> (%)	-11.1%	-8.0%	-6.9%	-6.4%	-6.2%	-6.0%	-5.8%
	η <sub>climate</sub> (%)	117.7%	109.8%	108.8%	108.5%	108.2%	108.1%	108.0%
	η <sub>int</sub> (%)	-6.5%	-1.8%	-2.0%	-2.0%	-2.1%	-2.1%	-2.1%

Q<sub>2</sub> refers to the daily extreme surface runoff for 2-years return period. Likewise, Q<sub>30</sub> is the daily extreme surface runoff for a 30-years return period. ΔQ<sub>t</sub> is the total change in surface runoff affected both by climate and land use changes. ΔQ<sub>lulc</sub> is the change in surface runoff caused by land use change, and ΔQ<sub>climate</sub> is the change in surface runoff caused by climate change.

#### 4. Discussion

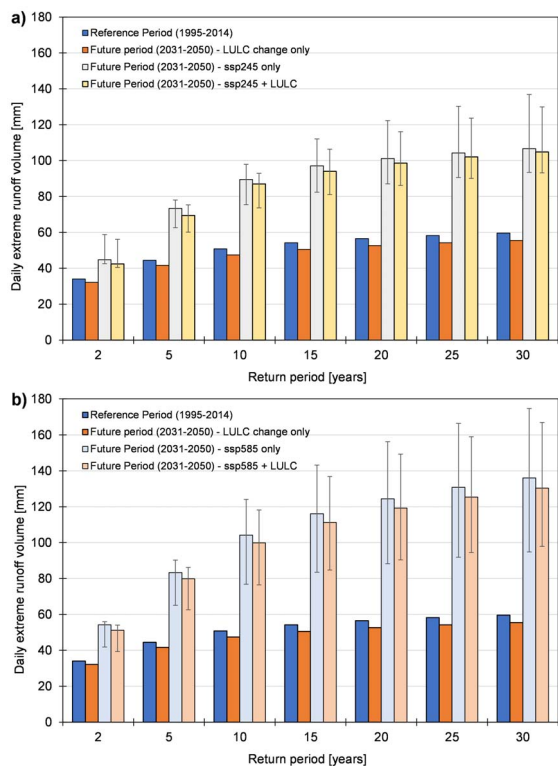
##### 4.1. On the contributions of climate and land use changes to surface runoff

In this study, using a hydrological modelling framework, the combined and isolated impact of climate and land use changes on surface runoff have been assessed between a baseline period 1995–2014 and a future period 2031–2050, under two warming scenarios (SSP2-4.5 and SSP5-8.5). The overall analysis results showed that rainfall and PET are projected to increase (under both SSPs), which is likely to cause an increase in actual evapotranspiration as well. Furthermore, the increase in rainfall is likely causing an increase in surface runoff, overriding the minor streamflow reduction caused by land use changes.

In previous studies, the hydrological response of West African watersheds to changes in climate and land use have been reported for past (historical) and current conditions. Many studies established that surface runoff generating mechanisms in the Sahel region are tightly dependent on soil surface conditions, hence explaining the dominant effect of land use changes over past and current climate conditions on changes in hydrological balance [Descroix et al.,

2018, Gal et al., 2017, Karambiri et al., 2011, Patuarel et al., 2017, Séguis et al., 2004, Yonaba et al., 2021a]. Yet, few studies have assessed the specific contributions of climate and land use changes on surface runoff variability and changes in West Africa.

Aich et al. [2015] used the Soil and Water Integrated Model (SWIM) with dynamic land use update to assess the relative contributions of climate and land use to the change in trends in floods magnitude. Without drawing a general answer, roughly equal shares of contributions for land use and climate contributions are reported for Gorouol and Sirba catchments. Sidibe et al. [2019] showed that in most of the West African Sahel, changes in land use are the major driver in surface runoff fluctuations over the period 1950–1990, further amplified by climate variability that occurred within the same period. Yonaba et al. [2021a] used the SWAT model in the Tougou watershed, over the historical period 1952–2005 and found that the fluctuations in surface runoff along with the hydrological paradox that occurred within the 1970s (likewise, in most Sahelian hydrosystems) could mainly be attributed to land use changes. Gbohoui et al. [2021], in the Nakanbe River Basin, reported similar conclusions using a Budyko-type separation approach; however, the same au-



**Figure 15.** Projected changes in maximum surface runoff quantiles under climate and land use changes. (a) Changes under SSP2-4.5 scenario. (b) Changes under SSP5-8.5 scenario. The error bars delineate the 90% confidence interval around the estimates across the models in the GCM ensemble.

thors observed an increasing contribution of climate-environment interaction effects on surface runoff fluctuations.

Regarding future climate projections over the Sahel area, essentially for rainfall, forecasts are mitigated, with high zonal contrasts. For the 2050 timeline, Mbaye *et al.* [2016] projected a decrease in rainfall over the western part of the Sahel; Tazen *et al.* [2013] projected a change in annual rainfall in the range of  $-3\%$  to  $+10\%$  over the Nakanbe River Basin; Badou *et al.* [2018] reported an increase in rainfall ( $1.7\%$  to  $23.4\%$ ), but with contrasting changes ( $-8.5\%$  to  $17.3\%$ ) over the Niger River basin. Recently, Dieng *et al.* [2022] projected a significant increase in rainfall between  $5\%$  to  $20\%$  for Sahara and Western Sahel regions using bias-corrected high-resolution

climate change simulations. Ultimately, there is still no consensus across studies on the direction in which rainfall is likely to evolve over the West African Sahel [Almazroui *et al.*, 2020], let alone the resulting impacts on hydrological processes [Stanzel *et al.*, 2018]. Also, large uncertainties across climate models are still persistent, bringing a supplemental layer of complexity in assessing future forecasts [Hattermann *et al.*, 2018, Laux *et al.*, 2021]. It should also be acknowledged that most of the previous studies use datasets issued from CMIP5 models. In this regard, this study sheds new insights in comparison, as it is among the first (to the best of our knowledge) to assess the latest available iteration of climate change simulations, that is CMIP6, in northern Burkina Faso.

In terms of future projections, this study reported an increase in surface runoff, more likely controlled by an increase in rainfall, although slightly mitigated by future land use changes. The dominant control of climate (i.e. rainfall) expected in the future is clearly at the opposite of what occurred in past and current conditions, and is likely attributable to: (1) the large increase in rainfall conditions; (2) the almost exclusive sensitivity of surface runoff to rainfall conditions; (3) the scale of observation which is, in this study, relatively small ( $37 \text{ km}^2$ ). As shown in Mounirou *et al.* [2020, 2021], surface runoff generating processes change with scale. Likewise, Gbohoui *et al.* [2021] showed that for varying nested watersheds scale, the control of environmental conditions (i.e., land use) is dependent on the size of the watershed. In contrast for example, in southwestern Burkina Faso, Idrissou *et al.* [2022] projected to the 2050 horizon an increase in surface runoff ( $12\%$  to  $95\%$ ) and also a decrease ( $24\%$  to  $44\%$ ), primarily driven by land use changes in both cases.

Regarding annual maximum surface runoff, an increase in the amplitude of runoff quantiles is reported in this study, largely caused by the increase in rainfall quantiles. These findings agree with previous studies, which reported an intensification of rainfall over the Sahel region, with a subsequent increase in surface runoff in most of the cases [Panthou *et al.*, 2018, Taylor *et al.*, 2017, Tazen *et al.*, 2019, Trambly *et al.*, 2020]. Hounkpè *et al.* [2019] showed that land use changes (especially the conversion of natural vegetation to croplands) could contribute to the increase in both the frequency and magnitude of floods, amplifying the contribution of the projected



rainfall increase.

#### 4.2. *Practical implications of this study*

In the light of the findings presented in this study, a few major practical implications to hydrologists and water resource managers can be drawn:

- (1) there is a need to assess, at larger and various scales and contexts, the direction of the future evolution of climate based on future climate projections. Yet, the cautious use of such data in impact studies is recommended, since they can be fraught with large discrepancies with past observations over their historical baselines [Dieng *et al.*, 2022]. Post-processing these simulations is always possible (bias-correction), even though the added value of such processing is questionable, and should be critically discussed depending on the impact study to be carried out [Laux *et al.*, 2021]. This also relates to the need for large sets of observations, which are scarce in the Sahelian context. Gridded datasets [Dembélé *et al.*, 2020, Satgé *et al.*, 2020], which provide estimations of rainfall from satellite observations, or bottom-up inversion from soil humidity [Brocca *et al.*, 2019, Yonaba *et al.*, 2022] can be considered as reliable alternatives.
- (2) There is also a growing need to include hydrological modelling at the core of water resource management. With the increase of available hydrological models of varying complexity, interesting insights can be drawn regarding how various processes and inputs affects a watershed balance, and therefore developing informed strategies for water resource management can be further considered [Karambiri *et al.*, 2011, Lèye *et al.*, 2021].
- (3) Also, regarding hydrological modelling, it is established that land use changes significantly affect surface runoff, especially in the Sahel region. Yet, in its current practice, most hydrological modelling applications still represent land use with static conditions, which leads to either inaccurate representation of spatial/temporal patterns of hydrological processes, or over/under parametrization of such models [Hounkpè *et al.*, 2019, Wagner

*et al.*, 2016, Yonaba *et al.*, 2021a]. Therefore, there is a need to develop integrated modelling frameworks, to fully account for dynamic land use changes during hydrological simulations, for more reliable outputs.

#### 4.3. *Limitations of this study*

Some of the limitations this study is fraught with should be acknowledged, to clarify the scope to which the reported results should be fully assessed.

First, for future rainfall projections, only the mean (or the median) of the model ensemble was considered most of the time, which is valid since the ensemble features a large number of models [Maraun, 2016]. However, an in-depth assessment of the uncertainty range around future projections should be carried out to draw a more complete picture of the possible contribution of climate change and land use to changes in surface runoff. In this study for example, uncertainties stem from multiple sources, including the GCM selection, the climate data pre-processing (statistical bias-correction), the land use/land cover maps (which embody spatial and structural uncertainties) and the hydrological model parametrization and simulations. The issue is acute in general to all impact studies which use a hydro-climatic modelling chain to assess the contributions of climate and land use on the water resource [Clark *et al.*, 2016]. In this study, in an effort to bridge this gap, uncertainties around GCM projections in rainfall and surface runoff are provided to some extent, based on aggregation of the individual models in the GCM ensemble. However, the uncertainty range around the hydrological model parameters range was not fully assessed on surface runoff estimates. However, dedicated research paths are proposing explicit approaches to decompose uncertainties (*i.e.* variance decomposition) according to its various sources to provide a better quantitative and qualitative insight regarding how trustful the projected impacts should be perceived [Bosshard *et al.*, 2013, Gao *et al.*, 2019].

Second, regarding future land use, a single business-as-usual narrative, was considered in this study [Yonaba *et al.*, 2021b]. This trajectory is actually closely similar to the land use narrative built within SSP2-4.5 [Popp *et al.*, 2017]. However, different land change scenarios should be assessed through the

framework used in this study, to provide a larger picture of the contribution of various land use narratives on surface runoff in the study area. Such future land use maps can be developed using statistical pattern-based approaches (cellular automata or artificial neural networks, for example), coupled with land demand models build around future scenarios: expansion of croplands to meet increasing food demand, mitigation policies to preserve natural vegetation for sustainability, etc. [Yonaba *et al.*, 2021b].

## 5. Conclusion

This study seeks to quantify the contribution of climate and land use to changes in surface runoff in the Tougou watershed, a typical Sahelian hydrosystem, located in the Upper Nakanbe River basin, in Burkina Faso (in West Africa). Between the baseline period (1995–2014) and the future period (2031–2050), a multi-model ensemble of 9 GCMs (issued from CMIP6 simulations) is used to drive a calibrated hydrological SWAT model, along with a dynamic land use update of LULC maps (for past and future conditions). An increase in annual rainfall (13.7% and 18.8%, respectively under SSP2-4.5 and SSP5-8.5 scenarios) and also in daily maximum rainfall (from 26.6% to 45.1% under SSP2-4.5 and from 35.5% to 77.2% under SSP5-8.5 scenarios), is projected. This increase is found to be the major cause of a significant increase in daily surface runoff, as shown by the rise of flow duration curve. The annual streamflow is projected to increase (24.2% and 34.3% under SSP2-4.5 and SSP5-8.5 respectively). In terms of relative contributions, land use change (mainly the conversion of bare lands to croplands) is expected to cause a decrease in annual surface runoff, with a contribution of –27.6% to –19.5% (under SSP2-4.5 and SSP5-8.5 scenarios respectively), and also in maximum daily surface runoff (SSP2-4.5: –22.8% to –9.1% and SSP5-8.5 11.1% to –5.8%). In contrast, the contribution of climate (i.e., the rainfall) to the annual surface runoff increase is dominant, with a contribution of +128.8% and +120.7% (respectively under SSP2-4.5 and SSP5-8.5 scenarios), and likewise for maximum daily surface runoff (SSP2-4.5: +127.2% to +104.0%; SSP5-8.5: +117.7% to +108.0%).

Overall, these results shed light on the isolated contributions of climate and land use, and their interaction to the changes in hydrological processes. It

also shows that such processes stem from the complex nature and interplay between climate and land change (both including natural and anthropogenic change), which should be fully assessed and integrated within an integrated modelling framework for more realistic and more reliable outputs. Finally, as the increase in surface runoff in the Tougou watershed is likely, future strategies should focus on: (i) mitigation of associated risks such as flood control; (ii) surface runoff management through farming practices (*zaï*, half-moons, stone rows) to increase soil water content for agricultural production; (iii) and finally, mitigation of runoff erosion. The results of this study might significantly help in framing such policies, as it unveils increased risks relating to climate change in the future, which could be effectively mitigated through local and effective land use decisions.

## Conflicts of interest

Authors have no conflict of interest to declare.

## Acknowledgements

This study has been carried out under the financial support of the World Bank, through the Government of Burkina Faso (grant IDA n° 5420-BF). The authors are also extremely grateful to the local community of the Tougou village (in Burkina Faso) for their helpful assistance and cooperation; especially since this local community is nowadays hardly affected by the ongoing security crisis in Burkina Faso.

## References

- Abbaspour, K. C., Vajdani, M., Haghghat, S., and Yang, J. (2007). SWAT-CUP calibration and uncertainty programs for SWAT. In *MODSIM 2007 International Congress on Modelling and Simulation, Modelling and Simulation Society of Australia and New Zealand*, pages 1596–1602.
- Aghsaee, H., Mobarghaee Dinan, N., Moridi, A., Asadolahi, Z., Delavar, M., Fohrer, N., and Wagner, P. D. (2020). Effects of dynamic land use/land cover change on water resources and sediment yield in the Anzali wetland catchment, Gilan, Iran. *Sci. Total Environ.*, 712, article no. 136449.

- Aich, V., Liersch, S., Vetter, T., Andersson, J., Müller, E., and Hattermann, F. (2015). Climate or land use?—Attribution of changes in river flooding in the sahel zone. *Water*, 7, 2796–2820.
- Akoko, G., Le, T. H., Gomi, T., and Kato, T. (2021). A review of SWAT model application in Africa. *Water*, 13, article no. 1313.
- Almazroui, M., Saeed, F., Saeed, S., Nazrul Islam, M., Ismail, M., Klutse, N. A. B., and Siddiqui, M. H. (2020). Projected change in temperature and precipitation over Africa from CMIP6. *Earth Syst. Environ.*, 4, 455–475.
- Amogu, O., Esteves, M., Vandervaere, J.-P., Malam Abdou, M., Panthou, G., Rajot, J.-L., Souley Yéro, K., Boubkraoui, S., Lapetite, J.-M., Dessay, N., Zin, I., Bachir, A., Bouzou Moussa, I., Faran Maïga, O., Gautier, E., Mamadou, I., and Descroix, L. (2015). Runoff evolution due to land-use change in a small Sahelian catchment. *Hydrol. Sci. J.*, 60, 78–95.
- Andréassian, V., Coron, L., Lerat, J., and Le Moine, N. (2016). Climate elasticity of streamflow revisited—an elasticity index based on long-term hydrometeorological records. *Hydrol. Earth Syst. Sci.*, 20, 4503–4524.
- Angelina, A., Gado Djibo, A., Seidou, O., Seidou Sanda, I., and Sittichok, K. (2015). Changes to flow regime on the Niger River at Koulikoro under a changing climate. *Hydrol. Sci. J.*, 60, 1709–1723.
- Arnold, J., Srinivasan, R., Muttiyah, R. S., and Williams, J. R. (1998). Large area hydrologic modeling and assessment Part I: model development. *J. Am. Water Res. Assoc.*, 34, 73–89.
- Ayugi, B., Jiang, Z., Iyakaremye, V., Ngoma, H., Babaousmail, H., Onyutha, C., Dike, V. N., Mumo, R., and Ongoma, V. (2022). East African population exposure to precipitation extremes under 1.5 °C and 2.0 °C warming levels based on CMIP6 models. *Environ. Res. Lett.*, 17(4), article no. 044051.
- Badou, D. F., Diekrüger, B., Kapangaziwiri, E., Mbaye, M. L., Yira, Y., Lawin, E. A., Oyerinde, G. T., and Afouda, A. (2018). Modelling blue and green water availability under climate change in the Beninese Basin of the Niger River Basin, West Africa. *Hydrol. Process.*, 32, 2526–2542.
- Belemtougri, A. P., Ducharne, A., Tazen, F., Oudin, L., and Karambiri, H. (2021). Understanding key factors controlling the duration of river flow intermittency: Case of Burkina Faso in West Africa. *J. Hydrol.: Reg. Stud.*, 37, article no. 100908.
- Bhagat, S. K., Tiyasha, T., Al-khafaji, Z., Laux, P., Ewees, A. A., Rashid, T. A., Salih, S., Yonaba, R., Beyaztas, U., and Yaseen, Z. M. (2022). Establishment of dynamic evolving neural-fuzzy inference system model for natural air temperature prediction. *Complexity*, 2022, article no. 1047309.
- Biasutti, M. (2019). Rainfall trends in the African Sahel: Characteristics, processes, and causes. *Wiley Interdiscip. Rev.: Clim. Change*, 10, article no. e591.
- Bosshard, T., Carambia, M., Goergen, K., Kotlarski, S., Krahe, P., Zappa, M., and Schär, C. (2013). Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections: uncertainty sources in climate-impact projections. *Water Resour. Res.*, 49, 1523–1536.
- Brocca, L., Filippucci, P., Hahn, S., Ciabatta, L., Mas-sari, C., Camici, S., Schüller, L., Bojkov, B., and Wagner, W. (2019). SM2RAIN–ASCAT (2007–2018): global daily satellite rainfall data from ASCAT soil moisture observations. *Earth Syst. Sci. Data*, 11, 1583–1601.
- Cannon, A. J. (2018). Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables. *Clim. Dyn.*, 50, 31–49.
- Cannon, A. J., Sobie, S. R., and Murdock, T. Q. (2015). Bias correction of GCM precipitation by quantile mapping: how well do methods preserve changes in quantiles and extremes? *J. Clim.*, 28, 6938–6959.
- Clark, M. P., Wilby, R. L., Gutmann, E. D., Vano, J. A., Gangopadhyay, S., Wood, A. W., Fowler, H. J., Prudhomme, C., Arnold, J. R., and Brekke, L. D. (2016). Characterizing uncertainty of the hydrologic impacts of climate change. *Curr. Clim. Change Rep.*, 2, 55–64.
- Connors, S., Dionne, M., Hanak, G., Musulin, R., Aellen, N., Amjad, M., Bowen, S., Carrascal, D. R., Coppola, E., Moro, E. D., Dosio, A., Faria, S. H., Gan, T. Y., Gomis, M., Gutierrez, J. M., Hope, P., Kopp, R., Krakovska, S., Leitzell, K., Maraun, D., Masson-Delmotte, V., Matthews, R., Maycock, T., Paddam, S., Plattner, G.-K., Pui, A., Rahimi, M., Ranasinghe, R., Rogelj, J., Ruane, A. C., Szopa, S., Turner, A., Vautard, R., Velichkova, Y., Weigel, A., and Zhang, X. (2022). *Climate Science: A Summary for Actuaries—What the IPCC Climate Change Report 2021 Means for the Actuarial Profession (Report)*. International Actuarial Association, Ottawa, Canada.

- Cornelissen, T., Diekkrüger, B., and Giertz, S. (2013). A comparison of hydrological models for assessing the impact of land use and climate change on discharge in a tropical catchment. *J. Hydrol.*, 498, 221–236.
- de Marsily, G. (2007). An overview of the world's water resources problems in 2050. *Ecohydrol. Hydrobiol.*, 7, 147–155.
- Dembélé, M., Schaeffli, B., van de Giesen, N., and Mariéthoz, G. (2020). Suitability of 17 gridded rainfall and temperature datasets for large-scale hydrological modelling in West Africa. *Hydrol. Earth Syst. Sci.*, 24, 5379–5406.
- Dembélé, M., Vrac, M., Ceperley, N., Zwart, S. J., Larsen, J., Dadson, S. J., Mariéthoz, G., and Schaeffli, B. (2022). Contrasting changes in hydrological processes of the Volta River basin under global warming. *Hydrol. Earth Syst. Sci.*, 26, 1481–1506.
- Descroix, L., Guichard, F., Grippa, M., Lambert, L., Panthou, G., Mahé, G., Gal, L., Dardel, C., Quantin, G., Kergoat, L., Bouaïta, Y., Hiernaux, P., Vischel, T., Pellarin, T., Faty, B., Wilcox, C., Malam Abdou, M., Mamadou, I., Vandervaere, J.-P., Diongue-Niang, A., Ndiaye, O., Sané, Y., Dacosta, H., Gosset, M., Cassé, C., Sultan, B., Barry, A., Amogu, O., Nka Nnomo, B., Barry, A., and Paturel, J.-E. (2018). Evolution of surface hydrology in the Sahelo-Sudanian strip: an updated review. *Water*, 10, article no. 748.
- Dey, P. and Mishra, A. (2017). Separating the impacts of climate change and human activities on streamflow: A review of methodologies and critical assumptions. *J. Hydrol.*, 548, 278–290.
- Diedhiou, A., Bichet, A., Wartenburger, R., Seneviratne, S. I., Rowell, D. P., Sylla, M. B., Diallo, I., Todzo, S., Touré, N. E., Camara, M., Ngatchah, B. N., Kane, N. A., Tall, L., and Affholder, F. (2018). Changes in climate extremes over West and Central Africa at 1.5 °C and 2 °C global warming. *Environ. Res. Lett.*, 13, article no. 065020.
- Dieng, D., Cannon, A. J., Laux, P., Hald, C., Adeyeri, O., Rahimi, J., Srivastava, A. K., Mbaye, M. L., and Kunstmann, H. (2022). Multivariate bias-correction of high-resolution regional climate change simulations for West Africa: performance and climate change implications. *JGR Atmos.*, 127(5), article no. e2021JD034836.
- Dosio, A., Turner, A. G., Tamoffo Tchio, A., Sylla, M. B., Lennard, C., Jones, R., Terray, L., Nikulin, G., and Hewitson, B. (2020). A tale of two futures: contrasting scenarios of future precipitation for West Africa from an ensemble of regional climate models. *Environ. Res. Lett.*, 15(6), article no. 064007.
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E. (2016). Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.*, 9, 1937–1958.
- Fovet, O., Belemtougri, A., Boithias, L., Braud, I., Charlier, J., Cottet, M., Daudin, K., Dramais, G., Ducharne, A., Folton, N., Grippa, M., Hector, B., Kuppel, S., Le Coz, J., Legal, L., Martin, P., Moatar, F., Molénat, J., Probst, A., Riotte, J., Vidal, J., Vinatier, F., and Detry, T. (2021). Intermittent rivers and ephemeral streams: Perspectives for critical zone science and research on socio-ecosystems. *WIREs Water*, 8(4), article no. e1523.
- Gal, L., Grippa, M., Hiernaux, P., Pons, L., and Kergoat, L. (2017). The paradoxical evolution of runoff in the pastoral Sahel: analysis of the hydrological changes over the Agoufou watershed (Mali) using the KINEROS-2 model. *Hydrol. Earth Syst. Sci.*, 21, 4591–4613.
- Gao, J., Sheshukov, A. Y., Yen, H., Douglas-Mankin, K. R., White, M. J., and Arnold, J. G. (2019). Uncertainty of hydrologic processes caused by bias-corrected CMIP5 climate change projections with alternative historical data sources. *J. Hydrol.*, 568, 551–561.
- Gbohoui, Y. P., Paturel, J.-E., Tazen, F., Mounirou, L. A., Yonaba, R., Karambiri, H., and Yacouba, H. (2021). Impacts of climate and environmental changes on water resources: A multi-scale study based on Nakanbé nested watersheds in West African Sahel. *J. Hydrol.: Reg. Stud.*, 35, article no. 100828.
- Gbohoui, Y. P., Yonaba, R., Fowé, T., Paturel, J.-E., Karambiri, H., and Yacouba, H. (2022). Comparison of One-Site vs Multi-Sites calibration/validation schemes for hydrological modelling of nested watersheds in the West African Sahel (No. IAHS2022-69). In *Presented at the IAHS2022, Copernicus Meetings, Montpellier, France*.
- Grippa, M., Kergoat, L., Boone, A., Peugeot, C., Demarty, J., Cappelaere, B., Gal, L., Hiernaux, P., Mougin, E., Ducharne, A., Dutra, E., Anderson, M., Hain, C., and ALMIP2 working group (2017). Mod-

- eling surface runoff and water fluxes over contrasted soils in the pastoral sahel: evaluation of the ALMIP2 land surface models over the Gourma Region in Mali. *J. Hydrometeorol.*, 18, 1847–1866.
- Hargreaves, G. H. and Samani, Z. A. (1985). Reference crop evapotranspiration from temperature. *Appl. Eng. Agric.*, 1, 96–99.
- Hattermann, F. F., Vetter, T., Breuer, L., Su, B., Daggupati, P., Donnelly, C., Fekete, B., Flörke, F., Gosling, S. N., Hoffmann, P., Liersch, S., Masaki, Y., Motovilov, Y., Müller, C., Samaniego, L., Stacke, T., Wada, Y., Yang, T., and Krysnova, V. (2018). Sources of uncertainty in hydrological climate impact assessment: a cross-scale study. *Environ. Res. Lett.*, 13, article no. 015006.
- Hounkpè, J., Diekrüger, B., Afouda, A. A., and Sintondji, L. O. C. (2019). Land use change increases flood hazard: a multi-modelling approach to assess change in flood characteristics driven by socio-economic land use change scenarios. *Nat. Hazards*, 98, 1021–1050.
- Ibrahim, B. (2002). Analyse de la variabilité climatique au Burkina Faso au cours de la seconde moitié du 20ème siècle - Sécheresse info. <http://www.secheresse.info/spip.php?article55865>. Master thesis, Institut International d'Ingénierie de l'Eau et de l'Environnement.
- Idrissou, M., Diekrüger, B., Tischbein, B., Op de Hipt, F., Näschen, K., Poméon, T., Yira, Y., and Ibrahim, B. (2022). Modeling the impact of climate and land use/land cover change on water availability in an inland valley catchment in Burkina Faso. *Hydrology*, 9, article no. 12.
- IGB (2002). Base de données d'Occupation des Terres (BDOT) 2002. <https://www.unoosa.org/documents/pdf/psa/activities/2007/morocco/presentations/4-6.pdf>. Burkina Faso.
- IPCC (2022). Summary for policymakers. In Pörtner, H. O., Roberts, D. C., Tignor, M., Poloczanska, E. S., Mintenbeck, K., Alegria, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., and Rama, B., editors, *Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 3–33. Cambridge University Press, Cambridge, UK and New York, NY, USA.
- Kafando, M. B., Koïta, M., Le Coz, M., Yonaba, O. R., Fowe, T., Zouré, C. O., Faye, M. D., and Leye, B. (2021). Use of multidisciplinary approaches for groundwater recharge mechanism characterization in basement aquifers: case of sanon experimental catchment in Burkina Faso. *Water*, 13, article no. 3216.
- Karambiri, H., García Galiano, S. G., Giraldo, J. D., Yacouba, H., Ibrahim, B., Barbier, B., and Polcher, J. (2011). Assessing the impact of climate variability and climate change on runoff in West Africa: the case of Senegal and Nakambe River basins. *Atmos. Sci. Lett.*, 12, 109–115.
- Karambiri, H., Ribolzi, O., Delhoume, J. P., Ducloux, J., Coudrain-Ribstein, A., and Casenave, A. (2003). Importance of soil surface characteristics on water erosion in a small grazed Sahelian catchment. *Hydrol. Process.*, 17, 1495–1507.
- Laux, P., Rötter, R. P., Webber, H., Dieng, D., Rahimi, J., Wei, J., Faye, B., Srivastava, A. K., Bliefernicht, J., Adeyeri, O., Arnault, J., and Kunstmann, H. (2021). To bias correct or not to bias correct? An agricultural impact modelers' perspective on regional climate model data. *Agric. For. Meteorol.*, 304–305, article no. 108406.
- Lèye, B., Zouré, C. O., Yonaba, R., and Karambiri, H. (2021). Water resources in the sahel and adaptation of agriculture to climate change: Burkina Faso. In Diop, S., Scheren, P., and Niang, A., editors, *Climate Change and Water Resources in Africa*, pages 309–331. Springer International Publishing, Cham.
- Li, Z., Li, Z., Zhao, W., and Wang, Y. (2015). Probability modeling of precipitation extremes over two river Basins in Northwest of China. *Adv. Meteorol.*, 2015, article no. 374127.
- Mahé, G. and Paturel, J.-E. (2009). 1896–2006 Sahelian annual rainfall variability and runoff increase of Sahelian Rivers. *C. R. Geosci.*, 341, 538–546.
- Maraun, D. (2016). Bias correcting climate change simulations—a critical review. *Curr. Clim. Change Rep.*, 2, 211–220.
- Mbaye, M. L., Haensler, A., Hagemann, S., Gaye, A. T., Moseley, C., and Afouda, A. (2016). Impact of statistical bias correction on the projected climate change signals of the regional climate model REMO over the Senegal River Basin. *Int. J. Climatol.*, 36, 2035–2049.
- Moriasi, D. N., Pai, N., Steiner, J. L., Gowda, P. H., Winchell, M., Rathjens, H., Starks, P. J., and Verser, J. A. (2019). SWAT-LUT: a desktop graphical user interface for updating Land Use in SWAT. *J. Am.*

- Water Resour. Assoc.*, 55, 1102–1115.
- Mounirou, L. A. (2012). *Etude du ruissellement et de l'érosion à différentes échelles spatiales sur le bassin versant de Tougou en zone sahélienne du Burkina Faso: quantification et transposition des données*. PhD thesis, Université des sciences et techniques de Montpellier 2.
- Mounirou, L. A., Yonaba, R., Koïta, M., Paturel, J.-E., Mahé, G., Yacouba, H., and Karambiri, H. (2021). Hydrologic similarity: Dimensionless runoff indices across scales in a semi-arid catchment. *J. Arid Environ.*, 193, article no. 104590.
- Mounirou, L. A., Zouré, C. O., Yonaba, R., Paturel, J.-E., Mahé, G., Niang, D., Yacouba, H., and Karambiri, H. (2020). Multi-scale analysis of runoff from a statistical perspective in a small Sahelian catchment under semi-arid climate. *Arab. J. Geosci.*, 13, article no. 154.
- Ndiaye, P. M., Bodian, A., Diop, L., and Djaman, K. (2017). Évaluation de vingt méthodes d'estimation de l'évapotranspiration journalière de référence au Burkina Faso. *Physio-Géo*, 11, 129–146.
- Neitsch, S., Arnold, J., Kiniry, J. R., and Williams, J. R. (2011). *Soil and Water Assessment Tool theoretical documentation: Version 2009*. Texas Water Resources Institute, Texas, USA.
- Nyamekye, C., Thiel, M., Schönbrodt-Stitt, S., Zoungrana, B., and Amekudzi, L. (2018). Soil and water conservation in Burkina Faso, West Africa. *Sustainability*, 10, article no. 3182.
- Olivera, S. and Heard, C. (2019). Increases in the extreme rainfall events: Using the Weibull distribution. *Environmetrics*, 30, article no. e2532.
- Onyutha, C., Turyahabwe, C., and Kaweesa, P. (2021). Impacts of climate variability and changing land use/land cover on River Mpanga flows in Uganda, East Africa. *Environ. Challenges*, 5, article no. 100273.
- Pai, N. and Saraswat, D. (2011). SWAT2009\_LUC: a tool to activate the land use change module in SWAT 2009. *Trans. ASABE*, 54, 1649–1658.
- Panthou, G., Lebel, T., Vischel, T., Quantin, G., Sane, Y., Ba, A., Ndiaye, O., Diongue-Niang, A., and Diopkane, M. (2018). Rainfall intensification in tropical semi-arid regions: the Sahelian case. *Environ. Res. Lett.*, 13, article no. 064013.
- Paturel, J. E., Mahé, G., Diello, P., Barbier, B., Dezetter, A., Dieulin, C., Karambiri, H., Yacouba, H., and Maiga, A. (2017). Using land cover changes and demographic data to improve hydrological modeling in the Sahel: Improving hydrological modelling in the Sahel. *Hydrol. Process.*, 31, 811–824.
- Popp, A., Calvin, K., Fujimori, S., Havlik, P., Humpenöder, F., Stehfest, E., Bodirsky, B. L., Dietrich, J. P., Doelmann, J. C., Gusti, M., Hasegawa, T., Kyle, P., Obersteiner, M., Tabeau, A., Takahashi, K., Valin, H., Waldhoff, S., Weindl, I., Wise, M., Kriegler, E., Lotze-Campen, H., Fricko, O., Riahi, K., and van Vuuren, D. P. (2017). Land-use futures in the shared socio-economic pathways. *Glob. Environ. Change*, 42, 331–345.
- Raziei, T. and Pereira, L. S. (2013). Estimation of ETo with Hargreaves–Samani and FAO-PM temperature methods for a wide range of climates in Iran. *Agric. Water Manage.*, 121, 1–18.
- Rusagara, R., Koïta, M., Plagnes, V., and Jost, A. (2022). Groundwater recharge pathways to a weathered-rock aquifer system in a dryland catchment in Burkina Faso. *Hydrogeol. J.*, 30, 1489–1512.
- Satgé, E., Defrance, D., Sultan, B., Bonnet, M.-P., Seyler, F., Rouché, N., Pierron, F., and Paturel, J.-E. (2020). Evaluation of 23 gridded precipitation datasets across West Africa. *J. Hydrol.*, 581, article no. 124412.
- Sawadogo, B. and Barro, D. (2022). Space-time trend detection and dependence modeling in extreme event approaches by functional peaks-over-thresholds: application to precipitation in Burkina Faso. *Int. J. Math. Math. Sci.*, 2022, article no. 2608270.
- Séguis, L., Cappelaere, B., Milési, G., Peugeot, C., Massuel, S., and Favreau, G. (2004). Simulated impacts of climate change and land-clearing on runoff from a small Sahelian catchment: Climatic and anthropogenic effects on Sahelian runoff. *Hydrol. Process.*, 18, 3401–3413.
- Serdeczny, O., Adams, S., Baarsch, F., Coumou, D., Robinson, A., Hare, W., Schaeffer, M., Perrette, M., and Reinhardt, J. (2017). Climate change impacts in Sub-Saharan Africa: from physical changes to their social repercussions. *Reg. Environ. Change*, 17, 1585–1600.
- Sidibe, M., Dieppois, B., Eden, J., Mahé, G., Paturel, J.-E., Amoussou, E., Anifowose, B., and Lawler, D. (2019). Interannual to Multi-decadal streamflow variability in West and Central Africa: Interactions with catchment properties and large-scale climate variability. *Glob. Planet. Change*, 177, 141–156.

- Stanzel, P., Kling, H., and Bauer, H. (2018). Climate change impact on West African rivers under an ensemble of CORDEX climate projections. *Clim. Serv.*, 11, 36–48.
- Sylla, M. B., Pal, J. S., Faye, A., Dimobe, K., and Kuntmann, H. (2018). Climate change to severely impact West African basin scale irrigation in 2 °C and 1.5 °C global warming scenarios. *Sci. Rep.*, 8, article no. 14395.
- Taylor, C. M., Belušić, D., Guichard, F., Parker, D. J., Vischel, T., Bock, O., Harris, P. P., Janicot, S., Klein, C., and Panthou, G. (2017). Frequency of extreme Sahelian storms tripled since 1982 in satellite observations. *Nature*, 544, 475–478.
- Tazen, F., Diarra, A., Kabore, R. F. W., Ibrahim, B., Bologo/Traoré, M., Traoré, K., and Karambiri, H. (2019). Trends in flood events and their relationship to extreme rainfall in an urban area of Sahelian West Africa: The case study of Ouagadougou, Burkina Faso. *J. Flood Risk Manage.*, 12, article no. e12507.
- Tazen, F., Paturol, J.-E., and Karambiri, H. (2013). Impacts des changements globaux sur les ressources en eau dans la zone sahélienne en Afrique de l'Ouest. In *IAHS-AISH Publication*, pages 99–104. International Association of Hydrological Sciences, Gothenburg, Sweden.
- Todzo, S., Bichet, A., and Diedhiou, A. (2020). Intensification of the hydrological cycle expected in West Africa over the 21st century. *Earth Syst. Dynam.*, 11, 319–328.
- Tomer, M. D. and Schilling, K. E. (2009). A simple approach to distinguish land-use and climate-change effects on watershed hydrology. *J. Hydrol.*, 376, 24–33.
- Tramblay, Y., Villarini, G., and Zhang, W. (2020). Observed changes in flood hazard in Africa. *Environ. Res. Lett.*, 15, article no. 1040b5.
- Valentin, C. (2018). Soil surface crusting of soil and water harvesting. In Valentin, C., editor, *Soils as a Key Component of the Critical Zone 5*, pages 21–38. John Wiley & Sons, Inc., Hoboken, NJ, USA.
- Wagner, P. D., Bhallamudi, S. M., Narasimhan, B., Kantakumar, L. N., Sudheer, K. P., Kumar, S., Schneider, K., and Fiener, P. (2016). Dynamic integration of land use changes in a hydrologic assessment of a rapidly developing Indian catchment. *Sci. Total Environ.*, 539, 153–164.
- Wallach, D., Mearns, L. O., Ruane, A. C., Rötter, R. P., and Asseng, S. (2016). Lessons from climate modeling on the design and use of ensembles for crop modeling. *Clim. Chang.*, 139, 551–564.
- Yang, Q., Wang, Q. J., Hakala, K., and Tang, Y. (2021). Bias-correcting input variables enhances forecasting of reference crop evapotranspiration. *Hydrol. Earth Syst. Sci.*, 25, 4773–4788.
- Yin, J., He, F., Xiong, Y. J., and Qiu, G. Y. (2017). Effects of land use/land cover and climate changes on surface runoff in a semi-humid and semi-arid transition zone in northwest China. *Hydrol. Earth Syst. Sci.*, 21, 183–196.
- Yira, Y., Diekkrüger, B., Steup, G., and Bossa, A. Y. (2017). Impact of climate change on hydrological conditions in a tropical West African catchment using an ensemble of climate simulations. *Hydrol. Earth Syst. Sci.*, 21, 2143–2161.
- Yonaba, R., Belemtougri, A., Tazen, F., Mounirou, L. A., Koïta, M., Karambiri, H., and Yacouba, H. (2022). Assessing the accuracy of SM2RAIN (Soil Moisture to Rainfall) products in poorly gauged countries: the case of Burkina Faso in the West African Sahel. (No. IAHS2022-263). In *Presented at the IAHS2022, Copernicus Meetings, Montpellier, France*.
- Yonaba, R., Biaou, A. C., Koïta, M., Tazen, F., Mounirou, L. A., Zouré, C. O., Queloz, P., Karambiri, H., and Yacouba, H. (2021a). A dynamic land use/land cover input helps in picturing the Sahelian paradox: Assessing variability and attribution of changes in surface runoff in a Sahelian watershed. *Sci. Total Environ.*, 757, article no. 143792.
- Yonaba, R., Koïta, M., Mounirou, L. A., Tazen, F., Queloz, P., Biaou, A. C., Niang, D., Zouré, C., Karambiri, H., and Yacouba, H. (2021b). Spatial and transient modelling of land use/land cover (LULC) dynamics in a Sahelian landscape under semi-arid climate in northern Burkina Faso. *Land Use Policy*, 103, article no. 105305.
- Yonaba, R. O. (2020). *Dynamique spatio-temporelle des états de surface et influence sur le ruissellement sur un bassin de type sahélien: cas du bassin de Tougou (Nord Burkina Faso)*. PhD thesis, Institut International d'Ingénierie de l'Eau et de l'Environnement.
- Zipper, S. C., Motew, M., Booth, E. G., Chen, X., Qiu, J., Kucharik, C. J., Carpenter, S. R., and Loheide II, S. P. (2018). Continuous separation of land use and climate effects on the past and future water

- balance. *J. Hydrol.*, 565, 106–122.
- Zouré, C., Queloz, P., Koïta, M., Niang, D., Fowé, T., Yonaba, R., Consuegra, D., Yacouba, H., and Karambiri, H. (2019). Modelling the water balance on farming practices at plot scale: Case study of Tougou watershed in Northern Burkina Faso. *Catena*, 173, 59–70.
- Zouré, C. O. (2019). Étude des performances hydrologiques des techniques culturales dans un contexte de changement climatique en zone sahélienne du Burkina Faso. PhD thesis, Institut International d'Ingénierie de l'Eau et de l'Environnement.





Research article

## Geo-hydrological Data & Models

# Impacts of climate change and pumping on groundwater resources in the Kou River basin, Burkina Faso

Justine Tirogo<sup>✉,†,a</sup>, Anne Jost<sup>✉,\*,b</sup>, Angelbert Biaou<sup>✉,c</sup>, Youssouf Koussoubé<sup>✉,d</sup>, Pierre Ribstein<sup>✉,b</sup> and Denis Dakouré<sup>e</sup>

<sup>a</sup> Programme d'Approvisionnement en Eau et d'Assainissement, Ouagadougou, Burkina Faso

<sup>b</sup> Sorbonne Université, CNRS, EPHE, Metis, F-75005 Paris, France

<sup>c</sup> Institut International d'Ingénierie de l'Eau et l'Environnement, Ouagadougou, Burkina Faso

<sup>d</sup> Université Ouaga I Professeur Joseph KI-ZERBO, Ouagadougou, Burkina Faso

<sup>e</sup> Agence Française de Développement, Ouagadougou, Burkina Faso

*E-mails:* anne.jost@sorbonne-universite.fr (A. Jost), angelbert.biaou@zie-edu.org (A. Biaou), youssouf.koussoubé@gmail.com (Y. Koussoubé), pierre.ribstein@sorbonne-universite.fr (P. Ribstein), dakoure\_dc@hotmail.com (D. Dakouré)

*This article results from the PhD thesis of Justine Tirogo, who wrote the submitted version. Justine passed away suddenly in mid-December 2022, while she was working on the revision of the article. We wanted to publish this paper in her honour and to pay tribute to the high quality and rigour of her scientific work.*

**Abstract.** West Africa is subject to climatic variability with a long period of drought during the 1970s–1990s whose impact on groundwater remains poorly studied. This work focuses on the Kou basin in Burkina Faso, which holds a large groundwater resource resulting in exceptional springs. This resource shows signs of depletion that raise a critical question: What is the share of climate and that of withdrawals in the depletion of the resource? To answer this question, we used a hydrogeological model calibrated in steady state and in transient regime (monthly time step 1995–2014; annual time step 1961–2014). The results showed that pumping has a significant impact on the resource, especially in the vicinity of the pumping areas, and caused approximately 30% of the observed decrease in spring discharge. Drought periods and consecutive years without recharge also led to a decrease in groundwater levels. Thus, precipitation deficits and variability, combined with increasing pumping, have endangered the resource.

**Keywords.** Groundwater pumping, Groundwater modelling, Climate change, Sudanian region, Sedimentary aquifer, West Africa.

*Manuscript received 16 May 2022, revised 17 February 2023 and 16 March 2023, accepted 20 March 2023.*

\* Corresponding author.

† Deceased.

## 1. Introduction

Population growth leading to increased agricultural production [de Marsily, 2008, 2020] and demand for drinking water [He *et al.*, 2021] combined with climate change and variability poses a real threat to the availability of water resources worldwide and in Africa in particular [UN-WWDR, 2021, Sylla *et al.*, 2018, for West Africa]. Despite major disparities related to differences in physical environments and the climatic mosaic of the continent, the decline of a quality water resource is a reality and a threat in the most vulnerable arid to sub-humid areas, where more than 40% of the population lives.

West Africa is one of the regions most affected by climate variability and change. During the 20th century, the West African region experienced a succession of wet and dry periods marked by significant rainfall deficits of varying intensity depending on the sub-region, particularly since the 1970s [Nicholson, 1978, Servat *et al.*, 1997, Paturel *et al.*, 1997, 2002, Lebel and Ali, 2009, De Longueville *et al.*, 2016, Nicholson *et al.*, 2018]. There is a contrasting response of hydrological systems to this variability in rainfall between the Sahel region (annual rainfall < 700 mm) and the Sudanian region (annual rainfall > 700 mm) [Descroix *et al.*, 2009, Nka *et al.*, 2015]. In the Sahel region, an increase in runoff in most small and medium-sized watersheds is noted [Albergel, 1987, Descroix *et al.*, 2009, Amogu *et al.*, 2010]. In the Sudanian region, there has been a clearly greater decrease in runoff than in rainfall [Mahé, 2009, Descroix *et al.*, 2015]. The increase in runoff in the Sahelian zone is linked to a change in land use that has also caused an increase in the water table in endoreic basins such as those in southwest Niger [Leduc *et al.*, 2001, Favreau *et al.*, 2009]. In the Sudanian area, the significant decrease in river discharge is related to a decline in base flow caused by the drop in the water table [Mahé, 2009, Tirogo *et al.*, 2016].

However, in West Africa, little is known about the response of groundwater to climate-induced variations. Studies that have addressed this issue have shown that the relationships between precipitation and water table recharge are complex to determine because they are highly variable in space and dependent on several factors [Tirogo *et al.*, 2016, Cuthbert *et al.*, 2019, Ascott *et al.*, 2020].

These factors include anthropogenic influence (dam construction, changes in land use, increased water withdrawals) and natural factors such as the surface water drainage network as well as the spatial distribution of hydraulic properties and geometry of the aquifer. In addition to this variability in the evolution of the water table, there is also uncertainty in the estimation of the future climate of the area [e.g., Monerie *et al.*, 2020]. In a context where the demand for water continues to grow, groundwater resources, which are of paramount importance for water supply to populations, socio-economic activities and vital ecosystems [Calow and MacDonald, 2009, USAID, 2014], warrant better control for sustainable management.

In the present study, we focus on the groundwater resource of the Kou basin (1860 km<sup>2</sup>) in the western sedimentary basin of Burkina Faso. This groundwater is a major socio-economic asset, at the origin of important springs such as those of Nasso/Guinguette (three main springs), which have an exceptional discharge compared to that of the region ( $\approx 1.5$  m<sup>3</sup>/s for the Guinguette spring during the low-flow period in 2011). The groundwater resource is used for irrigation, but also by various industries and for the supply of drinking water to many communities, including the country's second largest city, Bobo-Dioulasso. Water needs at the basin level are constantly increasing while the groundwater level is dropping (ca. 0.5 m between 1990 and 2015) as is the discharge rate of springs. Through geochemical and isotopic analyses, Dakouré [2003] and Huneau *et al.* [2011] showed there is a predominance of ancient waters in the southeast part of the much larger Taoudeni sedimentary basin with a low renewal rate. In a more recent geochemical study, Kouanda [2019] nevertheless reported on the existence of significant ongoing recharge in the same area. In this context where the water table is influenced by the rainfall variability and by strong demand, the question arises as to what would be the cause of the decline in the discharge of springs. Analysis of the response of the water table to climatic variations based on groundwater level records has proved to be limited by the spatial heterogeneity of its behaviour and the factors that influence it [Tirogo *et al.*, 2016]. Therefore, hydrogeological modelling was recommended to study the behaviour of the water table with respect to climate and to pumping. In this area, several hydrogeological

models run in steady state [Sogreah Ingénierie, 1994, Dakouré, 2003, Derouane, 2008, Sauret, 2008] and aimed at understanding the hydrodynamic functioning of the aquifer system have not provided answers to the question.

Here, we propose to study the behaviour of the aquifer system using a hydrogeological model in both steady and transient states. Indeed, the response of aquifers to various demands is a dynamic and spatially variable process that can be effectively analysed by modelling [e.g., Bredehoeft, 2002, Dong *et al.*, 2012, Condon *et al.*, 2021]. Transient modelling enable the simulation of temporal and spatial variations in water table [Howard and Griffith, 2009] and hence the study of its behaviour in response to past or future climate change and anthropogenic activities [e.g., Scibek and Allen, 2006, Knowling *et al.*, 2015, de Graaf *et al.*, 2019]. The analysis of the combined impact of pumping and climate provides a better appraisal of groundwater dynamics subject to temporal trends (rainfall and pumping variability) in a context of spatial heterogeneity (drainage, aquifer geometry, hydrodynamic parameter distribution, etc.). In this article, we first present the hydrological, geological and hydrogeological framework of the study area, with particular emphasis on the discharges of the hydrosystem, at springs and as a result of withdrawals (Section 2). The results of the hydrogeological modelling, the development of which is described in Section 3, are presented in three main parts, namely, the calculation of the input data of the model, the calibration of the model in steady and transient states, and the evaluation of the impact of climate change and pumping (Section 4), before they are discussed (Section 5).

## 2. Study area

### 2.1. General information

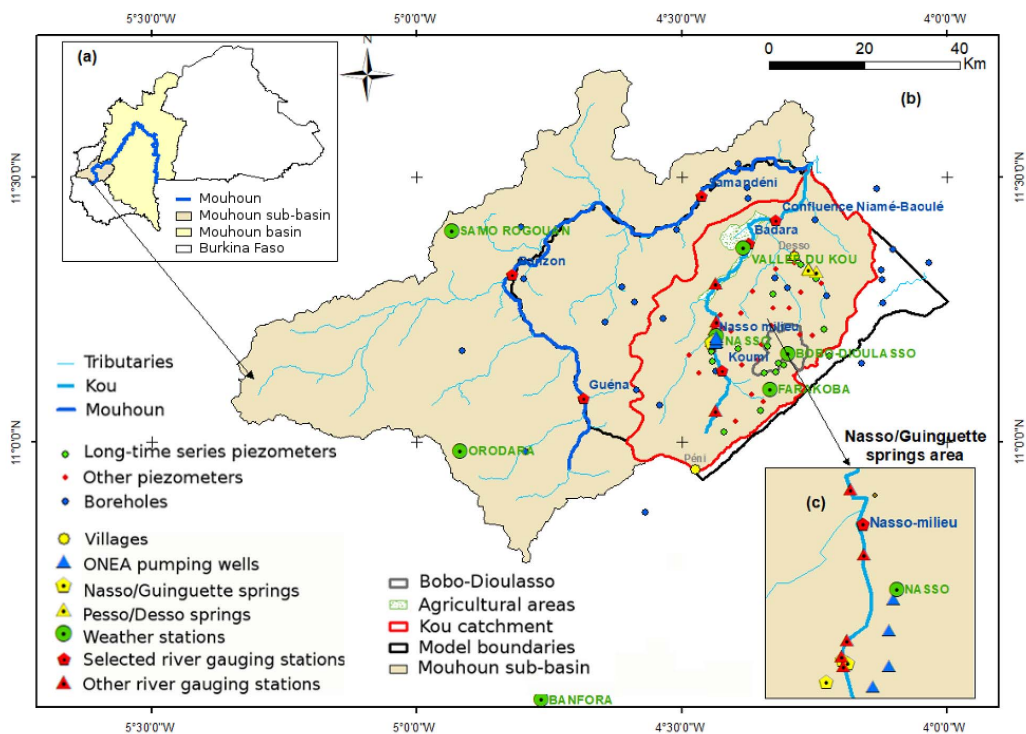
The study area covers approximately 3600 km<sup>2</sup> in southwest Burkina Faso between longitudes 4°00' W and 4°49' W and latitudes 10°56' N and 11°31' N (Figure 1). The Kou hydrological basin (1860 km<sup>2</sup>) is located in the administrative area of the Hauts-Bassins region and is home to the country's second largest city, Bobo-Dioulasso [904,920 inhabitants according to the results of the Fifth General Census of Population and Housing in 2019, i.e. approximately 92% of the population is urban in the Kou

basin, INSD, 2022]. It is in a Sudanian zone (800–1200 mm annual rainfall) characterized by two alternating seasons (dry season from November to April and rainy season from May to October). The average annual rainfall recorded between 1961 and 2014 at the Bobo-Dioulasso rainfall station is approximately 1025 mm and the monthly average temperatures range between 25 °C and 31 °C during the same period. Monthly potential evapotranspiration is highest in March (202 mm) and lowest in August (116 mm) [Tirogo *et al.*, 2016].

The Kou basin is home to an important groundwater resource that is the source of exceptional springs at the scale of the West African region. This resource is of primary socio-economic importance for Burkina Faso, in particular for the city of Bobo-Dioulasso and its surroundings. It is used for irrigation, industrial activities, and for drinking water supply. Water needs at the basin level are growing at the same rate as the population (growth rate of 3.1% per year) and as the development of irrigated agriculture (total irrigated area nearly 2000 ha in 2014). From about 2300 m<sup>3</sup>/day in 1960, the total amount of water withdrawn at the scale of the basin has increased to more than 75,000 m<sup>3</sup>/day in 2014 [Tirogo, 2016].

### 2.2. Hydrology

The Kou basin contains the perennial Kou River, its tributaries, and the Nasso/Guinguette springs. The Kou River (ca. 70 km long) is the first major right-bank tributary of the Mouhoun River, formerly the Black Volta [Moniod *et al.*, 1977]. The source of the river is in the locality of Péni at an altitude of approximately 500 masl. The Kou River is fed by several springs, the most important of which are the Nasso/Guinguette and Pesso/Desso springs (Figure 1). The springs constitute most of the base flow of the Kou River (about 2.1 m<sup>3</sup>/s of 2.5 m<sup>3</sup>/s between 1992 and 1997) and contribute to the downstream supply of important hydro-agricultural developments, notably the irrigated areas of the Kou Valley. The main tributaries of the Kou River are mostly temporary streams. At the confluence with the Kou River (286 masl altitude), the Mouhoun is approximately 147 km long and drains a watershed of 6647 km<sup>2</sup>. The average flow of the Kou (at the confluence of Niamé-Baoulé) and of the Mouhoun (at Samendéni) over a



**Figure 1.** (a) Study area location in Burkina Faso; (b) study area description and measurement stations (river, meteorology, and piezometers); (c) Nasso/Guinguette springs area.

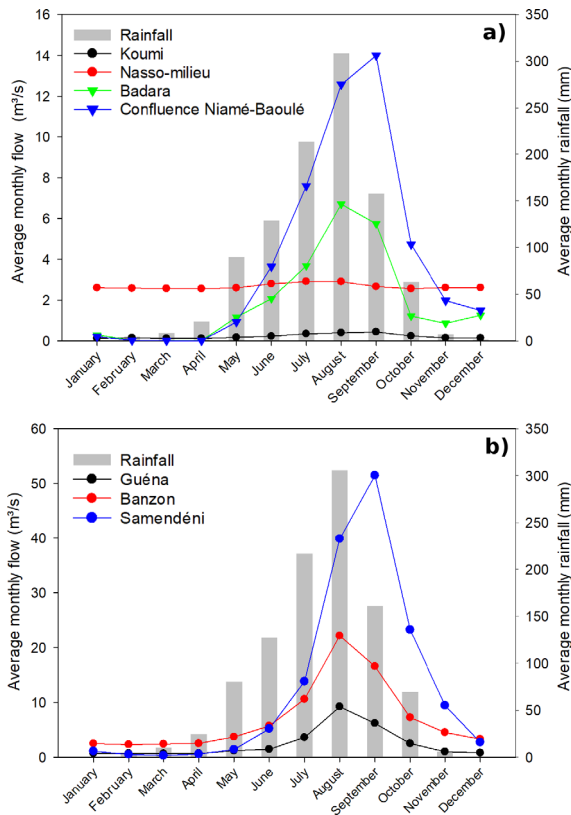
common period from 1985 to 1994 was  $3.9 \text{ m}^3/\text{s}$  and  $14.9 \text{ m}^3/\text{s}$ , respectively (Figure 2).

### 2.3. Geology and hydrogeology

The study area is part of the southeastern edge of the Taoudeni sedimentary basin (Figure 3a) [Trompette, 1973, Bronner *et al.*, 1980]. This Burkinabe part of the Taoudeni basin ( $45,000 \text{ km}^2$ ) consists of nine sandstone-dominated formations [Ouédraogo, 1983, 2006] with a monoclinical structure and a slight dip of  $2^\circ$  to the NW (Figure 4). It is limited to the east by the contact with the basement rocks (Figure 3b). From a hydrogeological point of view, the limits of the Kou watershed do not correspond to those of its hydrological basin. Indeed, based on the potentiometric surface map of the whole area, lateral groundwater flows would come from upstream of the basin and continue downstream to the Gondo Depression [Dakouré, 2003, Derouane, 2008, ANTEA, 2012].

The Kou basin has five outcrop formations, predominantly sandstone of terminal Precambrian (Infracambrian), with a total thickness that could reach

2000 m deep in its western part. These formations constitute its main aquifers separated locally by discontinuous layers of mudstone that make them locally confined (Figure 3c). From the oldest to the most recent level, it is made up of: Kawara-Sindou sandstone aquifer (GKS), glauconitic fine sandstone aquifer (Gfg), quartz granular sandstone aquifer (Ggq), Guéna-Sourkoudinga silty clay and carbonate aquifer (SAC1), and pink fine sandstone aquifer (Gfr). The GKS multi-layered aquifer is made of sandstone. The upper part of the Gfg aquifer is very productive while the lower clayey part is not very permeable and more or less sandstone-like. The Ggq aquifer is made up of two confined aquifers (lower Ggq and upper Ggq) separated by a layer of fine, highly consolidated sandstone. The base of the SAC1 aquifer consists of sandstone and the top consists of dolomitic limestone and clay. The Gfr aquifer, less known than the first four aquifers, contains a thin and shallow water table [Sogreah Ingénierie, 1994, Gombert, 1998]. All along the Kou River, the aforementioned formations are covered by alluvial deposits whose lateral extension is estimated at 350–



**Figure 2.** Average monthly flow of the Kou (a) and Mouhoun (b) rivers from 1985 to 1994 and average monthly rainfall of Kou and Mouhoun catchments during the same period.

750 m on each side of the river. Consisting of clay, sandy-clay, and altered sandstone deposits, from a few metres to nearly 40 m thick, the alluvial plain functions as a capacitive reservoir [Sauret, 2013].

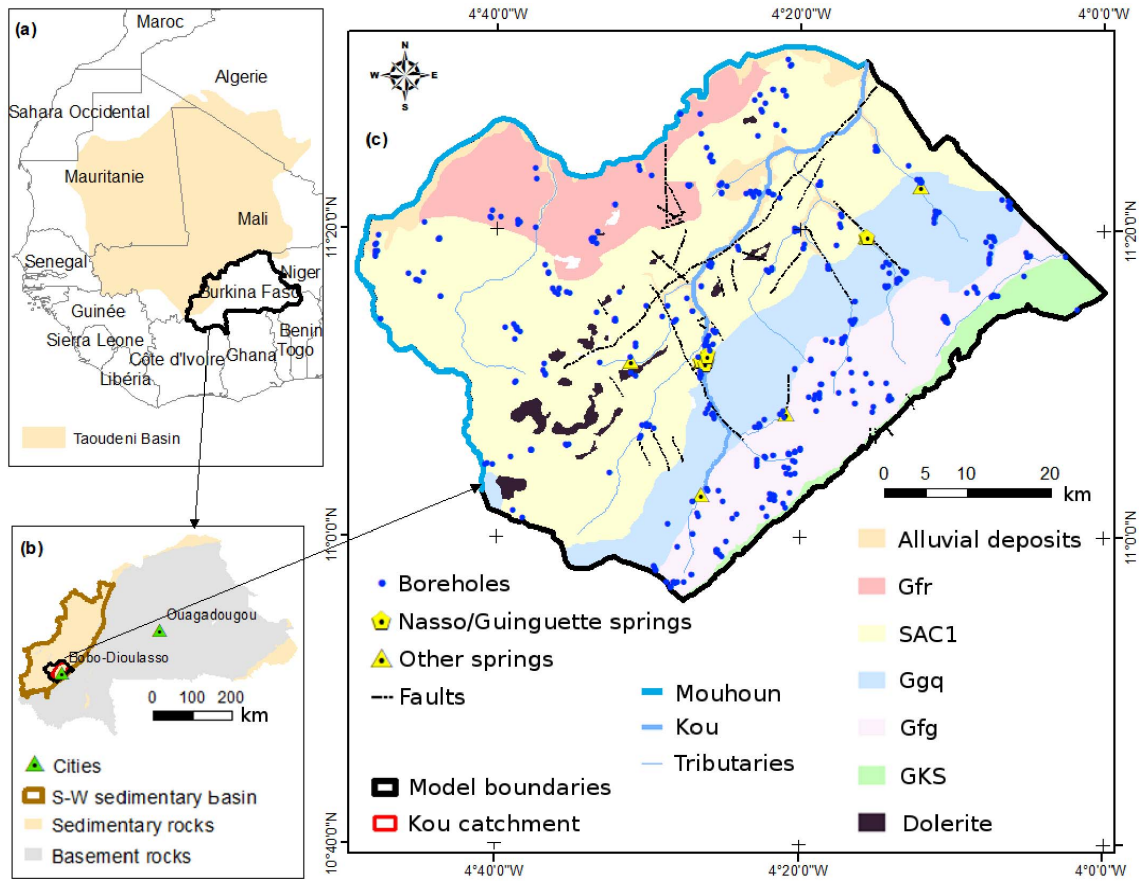
Thus, from the basement to the base of the SAC1, sandstone layers overlap over approximately 1000 m [Gombert, 1998], only separated by discontinuous layers of mudstone at the base of the Gfg in the eastern edge of the basin, which makes the GKS locally confined. Within the SAC1, the presence of thick layers of locally impermeable claystone suggests a compartmentalization of this aquifer. Also, dolerites outcrop abundantly at the western and northern ends of the Kou basin and occur as roughly concordant sills, dykes (veins), and necks [Sogreah Ingénierie, 1994, Gombert, 1998]. Their presence in the study area could also constitute real screens impermeable

to groundwater flow. However, their vertical and even horizontal extension remains poorly known. At the scale of the study area, according to the current state of knowledge, the existing data are insufficient to define the extension of these heterogeneities as well as their role in the functioning of the aquifer system.

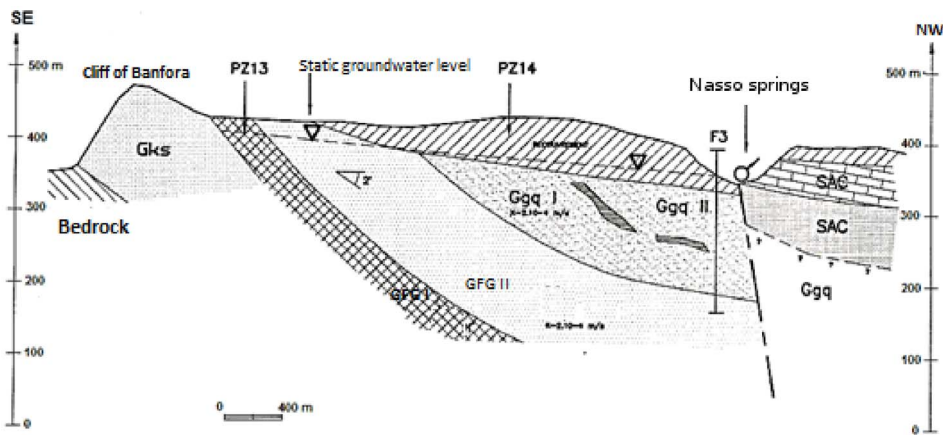
In general, the outcrop formations are fractured. There are several major faults, generally oriented SSE–NNW, which promote the circulation of water at depth and create important hydraulic connections between the different aquifers [Sogreah Ingénierie, 1994]. These are also areas favourable to the recharge of the water table and to the ascent of deep and older waters [Huneau *et al.*, 2011]. They are considered to be the main causes of resurgence of water from the Nasso/Guinguette springs [e.g., Talbaoui, 2009, Figure 4].

#### 2.4. Discharge of the aquifer system by springs

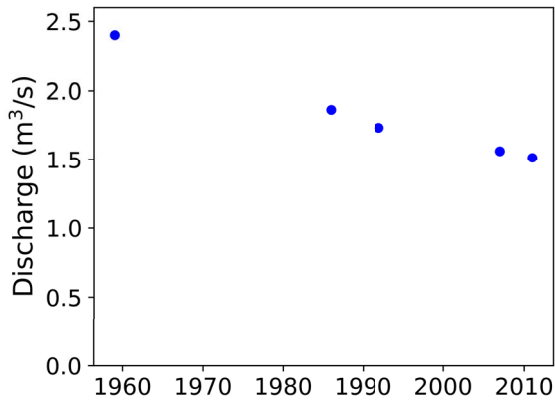
The aquifer system is drained by numerous springs. Apart from the Nasso/Guinguette springs already mentioned, there are about 30 other springs in the area with discharge varying between 1 m³/h and 500 m³/h according to a survey conducted in 1996. From this inventory to today, some of these springs have dried up and others have experienced a significant decrease in their discharge rates. Unfortunately, there is no recent inventory of all the springs with measurements of discharge. Only the Guinguette spring allows us to appreciate the evolution of the discharge over the long term thanks to sporadic measurements of the discharge rate realized between 1959 and 2011 [Sogreah Ingénierie, 1994, Sauret, 2013]. The discharge rate has decreased by approximately 1 m³/s in 50 years, from ca. 2.4 m³/s in 1959 to ca. 1.6 m³/s in 2011 (Figure 5), with low seasonal variation (less than 15% of the average annual flow in 1992). The two springs of Nasso, operated by the National Office of Water and Sanitation (ONEA), whose discharge was estimated at approximately 0.5 m³/s in 1993 [Sogreah Ingénierie, 1994], could not supply more than 0.33 m³/s in 2012. Apart from the two Nasso springs, which are used to supply drinking water to the city of Bobo-Dioulasso, most of the other springs, including the Guinguette, discharge into the river system where the water is used for irrigation in the downstream part of the basin.



**Figure 3.** Study area description: (a) location of the Taoudeni basin, (b) boundaries of the Western Burkinabe sedimentary basin, (c) geology of the study area described in Section 2.3 (in particular with regard to the short names of the outcropping geological formations).



**Figure 4.** Synthetic geological section of the Bobo-Dioulasso area [from Sogreah Ingénierie, 1994]. Hydrogeological units are referred to by their abbreviated names (see Section 2.3).

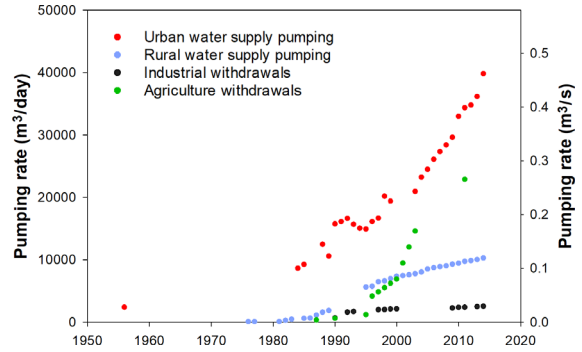


**Figure 5.** Evolution of the low-flow discharge rate of the Guinguette spring between 1959 and 2011.

### 2.5. Discharge of the aquifer system by withdrawals

The groundwater resources of the Kou watershed are exploited by many users (water supply, industry, agriculture). As stated in Section 2.1, water withdrawals have increased since the first drilling in the 1950s. Indeed, at that time, water abstractions were made only from springs and from a few traditional wells [Palausi, 1957].

ONEA's pumping system for supplying drinking water to urban areas including the city of Bobo-Dioulasso is the most important and best monitored. The quantities of water withdrawn by ONEA have evolved from ca. 2300 m<sup>3</sup>/day in 1956 [Palausi, 1957] to ca. 40,000 m<sup>3</sup>/day in 2014 (Figure 6) [Tirogo, 2016]. Between 2000 and 2014, ONEA's set-up evolved with the construction of two boreholes of approximately 250 m<sup>3</sup>/h each in 2000, followed by two others of 250 m<sup>3</sup>/h in 2014, all located within a radius of less than 2 km. Their depths vary between 200 and 300 m. In rural areas, drinking water is drawn from shallower boreholes (generally between 40 and 80 m) with flow rates varying between 0.5 m<sup>3</sup>/h and 20 m<sup>3</sup>/h. Most of them are equipped with manual pumps. The number of boreholes has grown since the first one in 1976 to a total of 346 in 2014. The withdrawals in 2014 were evaluated at approximately 10,000 m<sup>3</sup>/day to the nearest 20%, based on well flow rates and estimated pumping hours per day [Tirogo, 2016]. In addition, there are agricultural withdrawals, mainly located in the alluvial plain downstream of the Kou



**Figure 6.** Evolution of total drinking water production for the city of Bobo-Dioulasso (from springs and boreholes), rural drinking water supply withdrawals, as well as industrial and agricultural withdrawals.

basin, estimated at approximately 22,900 m<sup>3</sup>/day in 2011 during the irrigation period (November–April) [Sauret, 2013]. Groundwater is used by many industries located in the Bobo-Dioulasso area, whose withdrawals are ca. 2500 m<sup>3</sup>/day according to an estimate made in 2014 [Tirogo, 2016].

In 2014, total pumping by all users was approximately 75,000 m<sup>3</sup>/day on average. Taking into account seasonal variations, the total volume withdrawn for the year 2014 was estimated to be approximately 84,000 m<sup>3</sup>/day in the dry season (low-flow periods) and approximately 44,000 m<sup>3</sup>/day in the rainy season (high flow periods).

## 3. Materials and methods

### 3.1. Numerical code

The code chosen for the modelling is Modflow-2005 [Harbaugh, 2005], with its Visual Modflow© interface. Modflow-2005 solves the groundwater flow equation using the finite difference method for both confined and unconfined aquifers.

### 3.2. Geological model and discretization

The geological model is built in three dimensions considering the five aquifers. The basement on which the GKS aquifer rests is considered as the lower impermeable boundary of the domain. The top of the aquifer system is the topographic surface, based on

30 m SRTM data. The geometry of the multi-layered aquifer system was developed using data from six geological sections [Ouédraogo, 2006, Derouane, 2008] and about 60 borehole logs, with depths not exceeding 300 m. Beyond this depth, the vertical position of contacts between deep horizons was extrapolated from the observed average dip. The total thickness of the modelled domain is 1050 m in its deepest part to the NW of the basin where it is distributed as follows: 350 m of GKS, 150 m of Gfg, 200 m of Ggq, 250 m of SAC1, and 100 m of Gfr (Figure 7b).

The study area was discretised into regular 500 m  $\times$  500 m grid cells according to the six layers generated from the 3D geological model, for a total of 78,732 active grid cells (Figure 7). The GKS formation was divided into two layers to represent the semi-permeable layer that separates the GKS unit from the Gfg unit, as its impact on groundwater flow was established in the southeast of the basin. Other vertical heterogeneities were not represented in the model because they are unknown at the scale of the area. They are indirectly accounted for in the five aquifer horizons of the model by an equivalent hydraulic conductivity value generally lower than the one estimated from pumping tests.

### 3.3. Hydrodynamic parameters

The modelled domain is represented as a continuous porous medium in which the heterogeneities are taken into account through the variations of the values of the hydrodynamic parameters. The a priori values of horizontal hydraulic conductivity are estimated from ca. 20 long-term pumping tests. We deduced an average value for each layer:  $2.4 \times 10^{-5}$  m/s for GKS,  $9 \times 10^{-6}$  m/s for Gfg,  $2.1 \times 10^{-4}$  m/s for Ggq,  $3.6 \times 10^{-5}$  m/s for SAC1, and for Gfr,  $1.3 \times 10^{-5}$  m/s. The hydraulic conductivity in the direction parallel to the layers ( $K_x = K_y$ ) is assumed to be 10 times higher than the conductivity in the perpendicular direction ( $K_z$ ) for all layers. The major faults are implicitly considered in the model using a horizontal anisotropy factor ranging from 1 to 2 orders of magnitude. The a priori values of the storage coefficient are based on five values, available for GKS ( $1.5 \times 10^{-4}$ ;  $3.8 \times 10^{-3}$ ), Gfg ( $3.8 \times 10^{-3}$ ), and Ggq ( $5.5 \times 10^{-4}$ ;  $2.3 \times 10^{-3}$ ). The specific yield of the aquifers (outcrop parts) was initially considered to be between 2% and 5% according to their lithology.

### 3.4. Boundary conditions

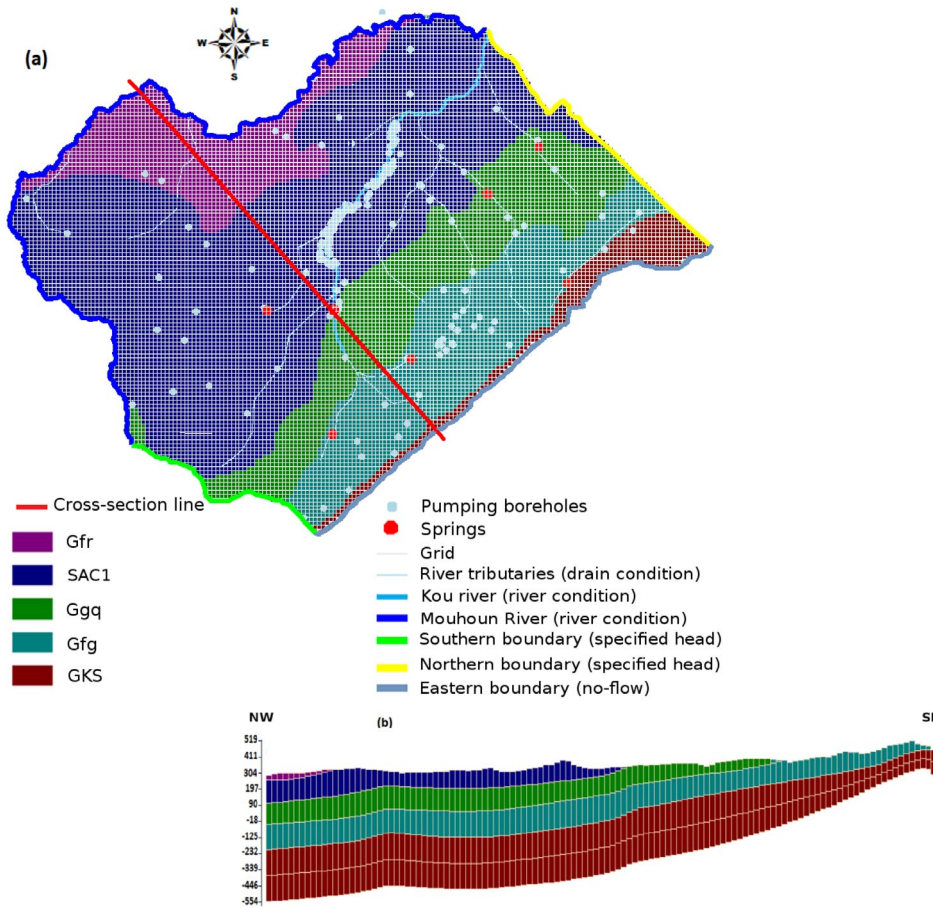
Since the hydrological boundaries of the Kou watershed do not match those of the hydrogeological basin (Section 2.3), the model domain was chosen in such a way that boundary conditions could be more easily defined (Figure 7). They were extended mainly to the west and east of the basin in order to find natural boundaries, which are, respectively, the Mouhoun River and the limit of the sedimentary basin (Bobo-Dioulasso cliff).

The eastern boundary of the sedimentary basin is prescribed as a no-flow boundary condition, assuming very limited exchanges between the basement and the sedimentary basin. On the Mouhoun River, which is a perennial river, a specified head boundary condition is assigned based on river heights. Daily water level and discharge data, provided by the General Directorate of Water Resources, are available for stations on the Kou and Mouhoun rivers for various periods between 1961 and 2014, ranging from 9 to 54 years. The data selected according to a data quality criterion from four stations on the Kou (Koumi, Nasso-milieu, Badara, Confluence Niamé-Baoulé) and three stations on the Mouhoun (Guéna, Banzon, Samendéni) were used to define these boundary conditions (water levels) (Figure 1). On the northern and southern limits of the area, specified steady state head boundaries are defined from the observed groundwater levels and the values are extracted from the reference potentiometric map. Groundwater level data from 100 observation points (87 piezometers and boreholes intersecting the phreatic aquifer, and 13 springs) for the period 1995–1999 were used to draw this reference potentiometric map for the low-flow period of this time interval.

Streams in the Kou basin are represented using a Robin boundary condition, i.e. a head-dependent flux condition. In the sections of the Kou River assumed to be in equilibrium with the water table, the river hydraulic head is derived from hydrometric data. Its tributaries, mostly temporary, are represented with a drain boundary condition and the drain elevation is equal to the topographic level. According to the classic approach proposed in Modflow, the amount of flux exchanged is regulated by the value of the conductance  $C$  ( $L^2 \cdot T^{-1}$ ) of the river bed, which is written as:

$$C = K_z \cdot l \cdot L/e, \quad (1)$$





**Figure 7.** Numerical model: (a) geological map with main springs, pumping boreholes, rivers, and drains, south, north, and east boundaries; (b) SE–NW cross-section with altitude in metres above mean sea level. Hydrogeological units are referred to by their abbreviated names (see Section 2.3).

with  $K_z$  the vertical hydraulic conductivity of the river bed materials ( $L \cdot T^{-1}$ ),  $l$  the width of the river within the mesh ( $L$ ),  $L$  the length of the river within the mesh ( $L$ ), and  $e$  the thickness of the river bed ( $L$ ). The main springs in the area are also represented with a drain boundary condition.

All the pumping boreholes listed in the area were taken into account in the model by considering their exact position, except for boreholes equipped with manual pumps. The latter have low pumping rates (ca.  $0.7 \text{ m}^3/\text{h}$ ), and were therefore grouped by village by summing the pumped rates, which reduced the number of wells from 346 to 91.

### 3.5. Natural groundwater recharge assessment

The recharge values introduced into the model are derived from the Thornthwaite balance [Thornthwaite, 1948]:

$$P = AET + R + I + \Delta S, \quad (2)$$

with  $P$  the total rainfall,  $AET$  the actual evapotranspiration,  $R$  the runoff,  $I$  the infiltration, and  $\Delta S$  the water stock variation in the available water content (AWC), all terms of dimension ( $L \cdot T^{-1}$ ). Infiltration is considered as diffuse groundwater recharge. The water balance is calculated on a monthly basis. For a given month  $j$ , the infiltration is calculated as follows:

$$AWC_j = \min(\max(AWC_{j-1} + P_j - ETP_j, 0), AWC_{\max}) \quad (3)$$

and

$$I_j = \max(AWC_{j-1} + P_j - ETP_j - AWC_{\max}, 0) - R_j, \quad (4)$$

where  $AWC_{\max}$  is the soil maximum available water capacity. Climate data are made available through permanent monitoring carried out by the technical services of the National Agency of Meteorology. Daily climate data from the Bobo-Dioulasso meteorological station (rainfall, temperature, Penman-Monteith potential evapotranspiration; 1961–2014) and monthly rainfall data from five other rainfall stations (Nasso, Farakoba, Vallée du Kou, Orodara, and Samoroguan; variable periods between 1961 and 2014) were used. The latter were completed using Bobo-Dioulasso time series with which they are correlated at over 0.8 ( $R^2$ ) [Tirogo *et al.*, 2016]. These data were used to assess groundwater recharge during the period 1961–2014 for both steady state and transient modelling. The infiltration was evaluated from the average basin rainfall of the whole modelled area calculated from the Thiessen polygon method and the data from the above-mentioned stations. A time-varying runoff coefficient in the range of 3%–9% is derived from the hydrometric data of the Kou River for the period 1961–2014. The maximum available water content of the soil ranges from 10 to 240 mm, with more than 80% of the basin in the 100–200 mm range [Wellens *et al.*, 2007]. In this work, we considered values between 100 and 200 mm with an average of 130 mm for the whole area.

### 3.6. Steady state and transient simulation configurations

A two-stage trial-and-error calibration procedure was used. The model was first calibrated in steady state using data from 1995 to 1999. It was then calibrated in transient state at a monthly time step during the period 1995–2014 and applied to estimate the different flows that transit in the domain. The choice of these periods for the history matching was driven by data availability [Tirogo *et al.*, 2016]. Finally, the model was used over the period 1961–2014 in order to analyse the effect of climate variability on the water table and in particular the impact of the great drought of the 1970s and 1980s.

For the transient simulation over the period 1995–2014, the boundary conditions are of the same type

as those defined in steady state (Section 3.4) but evolving with time, based on the observed water level in the rivers and on water table fluctuations in the piezometers near the boundaries. All other time-varying data (recharge, pumping) were also introduced into the model at a monthly time step. Initial conditions were obtained from the steady state solution for the period 1995–1999.

The 1961–2014 transient simulation was run at an annual time step. No groundwater-level time series was available for this period, preventing the use of specified head boundary conditions along the northern and southern limits. The evolution of the simulated inflow and outflow through the specified head boundaries in response to changes in recharge rate over the period 1995–2014 (decrease of ca. 20 mm between 1995–1999 and 2014) was analysed. It served as a basis for prescribing their temporal evolution as a Neumann specified-flux boundary condition, based on the estimated change in recharge from 1961 to 2014. With these boundary conditions, a steady state simulation was performed for the year 1961 and used as initial conditions for the 1961–2014 transient run.

### 3.7. Model calibration

Considering the availability of data, the steady state model is calibrated under low-flow conditions, using data from 68 observation points between 1995 and 1999. The discharge data of the main springs of the area and river flow were also used to calibrate the model. The input parameters for the steady state calibration were recharge, hydraulic conductivity, and river and drain conductances. The steady state calibration consisted in adjusting these parameters to reproduce the observed groundwater level and the flows of the Kou River, the drains (temporary tributaries), and the springs.

The a priori and uniform value of the recharge was adjusted following a sensitivity analysis, performed with a range of recharge values calculated using maximum available water contents between 100 and 200 mm and runoff coefficient values from 3% to 9%. These tests led to the definition of higher recharge values in the high-altitude areas of the domain than in the rest of the modelled area.

After setting the values of the recharge, the hydrodynamic parameters were calibrated, starting from a priori average hydraulic conductivity values from

the pumping tests (Table 1). The single hydraulic conductivity value applied to each layer was subsequently modified, especially in the outcrop areas, to reproduce the observed groundwater levels. In the absence of measured hydraulic conductivity values and groundwater level data in the deep levels (more than 300 m), the initial  $K$  values were not modified outside the outcrop areas, except in the upper GKS layer in the eastern part of the domain. In this area, the reproduction of the observed groundwater levels within the GKS and Gfg layers required the inclusion of a semi-permeable level that plays an important role in the hydrodynamic functioning of the aquifer system. The faults also play an important role in simulating the observed water table and flow rates and were progressively introduced in the calibration procedure. Finally, the conductance values were adjusted to reproduce the flows of 10 springs and those of the Kou and Mouhoun rivers.

Under transient conditions, the model was calibrated using the 20-year groundwater level time series of 30 observation points from April 1995 to December 2014 at a monthly time step, based on their detailed analysis by Tirogo *et al.* [2016]. Only the storage coefficient and the specific yield values were adjusted in this second calibration step, starting from the *a priori* values presented in Section 3.3. Particular attention was paid to reproducing the heterogeneous behaviour of the aquifer system after 2005 between the upstream and downstream parts of the basin, as described by Tirogo *et al.* [2016]. Indeed, it has been shown that, despite a slight increase in recharge after 2005, the water table still decreases upstream of the Nasso/Guinguette springs while it increases downstream of the basin. This led to the delineation of new recharge zones upstream and downstream of the Nasso/Guinguette springs. This revised spatial distribution of recharge was then taken into account in the entire steady state calibration procedure and in the design of the resulting transient simulations.

## 4. Results

### 4.1. Observed groundwater level map (1995–1999)

From groundwater level data, a reference average potentiometric map was drawn for the 1995–1999 low-flow period (Figure 8). A single map can be drawn

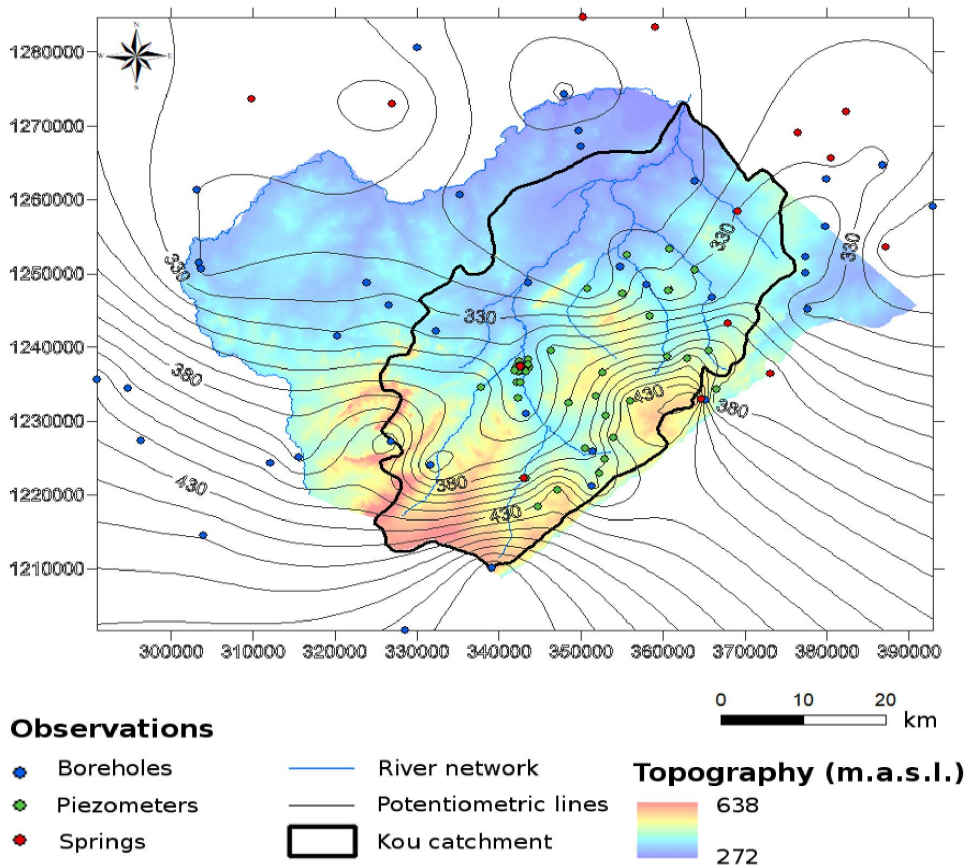
consistently over the entire study area, regardless of the outcrop aquifer. The assumption of a single water table was the basis for previous modelling works [Dakouré, 2003, Derouane, 2008, Sauret, 2008], given the absence of data on deep levels (beyond 300 m) or data that would make it possible to distinguish differences in heads between aquifers. In general, the observation points capture the first aquifer encountered and the static levels are shallow. The depth of the measured static levels is generally less than 10 m, but in a few rare places, it can reach 40–60 m. The direction of groundwater flow generally follows the topography. The highest groundwater levels are found in the south and east of the basin, in the topographic heights. On the potentiometric map for the period 1995–1999, groundwater levels range from 499 masl in the south to 286 masl in the north with a gradient of 3‰.

### 4.2. Recharge of the aquifer system

The monthly recharge was calculated during the period 1961–2014 using the Thornthwaite balance. The results show that recharge is highly variable from year to year and represents on average 12% of the rainfall (Figure 9). Over this period, on an annual scale, the balance terms are distributed as follows:

- In mm,  $P (1010) = AET (812) + R (75) + I (123)$ ,
- In percent,  $P (100\%) = AET (81\%) + R (7\%) + I (12\%)$ .

The proportion of rain that infiltrates, estimated by several authors [e.g., Gombert, 1998, Dakouré, 2003] at different periods between 1950 and 2000, is in the same order of magnitude (10–15%). The value of the annual recharge could thus reach 240 mm in the 1950s when rainfall was abundant [Sogreah Ingénierie, 1994] whereas in the recent period, it is less than 120 mm/year [Dakouré, 2003, Derouane, 2008]. There is a decreasing trend in recharge from 1961, which was attributed not only to a decrease in rainfall but also to a slight increase in evapotranspiration [Tirogo *et al.*, 2016]. This decline is also reflected in the presence of years without recharge after 1980. For the 1995–1999 steady state regime, the mean recharge value determined from these results is 95 mm/year but can vary between 60 mm/year and 150 mm/year. It is in this range of values that the sensitivity analysis of the model to recharge was performed.



**Figure 8.** Topography level map (colour) and average groundwater level map (lines) (1995–1999) in metres above mean sea level.

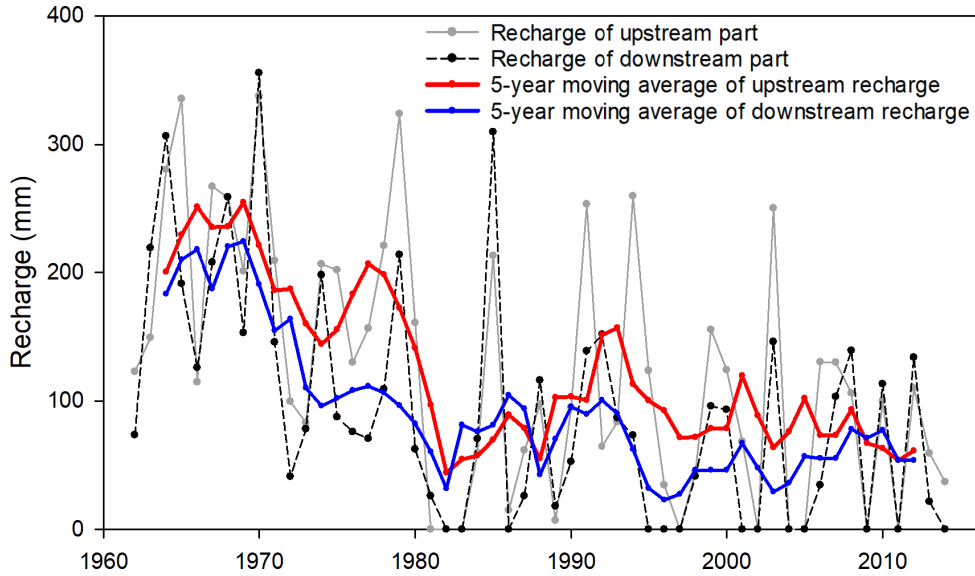
### 4.3. Model results

#### 4.3.1. Steady state calibration results

Following the sensitivity analysis, two preferential recharge zones were delineated around the high-altitude zones, to the south (around Péni) and to the east (Bobo-Dioulasso area), with recharge values of 115 mm/year compared to 95 mm/year in the other areas. This result is in agreement with the hypothesis put forward by Dakouré [2003] from isotopic analysis results, according to which the high-altitude zones are preferential recharge zones. The calibrated hydraulic conductivity values are summarized in Table 1. They are generally within the range of values obtained from the long-term pumping tests. In the outcrop areas, the calculated values generally vary within the same layer by 1–2 orders of magnitude. Lower values of hydraulic conductivity

( $<10^{-6}$  m/s) are predicted locally, reflecting the heterogeneous nature of the aquifer layers, which is also supported by observations from short-term pumping tests [Dakouré, 2003, Tirogo, 2016].

A total of 18 faults were introduced into the model. Their role in the reproduction of the observed groundwater levels proved important. As a result of the calibration, most of the faults were found to be conductive, especially those connected with the Nasso/Guinguette springs ( $K$  is on the order of  $10^{-2}$  m/s). Overall, their calculated hydraulic conductivity values range from  $10^{-4}$  to  $10^{-2}$  m/s, generally 1–2 orders of magnitude higher than those of the aquifer. Only a low-conductivity fault, intersecting Gfg and GKS, was introduced (Figure 10). In these ancient sandstones separated in places by mudstones, some faults could indeed be clogged and then considered of low permeability.



**Figure 9.** Variation in annual recharge between 1961 and 2014 for the upstream and downstream parts of the study area (with  $AWC_{max} = 130$  mm, see Section 3.5).

**Table 1.** Hydraulic conductivity values from measurements and from steady state calibration

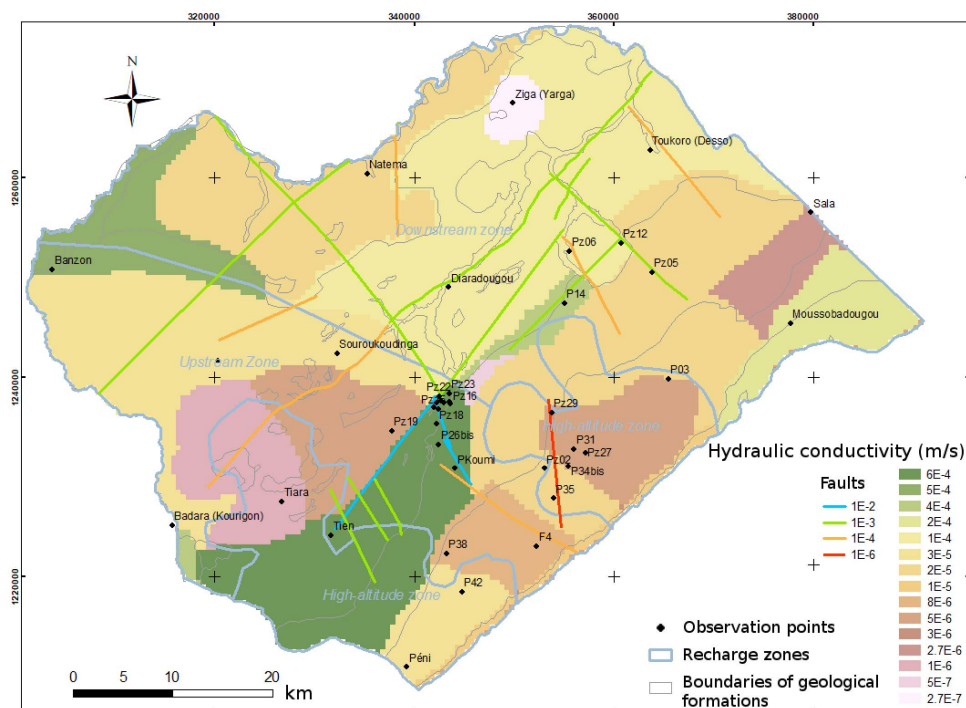
Aquifer layer	Measured values (m/s)		Calibrated values (m/s)	
	Minimum	Maximum	Minimum	Maximum
Gfr	$1.3 \times 10^{-5}$		$2.0 \times 10^{-5}$	
SAC1	$3.0 \times 10^{-5}$	$5.6 \times 10^{-3}$	$2.7 \times 10^{-7}$	$5.0 \times 10^{-4}$
Ggq	$9 \times 10^{-5}$	$5.2 \times 10^{-4}$	$2.0 \times 10^{-5}$	$4.0 \times 10^{-4}$
Gfg	$2.9 \times 10^{-8}$	$3.3 \times 10^{-5}$	$3.0 \times 10^{-6}$	$2.0 \times 10^{-5}$
Impermeable layer			$1.0 \times 10^{-8}$	$5.0 \times 10^{-8}$
GKS	$1.1 \times 10^{-5}$	$3.6 \times 10^{-5}$	$5.0 \times 10^{-6}$	$2.0 \times 10^{-4}$

The flow rates of the springs were adjusted using the drain conductance at the spring and the calculated values are of the order of  $10^{-1} \text{ m}^2/\text{s}$ . For the Kou River and the drains (river tributaries), these values are between  $4 \times 10^{-4}$  and  $2 \times 10^{-3} \text{ m}^2/\text{s}$ .

The hydrodynamic parameters presented above allowed us to reproduce the observed groundwater levels in a satisfactory way. The mean absolute error is 1.91 m for groundwater levels ranging between 286 m and 499 m. In addition, the root mean square error (RMSE) of 2.35 m is lower than the uncertainty in the topographic data, which could be as high as 4 m in some locations [Tirogo, 2016]. The general groundwater flow is reproduced for the period 1995–1999, with an overall direction of flow from south to

north. The horizontal hydraulic gradient is higher upstream than downstream of the domain.

The global water balance of the model is presented in Table 2. The total flow in the domain is estimated at  $19.65 \text{ m}^3/\text{s}$ . Recharge represents almost half of the water inflow to the aquifer ( $9.55 \text{ m}^3/\text{s}$  or 49%). The rest of the inflow comes from lateral flows and is estimated at  $3.95 \text{ m}^3/\text{s}$  from the south (20%),  $4.43 \text{ m}^3/\text{s}$  from the southwest via the Mouhoun (22%), and  $1.77 \text{ m}^3/\text{s}$  from the north (9%). The inflow from the Kou River is low ( $0.05 \text{ m}^3/\text{s}$  or 0.4%). As mentioned earlier, the study area is not a hydrogeological basin and the Kou basin is fed upstream by groundwater flow ( $\approx 42\%$  of total inflow). The main discharge point for



**Figure 10.** Hydraulic conductivity at outcrop from steady state calibration during 1995–1999.

groundwater is the Mouhoun (52%). Discharge to the northern boundary (downstream part of the domain) is only half of that into the Mouhoun (23%). Springs account for 13% of outflows, and the Kou River and its main tributaries account for 11%. Pumping accounts for only 2% of total outflows.

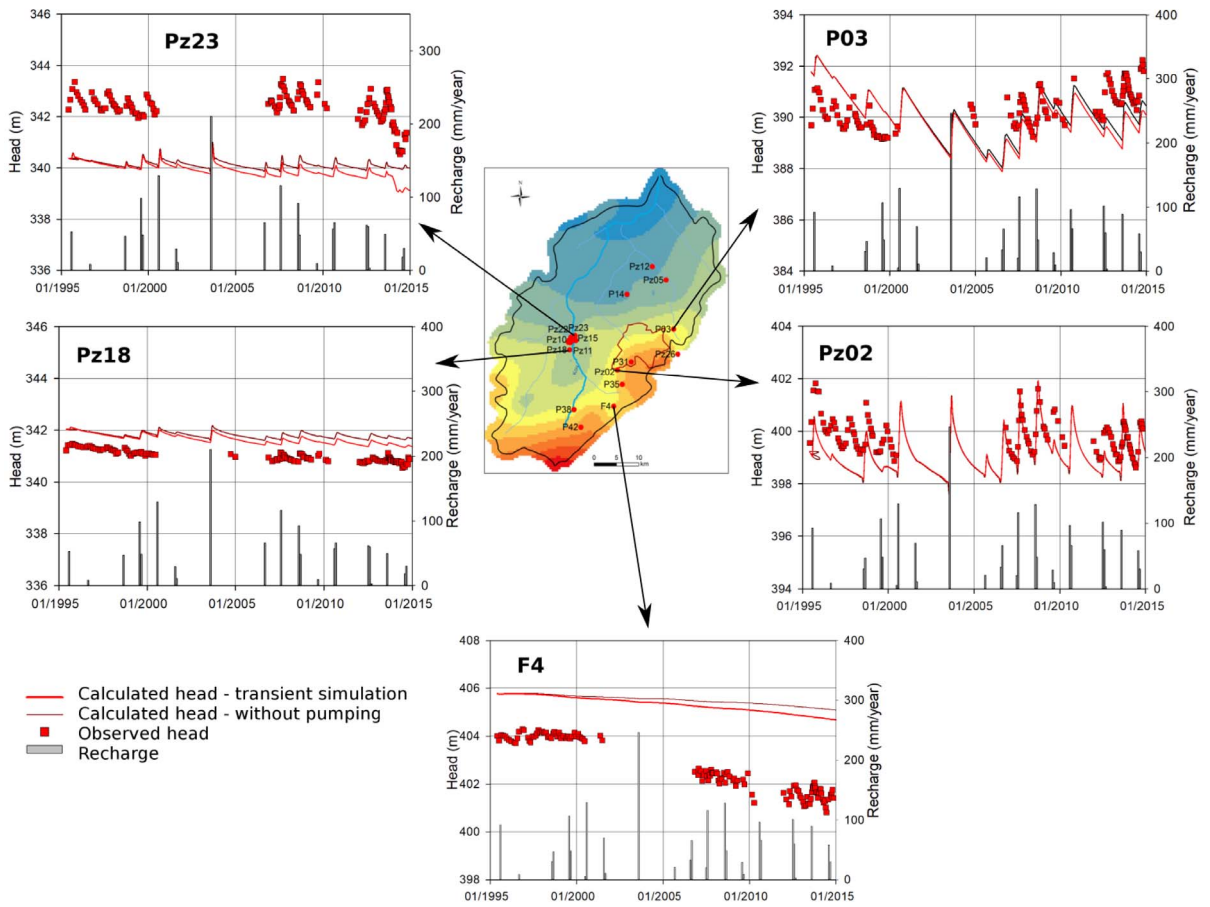
The inflows from the southern and western lateral boundaries would correspond to a recharge area of nearly 1100 km<sup>2</sup>, which extends therefore beyond the limits of the study area, in the Orodara area south of the study area, with a recharge assumed to be higher than in the Kou catchment (ca. 170 mm/year).

Regarding aquifer–river exchanges in the Kou basin, the model results indicate river losses upstream of Nasso, then mainly gaining from Nasso to Diaradougou (0.06 m<sup>3</sup>/s), in the order of magnitude of the flows calculated by Sauret [2013] on this section (0.09 m<sup>3</sup>/s). From Diaradougou to the confluence with the Mouhoun, exchanges are in one direction only, from the aquifer to the river. In the alluvial plain, downstream of the Kou basin, the water

table is very close to the surface (0–5.3 m). Overall, the Kou River drains the aquifer. The net flow from the aquifer to the river is estimated at 0.67 m<sup>3</sup>/s, including 0.36 m<sup>3</sup>/s upstream of the Niámé-Baoulé confluence.

#### 4.3.2. Transient calibration results

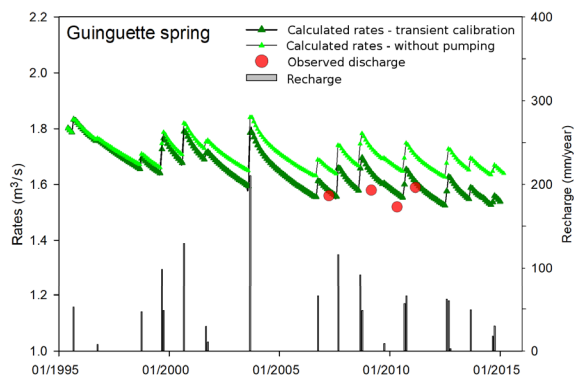
The calibrated distribution of specific storage ( $S_s$ ) shows a high spatial variability, with values ranging from  $5 \times 10^{-5}$  to  $5 \times 10^{-3}$  m<sup>-1</sup>. The values of  $S_s$  can be considered reasonable when compared with  $S_s$  deduced from available field values ( $8 \times 10^{-5}$  to  $1 \times 10^{-4}$  m<sup>-1</sup>). The specific yield ( $S_y$ ) varies between  $3 \times 10^{-2}$  and  $5 \times 10^{-2}$ . The differences between the aquifers (GKS, Gfg, Ggq, SAC1), however, are not very marked. Based on these values, the results of the comparison between observed and simulated groundwater level fluctuations are also considered satisfactory, although the dynamics of each piezometer could not always be reproduced (Figure 11), either because of the presence of several piezometers with different transient behaviour in the vicinity of each



**Figure 11.** Selection of simulated vs. observed hydraulic heads (simulation with pumping and simulation without pumping). In the background map, the steady state water table contour lines, 20 m apart, from 300 to 480 m.

**Table 2.** Groundwater balance at steady state during the period 1995–1999

		Inflow (m <sup>3</sup> /s)	Outflow (m <sup>3</sup> /s)
Specified hydraulic head	South boundary	3.95	0.14
	North boundary	1.77	4.43
	Mouhoun	4.33	10.16
Pumping		0	0.35
Drains (tributaries)		0	1.29
Springs			2.56
River (Kou)		0.05	0.72
Recharge		9.55	0
	<b>Total</b>	<b>19.65</b>	<b>19.65</b>



**Figure 12.** The Guinguette spring observed flow and simulated flow (simulation with pumping and simulation without pumping).

other (e.g., Pz23), or because of local specificities that could not be reproduced in the model (e.g., F4).

The components of the water balance are subject to seasonal and long-term variation related to variations in recharge. Inflows are dominated by those from the west with seasonal variations. In the downstream part of the study area, where the water table has been rising since 2006, the inflow from the Mouhoun River plays an important role in reproducing the evolution of the groundwater level. Therefore, there would be an additional filter between the rainfall signal and the water table response through the Mouhoun River, and subsequently an indirect recharge by this river.

The simulated discharge of the Guinguette spring decreases by approximately  $0.31 \text{ m}^3/\text{s}$ , from  $1.78 \text{ m}^3/\text{s}$  in 1995 to  $1.47 \text{ m}^3/\text{s}$  in 2014 in the low-flow period (Figure 12). For the Nasso springs, the simulated discharge decreases by approximately  $0.1 \text{ m}^3/\text{s}$ , from about  $0.5 \text{ m}^3/\text{s}$  to  $0.4 \text{ m}^3/\text{s}$  between 1995 and 2014. The simulated values are in the same order of magnitude as the observed data, both in terms of discharge decay (decrease by  $0.36 \text{ m}^3/\text{s}$  for the Guinguette spring and  $0.14 \text{ m}^3/\text{s}$  for the Nasso springs during the same period) and in absolute values where available.

#### 4.4. Response of the water table to pumping and recharge fluctuations

To test separately the influence of pumping and recharge variations on the water table, two simula-

tions were performed, each including only one of these forcings. Thus, the influence of recharge was evaluated in a transient simulation without pumping, and then the impact of pumping was tested by removing the fluctuations of recharge and keeping the value that determines the initial conditions in 1995.

##### 4.4.1. Response of the water table to recharge fluctuations

The response of the water table to the variation in recharge is spatially variable both in seasonal amplitude and in interannual variation (Figure 11). In general, the observed groundwater level decreases during the period 1995–1999 (period for which data are available). From the simulated groundwater level, we note that this decreasing trend continues until 2006. After the year 2006, the simulated groundwater level indicates a general trend toward stabilisation or a slight decrease in the upstream part of the basin (e.g., Pz02 and Pz18) while in the downstream part, there is an increase in the water table (e.g., P03). This evolution of the water table, both simulated and observed, follows the same trend as the recharge that drives it.

The simulated seasonal fluctuations vary between 0.2 and 2 m corresponding to the observed fluctuations. The highest seasonal amplitudes are recorded in the highest water level zone located in the Bobo-Dioulasso area (up to 2 m at Pz02 and P03). This reflects the reactivity of the water table in this zone, which is considered a preferential recharge zone. The groundwater response seems to depend on the depth to water table, which may reflect the variation of the transmissivity with the thickness of the saturated zone. Indeed, the seasonal fluctuations observed in the same zone (e.g., Pz02, P35, P31 with equal  $K$ ) are more important where the water table is a few meters deep (5 m in Pz02) compared to where the water table is more than 10 m deep (P35 at 12 m, P31 at 20 m and more).

The simulations also illustrate how the discharge of springs responds to variations in recharge. These variations over the past 20 years would have resulted in a decrease in the low-flow discharge of the Guinguette and Nasso springs of about  $0.16 \text{ m}^3/\text{s}$  and  $0.05 \text{ m}^3/\text{s}$ , respectively. The most significant decreases are linked to the successions of dry years over



at least 2 years (1995–1998, 2001–2002, 2004–2005) but it should also be noted that it only takes one wet year (e.g., 2003) to return to the discharge levels of 1995 (Figure 12). From 2006 onwards, the annual recharge, which has become more regular from year to year, would contribute to stabilising the discharge of the springs.

#### 4.4.2. *Response of the water table to pumping*

The results of the model simulation without variation in recharge show that withdrawals, although relatively low (ca. 2% of total flows), influence the water table (Figure 13). Over the period 1995–2014, this influence is reflected in a moderate drop in the water table ranging over the entire area from a few centimetres to approximately 1 m, depending on the distance to pumping (e.g., Pz23 within 200 m of ONEA wells) and on the hydrodynamic parameters. In general, the northern and eastern parts of the basin are the most sensitive to pumping, related to their relatively low hydraulic conductivity values compared to the southern and western areas. For the same period (1995–2014), in the Bobo-Dioulasso area where there is a lot of industrial pumping and in Darsalamy (F4, Figure 13), the simulated drop in the water table is of the same order of magnitude (ca. 0.5 m) as that of the ONEA pumping area. Yet pumping in these areas (ca. 2500 m<sup>3</sup>/day in 2014) is significantly lower than in the ONEA pumping area (ca. 40,000 m<sup>3</sup>/day in 2014). Indeed, ONEA boreholes tap into the Ggq, which are a very good aquifer with a hydraulic conductivity 10 times higher than that of the Gfg in which the pumping in the Bobo-Dioulasso area is done.

These simulations also allowed us to analyse the flow variations between 1995 and 2014 in order to evaluate the influence of the withdrawals on the springs' discharge. The same inflections observed on the water table are present on the simulated flow rate of the Nasso/Guinguette springs. The evolution of the springs' flow is marked by three phases (1995–2000, 2000–2014, from 2014), characterised by an accelerated decrease in the discharge rate (Figure 14, Guinguette spring). In this simulation without recharge variation, the discharge of the Guinguette decreases by about 0.10 m<sup>3</sup>/s and that of Nasso by about 0.04 m<sup>3</sup>/s under the effect of pumping, i.e., a contribution of pumping to the total decrease in the

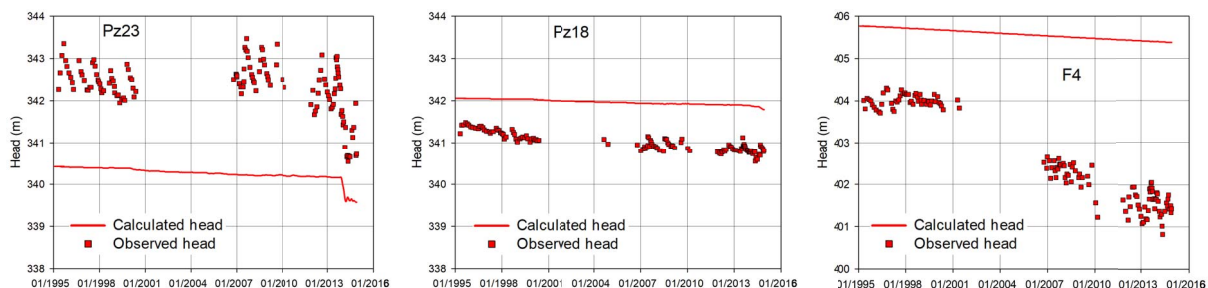
discharge of the Guinguette and Nasso springs between 1995 and 2014 (see Section 4.3.2) estimated at around 30–40%.

#### 4.4.3. *Impact of climate on the water table from 1961 to 2014*

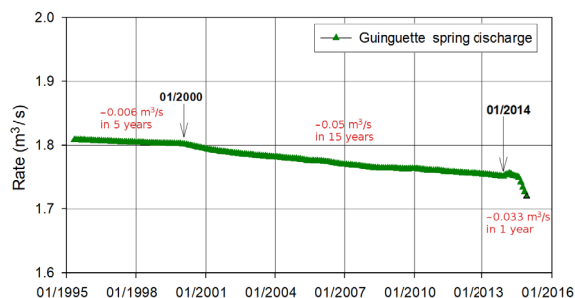
The simulation over the period 1961–2014 at annual time step allowed us to analyse the evolution of the water table and the discharge of springs related to the climate variability since the 1960s. The simulated groundwater levels are fairly well adjusted to the observed levels during the period 1995–2014 (Figure 15), and the simulated spring discharges are in agreement with the observed ones (Figure 16). The simulated results show that the water table is sensitive to interannual variations in recharge, which has been on a long-term downward trend since 1979.

The magnitude of the groundwater level drop, from 1 to 5 m between 1970 and 2014, is a function of hydrodynamic parameters. The drop in the water table has led to a decrease in the discharge of springs (Figure 16). Indeed, the simulated flows of the Guinguette and Nasso springs decrease, respectively, from 2.3 m<sup>3</sup>/s to 1.42 m<sup>3</sup>/s and from 0.6 m<sup>3</sup>/s to 0.37 m<sup>3</sup>/s between 1979 and 2014. The largest decreases occur during periods of zero or very little recharge (below 100 mm/year) over several consecutive years (–0.38 m<sup>3</sup>/s during 1980–1984, –0.20 m<sup>3</sup>/s during 1986–1990 for the Guinguette spring), and since 1979, the decrease in the discharge is continuous with occasional recoveries linked to relatively important recharge (over 100 mm/year).

As pumping was almost nil before the 1990s (Figure 6), the decrease in the springs' discharge was controlled by the decrease in recharge since the 1960s. Indeed, periods of drought (rainfall deficit) are generally accompanied by prolonged periods without recharge or with a low recharge (below 100 mm/year), resulting in a significant drop in the water table. From the 2000s onward, the evolution of the water table and of the discharge rate of the springs is the result of the combined effect of pumping (in the vicinity of the boreholes) and of a lower recharge than in the 1960s. Therefore, it appears that the combined effects of climate and pumping are contributing to the observed deterioration in the groundwater level and the discharge of springs.



**Figure 13.** Examples of simulated groundwater level variations under the influence of pumping.



**Figure 14.** Simulated variation in the discharge of the Guinguette spring under the influence of pumping.

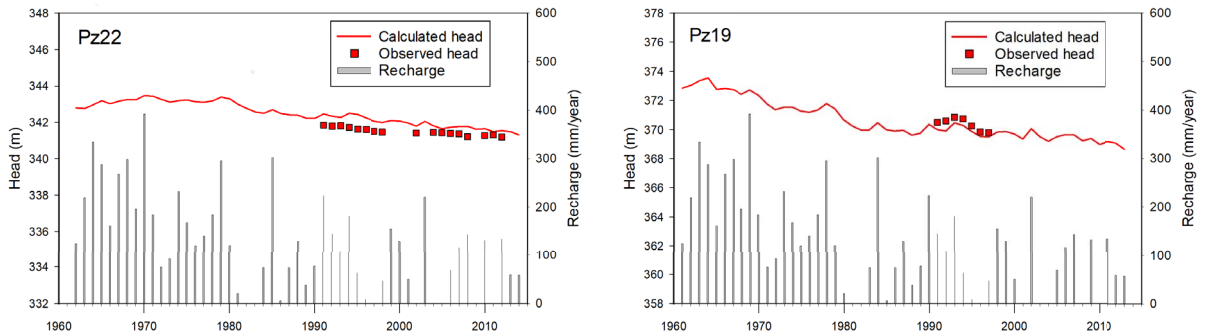
### 5. Discussion

The model run in steady state with Modflow was used to establish the water balance of the extended Kou catchment in the late 1990s. The simulation has shown that half of the inflows in the area come from the diffuse recharge of the aquifer and almost half from the upstream part of the basin through the Mouhoun and the upstream boundaries. The springs, the Kou River and its tributaries drain about a quarter of the flows, while pumping only accounts for approximately 2%.

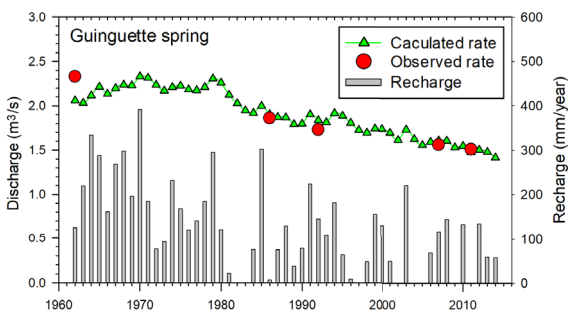
This water balance highlights the importance of inflows from outside the study area. Imposed by the prescribed hydraulic head boundary conditions, themselves derived from the observed groundwater levels, these inputs are indeed consistent with the observed and simulated regional potentiometric map of the south-eastern margin of the Taoudeni sedimentary basin, according to which the Mouhoun, in equilibrium with the main aquifer, is identified as a general drainage axis oriented SW-NE [Dak-

ouré, 2003, Derouane, 2008]. At the upstream end of the Kou watershed, the sedimentary structures extend south to the outcrop limits of the Taoudeni basin in Burkina Faso. Derouane [2008] places the groundwater divide on the topographical ridges, in line with the surface Mouhoun catchment, between 500 and 600 masl, upstream of Orodara and about 25 km from the boundary of the study area. This is where the excellent Ggq aquifer outcrops. The highlands of the Orodara region have been clearly identified as the main recharge area of the regional aquifer by previous geochemical studies in relation to their young age [Dakouré, 2003, Kouanda, 2019], which is also consistent with the regional water table pattern. In addition, these topographical heights are wetter than the Kou watershed (1140 mm/year at the Orodara weather station versus 1025 mm/year at the Bobo-Dioulasso station on average since the 1960s), with a clearer improvement in rainfall conditions after 1990 [Tirogo *et al.*, 2016]. With the rainfall recorded at the Orodara station, higher recharge values of around 150 mm/year are calculated for the study period (compared to 95 mm/year in the Kou catchment), all other conditions being equal. However, these more favourable current recharge conditions are still not sufficient to justify the inflow values upstream of the study area, given the surface area involved (approximately 1100 km<sup>2</sup>). They underline the high uncertainties associated with each term of the water balance calculated with the model (Table 2).

Indeed, the results of the model are highly dependent on the assumptions made. Missing information leads to non-unique solutions to the inverse groundwater flow problem, where different recharge and hydraulic conductivity patterns are possible to



**Figure 15.** Overview of the evolution of hydraulic heads from 1961 to 2014.



**Figure 16.** Simulated discharge of the Guinguette spring from 1961 to 2014.

reproduce the observations. This is evidenced, for example, by the differences with the water balance from the numerical groundwater flow model of Sauret [2008], which calculates 75% higher flows over an area similar to the one considered here, 2/3 of which come from the Mouhoun and Kou rivers, even though the estimated direct recharge is divided by 3. The possibility that the Mouhoun recharges the water table upstream of the study area during the rainy season, when floods are often severe [Kouanda, 2019], remains to be confirmed, as there are no observations of the exchange between the water table and the river in the vicinity and the data series from the Guéni hydrometric station are incomplete and sometimes questionable (see Figure 1). As for the Kou, the direction of groundwater-river exchanges seems to vary spatially from one section to another, according to Sauret [2013], who studied them in the alluvial plain downstream of the Nasso/Guinguette springs.

Modern recharge estimates using various methods (Thornthwaite balance, hydrological or hydro-

geological modelling) have not yet allowed a narrower range of values to be proposed (from 26 to 170 mm/year) and show some disparity, both in magnitude and in spatial distribution [Dakouré, 2003, Derouane, 2008, Sauret, 2008, 2013, Kouanda, 2019]. In particular, recharge estimates are highly dependent on the soil maximum available water capacity used in the Thornthwaite method or the calculation time step. It is therefore difficult to limit the adjustment of the model parameters to hydrodynamic parameters only. A special case are the relief areas coinciding with the presence of groundwater mounds. The relatively shallow depths to water table in these areas were here explained by higher recharge following the sensitivity analysis, given the range of possible recharge values and the observed response of the water table to temporal variations in recharge, rather than by lower hydraulic conductivities. However, to the east of the study area, it was necessary to introduce a low conductivity layer in order to maintain sufficiently high groundwater levels while maintaining hydraulic conductivities appropriate for a surface aquifer and consistent with the results of pumping tests. The presence of such low-permeability layers locally dividing the multilayer reservoir is recognised, particularly at the interface between the GKS and the Gfg [e.g., Derouane, 2008].

A modern automatic parameter estimation method, rather than the two-stage trial-and-error calibration procedure used in this study, would also help to quantify the degree of uncertainty associated with a given prediction of the calibrated model. It would provide valuable feedback on the strengths and weaknesses of the input datasets, allowing future data collection to be optimised. The lessons

learnt from this modelling study will also be relevant to the selection of an adequate level of complexity for future models to be developed in the study area [e.g. Lecca, 2005].

In this respect, although our model provides a quantification of flows that is not possible from groundwater level observations alone, but is consistent with them, it is subject to strong assumptions about the boundary conditions upstream of the domain that intersect with the flow lines. The mismatch between the hydrological and hydrogeological catchments of the Kou has already led us to extend the study area westwards, but it would also benefit from being extended southwards to the boundaries of the Mouhoun catchment, on the aforementioned Orodara heights. Another solution would be to include it in a regional model, such as those previously constructed, which cover the south-eastern edge of the Taoudeni sedimentary basin. This would make it possible to better represent the organisation of groundwater flows, particularly at depth, and to use the model to study scenarios relating to the origin of the water, particularly in the Nasso/Guinguette springs area, and to question the results obtained here concerning the relative influence of climatic variability and abstraction.

Indeed, this exceptional site for the sub-region is also recognised as being particularly vulnerable. It is influenced by the four nearby 250 m<sup>3</sup>/h ONEA wells, which in 2014 tapped the Ggq aquifer for about a third of the study area's water withdrawals. This is evidenced by the significant decline in spring discharge compared to the moderate drop in groundwater levels. In the absence of direct annual recharge, the model also predicts a continuous decline in their discharge, the preservation of which would therefore be partly contingent upon the annual recharge. However, the catchment area of the springs also extends beyond the Kou to the Orodara heights [Dakouré, 2003]. The springs are located along major faults that bring the Ggq aquifer into contact with the SAC1, as indicated by geophysical prospection, and are believed to be fed by upward circulation [Talbaoui, 2009]. Analyses of water and carbon isotopes also confirm the resurgence of a predominantly old water (between 300 and 500 years old) [Dakouré, 2003], which is all the more fragile because of its low renewal rate [Dakouré, 2003, Huneau *et al.*, 2011]. Considering the size of the reservoir

(more than 1000 m thick), the renewable water resources of the sedimentary aquifer in SW Burkina are in fact only about 1% [Gombert, 1998]. The decline in spring discharge could therefore also be due to the depletion of a resource isolated from modern recharge and palaeo-recharged during wetter periods, located in the deepest and most confined areas of the south-eastern margin of the Taoudeni sedimentary basin [Huneau *et al.*, 2011]. However, it is difficult to describe the organisation of the deep groundwater flow, which may be separated from the surface in places, as evidenced by the existence of numerous artesian wells in the basin. Following Derouane [2008], Tirogo [2016] or Kouanda [2019], we can propose a nested flow system according to the scheme of Tóth [1963], with a combination of recent superficial groundwater drained locally in the valleys and ancient water following regional flow paths that ensure the connection at the scale of the sedimentary system between the recharge zones of Orodara and its distant outlet in the Gondo depression [Derouane, 2008], some of which rise locally thanks to the large fault networks. Groundwater flow in this context is then similar to that of large systems with a wide range of water residence times [Maxwell *et al.*, 2016]. This hypothesis, which departs from the single multilayer reservoir model, needs to be confirmed by a better understanding of the basin's deep circulations and the properties of the aquifers that host them, which have not yet been identified by drilling. The development of local versus regional flow systems is sensitive to the exponential decay rate of hydraulic conductivity [Jiang *et al.*, 2009], a sensitivity that could be tested in our study case with the multilayer model extended to the southern recharge areas. By extending it also in time, back to the colder recharge periods mentioned by geochemists [Dakouré, 2003], one could also test the hypothesis of depletion of a palaeo-recharged resource and a system still in equilibrium with past boundary conditions. However, this is not the case for permeable surperficial sandstones and carbonates, for which short response times can be estimated, as well as short-term variability in groundwater levels, significant intra-annual variability, but limited changes on decadal scales [Ascott *et al.*, 2020]. An average transmissivity of  $5 \times 10^{-4}$  m<sup>2</sup>/s and a storage coefficient of  $10^{-4}$  for 30 km length gives a response time of about 5 years. A slower response, associated with lower hydraulic dif-

fusivity over longer distances, is possible at depth but needs to be demonstrated. By drilling deep boreholes through the sedimentary cover to the basement, it would be possible not only to confirm the thickness of the sedimentary cover, but also, by jointly estimating the storage coefficient, to determine the actual volumes of water contained in the sedimentary reservoir and thus its renewal rate, thereby improving the conceptual model of groundwater flow in the study area in the context of global change and increasing anthropogenic pressures.

## 6. Conclusion

Hydrogeological modelling has facilitated the analysis of the hydrodynamic functioning of the aquifer system of the Kou catchment using three factors that may influence it, namely, climate variability, anthropogenic withdrawals, and the spatial variation of the hydrodynamic parameters of the aquifer system. The response of the aquifer system to climatic and anthropogenic forcing is indeed quite heterogeneous all over the basin due to several factors, including the hydrodynamic parameters of the aquifers, the depth to the water table, or the exchanges with the drainage network, particularly the Mouhoun River.

Over the last 20 years, the variations observed in the water table, particularly through the decrease in the discharge of springs, are a combined effect of pumping and recharge evolution. Indeed, pumping increased significantly since the 2000s and the recharge of the last 20 years has often been marked by a succession of deficit years. Simulation performed since the 1960s indicates that these deficit years, more frequent after 1979, caused a drop in the water table, accompanied by a significant drop in the discharge of springs. The wet years, which remained isolated despite a slight increase in rainfall after the 1990s, were not sufficient to raise the water table, which is marked by a long-term downward trend. The aquifer system is sensitive to the variability of recharge, particularly to the cumulative deficit over several years. Pumping has also an impact on the water table and on the discharge of springs, even if the quantities of water withdrawn appear negligible in view of their small contribution to the water balance of the catchment. In particular, the concentration of boreholes with high flow rates in the same area significantly impacts the water table and could

in the long term cause a considerable decrease or even a drying up of some springs. Pumping is estimated to be responsible for up to 40% of the decline in low flows in springs during the period 1995–2014. Given the high contribution of the springs to the base flow of the Kou River (80%), pumping would thus affect their contribution to surface waters and could accelerate the depletion of the water resource in general [Sapriza-Azuri *et al.*, 2015], especially considering that these waters are ancient and deep [Huneau *et al.*, 2011].

In this context in which pumping is constantly increasing due to population growth, its impact on groundwater levels, although moderate so far, could be more important in the future if measures are not taken now to ameliorate groundwater declines. Also, although climate model projections of precipitation for the West African region in the 21st century indicate relatively small changes in the upstream part of the Mouhoun basin [e.g., Biasutti, 2013, Monerie *et al.*, 2020], regionally heterogeneous changes in water availability are expected [Sylla *et al.*, 2018] and there are large uncertainties in predicting the impact of climate change on runoff [Roudier *et al.*, 2014], let alone groundwater recharge. Indeed, the trend towards higher rainfall intensities accompanied by lower frequency of rainfall events is likely to result in more intense runoff and flooding events [Panthou *et al.*, 2018], which could deteriorate the recharge of the aquifer system [Kouanda, 2019]. Given its sensitivity to the high variability in recharge, as shown in this study, further assessment of groundwater level evolution in the context of climate change is also required.

Therefore, appropriate measures will have to be taken for adequate quantitative management of the water resource of the Kou River catchment to ensure its long-term sustainability. It will be necessary to turn to a more efficient use of the resource (Sustainable Development Goals – SDG 6.4) and to integrated management (SDG 6.5). A key part of the water resource management system is the collection, storage and analysis of the relevant groundwater data, including water levels and abstraction data, which must be continued over the long term and at the catchment scale. Given the high variability of the water table behaviour at the scale of the watershed, it is important to develop and maintain a network of well-distributed monitoring wells coupled with the

present model. In addition, as this study is based on data from shallow boreholes (less than 300 m), it would be appropriate to carry out further investigations over the entire sedimentary basin to gain a better understanding of the different aquifer levels at depth, their properties, and water potential. This could be used to improve the model to better fulfil its role as a resource management tool, and to identify areas of high productivity with modern groundwater where future high-flow-rate pumping could take place. Based on the results of the model, it appears indeed necessary to diversify the areas of withdrawals in the future to ensure the sustainability of the water resource, and to seek out the most productive and least vulnerable areas where pumping could be operated at optimal rates with the least impact, as far as the infrastructure allows. In this respect, the distance between production and demand must also be considered, depending on the different uses of the water resource. It is particularly important to preserve the area of springs on which the downstream flow of the Kou River depends, as well as the irrigated agriculture in its alluvial plain [Traoré, 2012, Wellens *et al.*, 2013]. Although the extent of managed aquifer recharge practice is relatively limited in Africa, and almost non-existent in West Africa [Ebrahim *et al.*, 2020], the potential for artificial recharge could be explored for its possible benefits in stabilising storage decline and mitigating the effects of inter-annual variability in water availability. However, the sandstone aquifer system is vulnerable to anthropogenic activity, particularly in highly-populated peri-urban areas, where subsurface contaminant loads may not meet the quality requirements for water used for recharge [Huneau *et al.*, 2011, Sako *et al.*, 2020].

The efforts already made in the field of integrated water resources management, such as the irrigation schemes in the Kou Valley [Wellens *et al.*, 2013], should be continued and extended to the entire catchment area. The modelling tool developed in this study could be used appropriately as a management tool to test scenarios for maintaining the coexistence of different water uses and defining allocation rules. It can also be used to predict the evolution of the water table under climate change and increased water abstraction. The current modelling, combined with alert thresholds for aquifer drawdown triggered by a remote monitoring network, can help to stabilise groundwater levels before they reach critical

levels, including when assessing new water project proposals. Volumetric quotas can be proposed and adjusted according to the state of resources at the beginning of the dry season [e.g., Verley, 2020]. The use of the model as an operational decision support tool to address issues raised by stakeholders involved in groundwater development and resource management will require further support and investment in applied problem-oriented research at the sub-regional level. Key to the success of this approach is the training of local operators in the use of the modelling tool, under the supervision of an expert hydrogeologist and in closer association with the academic sector.

### Conflicts of interest

The authors declare no competing financial interest.

### Acknowledgements

This work was carried out with the financial support of Institut de Recherche pour le Développement (IRD) through the *Programme Doctoral International "Modélisation des Systèmes Complexes" (PDI/MS)*, of Office National de l'Eau et de l'Assainissement (ONEA) and SIREA France. The realization of this work was possible thanks to the numerous data made available by the National Meteorological Agency (ANAM), Direction Générale des Ressources en Eau (DGRE) in Burkina Faso, ONEA, and the Regional Directorate in charge of water in "Hauts Bassins". We would also like to thank the three reviewers for their insightful comments, which have helped us to improve the quality of this article.

### References

- Albergel, J. (1987). Sécheresse, désertification et ressources en eau de surface – Application aux petits bassins du Burkina Faso. In *The Influence of Climate Change and Climatic Variability on the Hydrologic Regime and Water Resources. Proceedings of the Vancouver Symposium*, volume 168, pages 355–365. IAHS Publ., Wallingford, UK.
- Amogu, O., Descroix, L., Yéro, K. S., Le Breton, E., Mamadou, I., Ali, A., Vischel, T., Bader, J.-C., Moussa, I. B., Gautier, E., Boubkraoui, S., and Belleudy, P. (2010). Increasing river flows in the Sahel? *Water*, 2, 170–199.

- ANTEA (2012). Gestion Intégrée et Concertée des Ressources en Eau des Systèmes Aquifères d'Iullemeden, de Taoudeni/Tanezrouft et du fleuve Niger-(GICRESAIT)-Hydrogéologie (Rapport de synthèse). Sahara and Sahel Observatory, Tunis.
- Ascott, M. J., Macdonald, D. M. J., Black, E., Verhoef, A., Nakohoun, P., Tirogo, J., Sandwidi, W. J. P., Bliefernicht, J., Sorensen, J. P. R., and Bossa, A. Y. (2020). In situ observations and lumped parameter model reconstructions reveal intra-annual to multidecadal variability in groundwater levels in sub-Saharan Africa. *Water Resour. Res.*, 56, article no. e2020WR028056.
- Biasutti, M. (2013). Forced Sahel rainfall trends in the CMIP5 archive. *J. Geophys. Res. Atmos.*, 118, 1613–1623.
- Bredehoeft, J. (2002). The water budget myth revisited: why hydrogeologists model. *Groundwater*, 40, 340–345.
- Bronner, G., Roussel, J., and Trompette, R. (1980). Genesis and geodynamic evolution of the Taoudeni cratonic basin (upper Precambrian and Paleozoic) western Africa. *Dyn. Plate Inter. Geodyn. Ser. Amer. Geophys. Union*, 1, 81–90.
- Calow, R. and MacDonald, A. (2009). *What Will Climate Change Mean for Groundwater Supply in Africa?* Overseas Development Institute (ODI), London.
- Condon, L. E., Kollet, S., Bierkens, M. F. P., Fogg, G. E., Maxwell, R. M., Hill, M. C., et al. (2021). Global groundwater modeling and monitoring: Opportunities and challenges. *Water Resour. Res.*, 57, article no. e2020WR029500.
- Cuthbert, M. O., Taylor, R. G., Favreau, G., Todd, M. C., Shamsudduha, M., Villholth, K. G., Macdonald, A. M., Scanlon, B. R., Kotchoni, D. O. V., Vouillamoz, J.-M., Lawson, F. M. A., Adjomayi, P. A., Kashaigili, J., Seddon, D., Sorensen, J. P. R., Ebrahim, G. Y., Owor, M., Nyenje, P. M., Nazoumou, Y., Goni, I., Ousmane, B. I., Sibanda, T., Ascott, M. J., Macdonald, D. M. J., Agyekum, W., Kousoubé, Y., Wanke, H., Kim, H., Wada, Y., Lo, M.-H., Oki, T., and Kukuric, N. (2019). Observed controls on resilience of groundwater to climate variability in sub-Saharan Africa. *Nature*, 572, 230–234.
- Dakouré, D. (2003). *Etude hydrogéologique et géochimique de la bordure sud-est du bassin sédimentaire de Taoudéni (Burkina Faso - Mali) - Essai de modélisation*. PhD thesis, Université Pierre et Marie Curie Paris VI, France.
- de Graaf, I. E. M., Gleeson, T., (Rens) van Beek, L. P. H., Sutanudjaja, E. H., and Bierkens, M. F. P. (2019). Environmental flow limits to global groundwater pumping. *Nature*, 574(7776), 90–94.
- De Longueville, F., Hountondji, Y.-C., Kindo, I., Gemenne, F., and Ozer, P. (2016). Long-term analysis of rainfall and temperature data in Burkina Faso (1950–2013). *Int. J. Climatol.*, 36, 4393–4405.
- de Marsily, G. (2008). Eau, changements climatiques, alimentation et évolution démographique. *Rev. Sci. Eau.*, 21, 111–128.
- de Marsily, G. (2020). Will we soon run out of water? *Ann. Nutr. Metab.*, 76, 10–16.
- Derouane, J. (2008). Modélisation hydrogéologique du bassin sédimentaire. Programme VREO, SOFRECO - SAWES, Bobo Dioulasso.
- Descroix, L., Genthon, P., Peugeot, C., Mahé, G., Abdou, M. M., Vandervaere, J.-P., Mamadou, I., Tanimoun, B., Amadou, I., Galle, S., Vischel, T., Quantin, G., Dacosta, H., Bodian, A., Nazoumou, Y., and Zannou, A. (2015). Paradoxes et contrastes en Afrique de l'Ouest: impacts climatiques et anthropiques sur les écoulements. *Géologues*, 187, 47–52.
- Descroix, L., Mahé, G., Lebel, T., Favreau, G., Galle, S., Gautier, E., Olivry, J.-C., Albergel, J., Amogu, O., Cappelaere, B., Dessouassi, R., Diedhiou, A., Le Breton, E., Mamadou, I., and Sighomnou, D. (2009). Spatio-temporal variability of hydrological regimes around the boundaries between Sahelian and Sudanian areas of west Africa: a synthesis. *J. Hydrol.*, 375, 90–102.
- Dong, Y., Li, G., and Xu, H. (2012). An areal recharge and discharge simulating method for MODFLOW. *Comput. Geosci.*, 42, 203–205.
- Ebrahim, G. Y., Lautze, J. F., and Villholth, K. G. (2020). Managed aquifer recharge in Africa: taking stock and looking forward. *Water*, 12, article no. 1844.
- Favreau, G., Cappelaere, B., Massuel, S., Leblanc, M., Boucher, M., Boulain, N., and Leduc, C. (2009). Land clearing, climate variability, and water resources increase in semiarid southwest Niger: a review. *Water Resour. Res.*, 45, article no. W00A16.
- Gombert, P. (1998). Programme RESO. IWACOBURGEAP, Bobo-Dioulasso, Burkina Faso.
- Harbaugh, A. W. (2005). *MODFLOW-2005, The U.S.*

- Geological Survey Modular Ground-Water Model—The Ground-Water Flow Process: U.S. Geological Survey Techniques and Methods 6-A16*. U.S. Geological Survey, Virginia.
- He, C., Liu, Z., Wu, J., Pan, X., Fang, Z., Li, J., and Bryan, B. A. (2021). Future global urban water scarcity and potential solutions. *Nat. Commun.*, 12, article no. 4667.
- Howard, K. and Griffith, A. (2009). Can the impacts of climate change on groundwater resources be studied without the use of transient models? *Hydrol. Sci. J.*, 54, 754–764.
- Huneau, E., Dakouré, D., Celle-Jeanton, H., Vitvar, T., Ito, M., Traore, S., Compaore, N. F., Jirakova, H., and Le Coustumer, P. (2011). Flow pattern and residence time of groundwater within the south-eastern Taoudeni sedimentary basin (Burkina Faso, Mali). *J. Hydrol.*, 409, 423–439.
- INSD (2022). Cinquième Recensement Général de la Population et de l’Habitation du Burkina Faso. Synthèse des résultats définitifs. Technical report, Comité National du Recensement, Institut National de la Statistique et de la Démographie, Ouagadougou, Burkina Faso, [http://www.insd.bf/contenu/documents\\_rgph5/Rapport%20resultats%20definitifs%20RGPH%202019.pdf](http://www.insd.bf/contenu/documents_rgph5/Rapport%20resultats%20definitifs%20RGPH%202019.pdf).
- Jiang, X.-W., Wan, L., Wang, X.-S., Ge, S., and Liu, J. (2009). Effect of exponential decay in hydraulic conductivity with depth on regional groundwater flow. *Geophys. Res. Lett.*, 36, article no. L24402.
- Knowling, M. J., Werner, A. D., and Herckenrath, D. (2015). Quantifying climate and pumping contributions to aquifer depletion using a highly parameterised groundwater model: Uley South Basin (South Australia). *J. Hydrol.*, 523, 515–530.
- Kouanda, B. (2019). *Modélisation Intégrée du Complexe Mouhoun supérieur-Sourou dans le contexte des changements climatiques*. PhD thesis, Institut International d’Ingénierie de l’Eau et de l’Environnement, Ouagadougou, Burkina Faso.
- Lebel, T. and Ali, A. (2009). Recent trends in the central and western Sahel rainfall regime (1990–2007). *J. Hydrol.*, 375, 52–64.
- Lecca, G. (2005). Automatic calibration of CODESA-3D using PEST. SWIMED Deliverable 8, INCO Programme contract ICA3-CT2002-10004.
- Leduc, C., Favreau, G., and Schroeter, P. (2001). Long-term rise in a Sahelian water-table: the continental terminal in south-west Niger. *J. Hydrol.*, 243, 43–54.
- Mahé, G. (2009). Surface/groundwater interactions in the Bani and Nakambe rivers, tributaries of the Niger and Volta basins, west Africa. *Hydrol. Sci. J.*, 54, 704–712.
- Maxwell, R. M., Condon, L. E., Kollet, S. J., Maher, K., Haggerty, R., and Forrester, M. M. (2016). The imprint of climate and geology on the residence times of groundwater. *Geophys. Res. Lett.*, 43, 701–708.
- Monerie, P.-A., Sanchez-Gomez, E., Gaetani, M., Mohino, E., and Dong, B. (2020). Future evolution of the Sahel precipitation zonal contrast in CESM1. *Clim. Dyn.*, 55, 2801–2821.
- Moniod, E., Pouyaud, B., and Sechet, P. (1977). *Le Bassin du Fleuve Volta*. Monographie Hydrologique ORSTOM. ORSTOM, Paris.
- Nicholson, S. E. (1978). Climatic variations in the Sahel and other African regions during the past five centuries. *J. Arid Env.*, 1, 3–24.
- Nicholson, S. E., Fink, A. H., and Funk, C. (2018). Assessing recovery and change in west Africa’s rainfall regime from a 161-year record. *Int. J. Climatol.*, 38, 3770–3786.
- Nka, B. N., Oudin, L., Karambiri, H., Paturel, J. E., and Ribstein, P. (2015). Trends in floods in West Africa: analysis based on 11 catchments in the region. *Hydrol. Earth Syst. Sci.*, 19, 4707–4719.
- Ouédraogo, C. (1983). *Etude géologique des formations sédimentaires du bassin précambrien supérieur et paléozoïque du Taoudéni en Haute Volta*. PhD thesis, Université de Poitiers, France.
- Ouédraogo, C. (2006). Synthèse géologique de la région ouest du Burkina Faso (Rapport technique). Programme VREO, SOFRECO-SAWES, Bobo Dioulasso, Burkina Faso.
- Palausi, G. (1957). *Projet d’amélioration de l’adduction d’eau de Bobo-Dioulasso (Haute-Volta) : aspect géologique et hydrogéologique du problème*. Direction Fédérale des Mines et de la Géologie, Ouagadougou, Burkina Faso.
- Panthou, G., Lebel, T., Vischel, T., Quantin, G., Sane, Y., Ba, A., Ndiaye, O., Diongue-Niang, A., and Diopkane, M. (2018). Rainfall intensification in tropical semi-arid regions: the Sahelian case. *Environ. Res. Lett.*, 13(6), article no. 064013.
- Paturel, J.-E., Koukponou, P., Ouattara, F., L’Aour, A., Mahé, G., and Cres, F. N. (2002). Variabilité du climat du Burkina Faso au cours de la seconde moitié du XXème Siècle. *Sud. Sci. Technol.*, 8, 41–49.



- Paturel, J.-E., Servat, É., Lubès-Niel, H., and Delatre, M.-O. (1997). Variabilité climatique et analyse de séries pluviométriques de longue durée en Afrique de l'Ouest et centrale non sahélienne. *C. R. Acad. Sci. - Ser. IIA - Earth Planet. Sci.*, 325, 779–782.
- Roudier, P., Ducharne, A., and Feyen, L. (2014). Climate change impacts on runoff in west Africa: a review. *Hydrol. Earth Syst. Sci.*, 18, 2789–2801.
- Sako, A., Sawadogo, S., Nimi, M., and Ouédraogo, M. (2020). Hydrogeochemical and pollution characterization of a shallow glauconitic sandstone aquifer in a peri-urban setting of Bobo-Dioulasso, southwestern Burkina Faso. *Environ. Earth. Sci.*, 79, article no. 296.
- Sapriza-Azuri, G., Jódar, J., Carrera, J., and Gupta, H. V. (2015). Toward a comprehensive assessment of the combined impacts of climate change and groundwater pumping on catchment dynamics. *J. Hydrol.*, 529, 1701–1712.
- Sauret, E. S. G. (2008). Contribution à la compréhension du fonctionnement hydrogéologique du système aquifère dans le bassin du Kou (Mémoire de DEA). Université de Liège, Belgium.
- Sauret, E. S. G. (2013). *Etude des potentialités hydrogéologiques d'une plaine alluviale en relation avec les eaux souterraines et de surface dans un contexte d'agriculture irriguée (Burkina Faso)*. PhD thesis, Université de Liège, Belgium.
- Scibek, J. and Allen, D. (2006). Comparing modelled responses of two high-permeability, unconfined aquifers to predicted climate change. *Glob. Planet. Change*, 50, 50–62.
- Servat, E., Paturel, J.-E., Lubès-Niel, H., Kouamé, B., and Masson, J.-M. (1997). Variabilité des régimes pluviométriques en Afrique de l'Ouest et centrale non sahélienne. *C. R. Acad. Sci. - Ser. IIA - Earth Planet. Sci.*, 324, 835–838.
- Sogreah Ingénierie (1994). Etude des ressources en eau souterraine de la zone sédimentaire de la région de Bobo-Dioulasso (Rapport technique). Ministère de l'eau, Burkina Faso.
- Sylla, M. B., Pal, J. S., Faye, A., Dimobe, K., and Kunstmann, H. (2018). Climate change to severely impact west African basin scale irrigation in 2 °C and 1.5 °C global warming scenarios. *Sci. Rep.*, 8, article no. 14395.
- Talbaoui, M. (2009). Etude des périmètres de protection des sources de Nasso et des forages ONEAI et ONEAI. Rapport de la mission de mars 2009, Programme de Valorisation des Ressources en Eau de l'Ouest VREO, Bobo-Dioulasso, Burkina Faso, 24 p.
- Thornthwaite, C. W. (1948). An approach toward a rational classification of climate. *Geogr. Rev.*, 38, 55–94.
- Tirogo, J. (2016). *Etude du fonctionnement hydrodynamique de l'aquifère sédimentaire du bassin du Kou au sud-ouest du Burkina Faso*. PhD thesis, Université Pierre et Marie Curie Paris VI, France.
- Tirogo, J., Jost, A., Biao, A., Valdes-Lao, D., Kousoubé, Y., and Ribstein, P. (2016). Climate variability and groundwater response: a case study in Burkina Faso (west Africa). *Water*, 8, 171–191.
- Tóth, J. (1963). A theoretical analysis of groundwater flow in small drainage basins. *J. Geophys. Res.*, 68(16), 4795–4812.
- Traoré, F. (2012). *Optimisation de l'utilisation des ressources en eau du bassin du Kou pour des usages agricoles*. PhD thesis, Université de Liège, Belgium.
- Trompette, R. (1973). Le Précambrien supérieur et le Paléozoïque inférieur de l'Adrar de Mauritanie (bordure occidentale du bassin de Taoudeni, Afrique de l'Ouest). Un exemple de sédimentation de craton. Etude stratigraphique et sédimentologique. TOME 2 (Séries 2 et 3. Stratigraphie). Université de Provence, Aix-Marseille I.
- UN-WWDR (2021). *United Nations World Water Development Report 2021: Valuing Water*. UNESCO, Paris.
- USAID (2014). *Climate Change and Water Resources in West Africa: An Assessment of Groundwater Management*. African and Latin American Resilience to Climate Change, Burlington.
- Verley, F. (2020). Lessons from twenty years of local volumetric groundwater management: the case of the Beauce aquifer, central France. In Rinaudo, J. D., Holley, C., Barnett, S., and Montginoul, M., editors, *Sustainable Groundwater Management. Global Issues in Water Policy*, volume 24. Springer, Cham.
- Wellens, J., Diallo, M., Dakouré, D., and Compaoré, N. F. (2007). Renforcement structurel de la capacité de gestion des ressources en eau pour l'agriculture dans le bassin du Kou. Technical Report 1, AAPEFE-DRI/CGRI, Bobo-Dioulasso, Burkina Faso. 127 p.
- Wellens, J., Traoré, F., Diallo, M., and Tychon, B.

- (2013). A framework for the use of decision-support tools at various spatial scales for the management of irrigated agriculture in west-Africa. *Agri. Sci.*, 4, 9–15.



Research article

Geo-hydrological Data & Models

# Predictability of water resources with global climate models. Case of Northern Tunisia

Mustapha Besbes<sup>Ⓢ, a</sup> and Jamel Chahed<sup>Ⓢ, \*, b</sup>

<sup>a</sup> University of Tunis El Manar and Tunisian Academy of Sciences, Letters and Arts, Tunisia

<sup>b</sup> University of Tunis El Manar, National Engineering School of Tunis, Tunisia

E-mails: besbesmustapha9@gmail.com (M. Besbes), jamel.chahed@enit.utm.tn (J. Chahed)

**Abstract.** The objective of the research is to explore the predictability of water resources directly with GCMs by analysing long-term effects of climate change on Northern Tunisia's blue and green water. Hydrologic impacts rely on a rainfall-runoff lumped model using outputs of CMIP6 GCMs within the framework of the ssp2-45 scenario. Among the 30 CMIP6 models, the composite cnrm-esm2-1 and fgoals-g3 best restore observed runoff from 1995 to 2014 and give the best GCM. Hydrologic projections 2015–2100 show significant drops in rainfall (9%), runoff (21%), groundwater recharge (15%), as well as for green water (6%). The results show that the use of raw GCMs predictions on large basins is possible and provides precisions comparable to what is produced when using Regional Climate Models in medium size basins.

**Keywords.** Climate change, Global climate model, Earth system model, Tunisia, Blue water, Green water.

*Manuscript received 15 September 2022, revised 6 February 2023 and 22 March 2023, accepted 11 May 2023.*

*“The effects of climate change during the 21st Century are relatively well predicted with respect to temperature, but their hydrological effects are much less certain” [de Marsily, 2008].*

## 1. General introduction

Tunisia uses 80% of its national blue water resource for irrigation, which has almost reached its maximum water allocation. As for rain-fed agriculture, the green water, it plays an essential role in national food security: 65% of the national agricultural production and 80% of agricultural exports. However, blue and green water resources are still insufficient to meet the population's food needs, and Tunisia must import foodstuffs, that contain virtual water [Allan, 1998, Renault and Wallender, 2000, Hoekstra, 2003, Oki et al.,

2003]. Previous works on water prospects in Tunisia have shown [Chahed et al., 2008, Besbes et al., 2014, 2019b] that it is necessary to evolve the water paradigm towards a holistic vision that expands water resources to include green and virtual water; and consequently redefine the issue of Water Security in climate change context.

Global Climate Models (GCMs) or Global general circulation models are currently the best tools to anticipate the future impacts of climate change. Models with biophysical and biogeochemical processes, called Earth System Models (ESMs), include land and ocean carbon cycle, atmospheric chemistry, and other biogeochemical cycles [Watanabe et al.,

\* Corresponding author.

2011, Collins et al., 2011]. Climate Change effects [Collins et al., 2011, IPCC, 2021] are fairly well predicted as far as temperature is concerned. But GCMs have some difficulty in accurately estimating precipitation, a highly heterogeneous, non-smooth, and spatially discontinuous field [Hughes and Guttorp, 1994, Mandal et al., 2016]. However, to be accurate, hydrological models, which transform precipitation into runoff, need unbiased precipitation data [Mandal et al., 2016, Maraun and Widmann, 2018, Bruyère et al., 2014].

Much work has recently been devoted to producing aggregated predictions obtained by applying weightings to the results of a given ensemble of GCMs [McSweeney et al., 2015, Knutti et al., 2017, Herger et al., 2018, Laurent et al., 2021, Zhang et al., 2022]. Ensembles improve the results by smoothing out the uncertainties and biases of single models' predictions [Wang et al., 2021]. The other way to improve climate models' predictions is sought by increasing GCMs resolution or downscaling to build Regional Climate Models (RCMs). Nevertheless, Somot et al. [2018] argue that the reference datasets used for the evaluation of GCMs are often not suitable for RCMs, and some are run using specific higher-resolution data [Fathalli et al., 2019, Somot et al., 2018, Zhang et al., 2022]. Kim et al. [2020] and Cos et al. [2022] indicate that climate model errors propagate through the RCMs to become a fundamental source of uncertainty. The issue of climate models' uncertainties is especially concerning for impact studies at the watershed scale, for which unbiased precipitation is needed at small grids, requiring the implementation of RCMs. Many bias reduction methods have been applied in order to make climate models' outputs suitable for impact studies [Foughali et al., 2015, Dakhlaoui et al., 2022, Switanek et al., 2022].

Studies of climate change impacts with RCMs have been developed on medium-scale catchments in the Mediterranean basin [Deidda et al., 2013, Terink et al., 2013]. Some studies use RCMs outputs to run hydrological models and simulate observed runoff. In studies on Northern Tunisia, Ensemble RCMs provide inputs of hydrological models (temperature and precipitation): four RCMs from the European program ENSEMBLES [Foughali et al., 2015] and eleven RCMs from the Euro-Cordex Project [Dakhlaoui et al., 2022]. Otherwise, GCMs outputs are also used directly in predicting environmen-

tal impacts, in particular on large areas or basins, [Ramirez-Villegas et al., 2013, Farsani et al., 2019, Wang et al., 2021, Shokouhifar et al., 2022, Hamed et al., 2022, Li et al., 2022].

Despite uncertainties, Climate Models reveal significant impacts on the Mediterranean basin which would be among the most vulnerable regions in the world, with on average a sharp decline in rainfall and higher evapotranspiration due to the temperature increase. Countries on the southern and eastern shores are likely to be the most severely affected, resulting in an increase in food dependence [Besbes et al., 2010, de Marsily and Abarca-del Rio, 2016, de Marsily, 2020]. Data analysis for Tunisia shows an increase in mean temperature of 1.2 °C over the past century; while for the same period, no significant change in mean rainfall is detected even though an increase in variability (highest standard deviation) is observed for the latest period [King et al., 2007]. With growing concerns about impacts on the future of water resources in Tunisia, more specific climate change modeling and predictions are intended, to prepare adaptation and remediation measures [AFD-MA, 2021].

In the present research, we analyze the long-term effects of climate change using the predictions from CMIP6 on Northern Tunisia's water resources, including blue and green water. The region represents the essential source of surface water, which gives it the qualifier "water tower" of Tunisia. It is also the cereal region of the country, mainly cultivated in rain-fed: it is its "attic". Based on hydrological modeling, the analysis aims at determining the foreseeable climate change effect on the overall water resources of the northern region of Tunisia. The hydrologic model, a lumped rainfall-runoff model, is first calibrated and validated based on historical simulations, extensively confronted with field data. The simulations carried out by 30 GCMs that participated in the CMIP6 exercise provide annual chronological mean temperature and precipitation series for each sub-region (governorate). Among the various combinations of radiative forcing defined in the CMIP6 experiences, we consider the medium IPCC emission scenario SSP2-4-5, designed to prolong current trends. The analysis of these simulations on the hydrological model and comparison of their outputs to observed runoff data will lead to selecting the GCMs, the re-

sults of which fit best with observations. The couple “selected GCM-hydrologic model” is then applied in a prospective projection (2015–2100), where GCM outputs (temperature, precipitation) serve as inputs for the hydrologic Model, the outputs of which are surface runoff, aquifer recharge, and actual evapotranspiration.

GCM-hydrologic model coupling is not extensively used in the literature, considering that the bias induced by large dimensions of GCMs meshes is likely to introduce unacceptable uncertainties at the catchment areas scale. However, the downscaling alternative to GCMs usually comes with bias correction, the implementation of which is long and not devoid of uncertainty. Our objective is to explore to what extent, the direct GCM-hydrologic models’ coupling, more consistent and more accessible to interpretation, can be relevant for hydrological studies. It is to answer the following question: is it possible to predict Water Resources only with GCMs, without downscaling, and what would be the resulting uncertainties?

## 2. Northern Tunisia and its climate

### 2.1. *Setting up the context: Tunisia*

Tunisia has a surface area of 165,000 km<sup>2</sup> and a population of 12 M (2020). 5 M ha are cultivated area, 11% of which can be irrigated; 4.5 M ha are rangeland and grassland; 1.6 M ha are forests and steppes; 4.7 M ha are unproductive lands (wetlands, deserts, urban areas...). Agricultural needs are covered by domestic production at 70%; the country is a net importer of food products. Due to random climatic conditions, the level of cereal self-sufficiency ranges from 15 to 60% depending on annual precipitation. Tunisia has average rainfall estimated at 37 km<sup>3</sup>/y. According to Besbes et al. [2019b], the total hydraulic resources (blue water) are 4.8 km<sup>3</sup>/y: 2.7 km<sup>3</sup>/y are surface runoff, of which 85% runs off the northern Tunisia basin. In 2010, groundwater withdrawals were estimated at 2 km<sup>3</sup>, representing an average operating rate of 95% of renewable groundwater resources estimated at 2.1 km<sup>3</sup>/y, and many aquifer systems are severely overexploited. The water resource of rain-fed agriculture (green water), available for evaporation and consumption by plants, related to arable land (5 million ha), is estimated on average

at 13 km<sup>3</sup>/y, which increases to 19 km<sup>3</sup>/y when incorporating rangelands. The total blue water withdrawals were estimated at 3.0 km<sup>3</sup> in 2015, of which 21% by municipalities, 76% by irrigation, and 3% by industries [Besbes et al., 2019b].

Extending over seven degrees of latitude, Tunisia offers a succession of hydro-climatic types: (a) humid and sub-humid in the North, (b) semi-arid and arid in central Tunisia, and (c) hyper-arid or Saharan for the entire South. These climatic types are articulated with the three major hydrological provinces (Figure 1): In the North, the reliefs, oriented SW–NE, frame the country’s most fertile plains and make this area the wettest in the country. The main river, Oued Medjerda, fully controlled by a series of large storage dams, crosses the region from west to east to unplug the North of the Tunis Gulf.

From North to South, this variety results in a range of rainfall regimes from more than 1000 mm/y in the North, up to 50 mm/y at the southern tip of the country, with an overall average of 250 mm/y. The annual average temperature for the country is 19.5 °C, which varies by region and season between 12 °C in winter and 32 °C in summer.

### 2.2. *The study area: hydro-meteorological series and data sources*

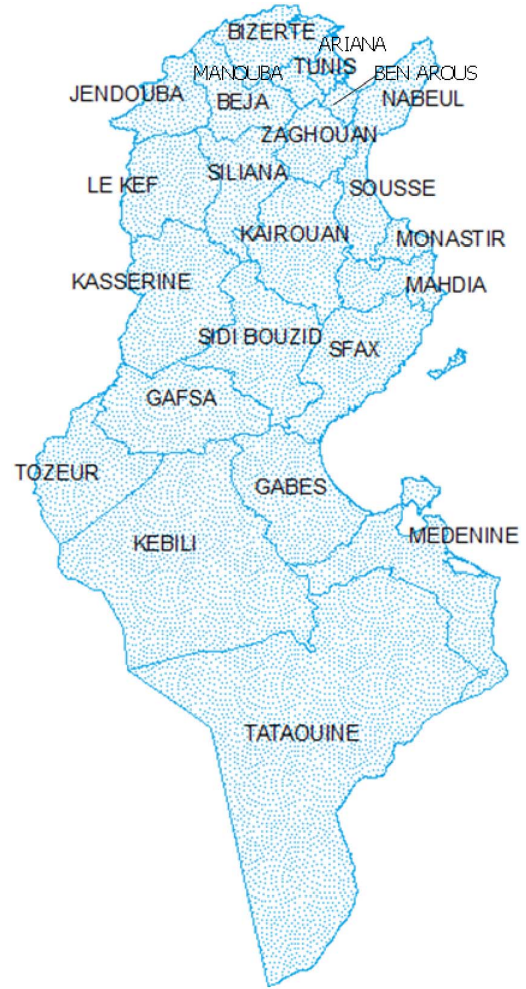
The study area, Northern Tunisia, forms an ensemble of three large hydrologic basins: Medjerda, Far North—Ichkeul and Cap Bon—Miliane (Figure 1), extending over 29,000 km<sup>2</sup> that is 18% of the area of Tunisia. With an average rainfall estimated at 510 mm/y, Northern Tunisia receives 40% of the national average rainfall, provides more than 80% of surface water, 40% of groundwater resources, and a large part of green water, used in particular to grow cereals. The region covers 11 governorates among the 24 in the country, namely those of Tunis, Ariana, BenArous, Manouba, Bizerte, Nabeul, Beja, Jendouba, Le Kef, Siliana, Zaghuan (Figure 2).

Monthly precipitation and temperature data at the national level for the 24 Tunisian governorates covering the period 1901–2020 can be downloaded from the Climate Change Knowledge Portal (CCKP) developed by the World Bank Group [WBG, 2022], which data source is the Climatic Research Unit (CRU) gridded time series, presented on a 0.5° grid. These monthly time series are developed by interpolation



**Figure 1.** Tunisia, the three major hydrological regions, with the Northern Tunisia study area limit [from Besbes et al., 2019b].

of values derived from observational data on temperature and rainfall in thousands of reference weather stations worldwide [Harris et al., 2020]. For Tunisia in particular, the original series of monthly rainfall depth are developed, criticized, and analyzed by the Tunisian National Institute of Meteorology [INM, 2022]. These data have been reanalyzed and integrated into the global networks respectively of NOAA in its global inventory [NOAA, 2022], and WMO [WMO, 2022]. All these databases are available to the



**Figure 2.** The governorates map of Tunisia.

public. Among the 7500 reference climate stations the rainfall data of which are published by the WMO, NOAA, and CRU, with historical series of more than 100 years, Tunisia has 15 rainfall reference stations, 8 of which cover (directly or in the immediate vicinity) the Northern Tunisia basin over periods exceeding one hundred years with monthly rainfall series.

Moreover, the European Centre of Medium-Range Weather Forecasts (ECMWF) provides the last analysis ERA5 of climate variables, covering the Earth on a 30 km grid. ERA5, which combines global historical observations into estimates by improved modeling and data processing systems [Jiao et al., 2021], has been shown to be among the best-performing reanalysis products [Hersbach et al., 2020, Tarek et al., 2020,

**Table 1.** Statistical parameters of CRU and ERA5 vs. DGRE precipitation series for the study area

Source of precipitation series (1970–2020)	CRU	ERA5	DGRE
Mean value mm/y	686	504	512
Standard deviation mm/y	136	108	107
Correlation coefficient	0.43	0.44	1
Root mean square error RMSE mm/y	217	110	
Average relative bias	24.9%	−1.4%	

Jiang et al., 2021]. For all governorates of Northern Tunisia, the CCKP provides also historical series of precipitation 1970–2020 labeled ERA5 [WBG, 2022].

Finally, the official Tunisian institution for water resources management, the General Directorate of Water Resources (DGRE, Ministry of Agriculture) has, among other missions, to maintain and develop national hydro-meteorological networks for Precipitation, Runoff, and Groundwater. The produced time series are regularly reviewed and updated. The latest update constitutes the main reference source for our case study [DGRE, Prointec-Comete, 2019] and will be considered as the historical reference observations series. However, these observations show some missing values. In order to complete missing data over 1970–2020, we use the global series CRU and ERA5. Table 1 presents the statistical parameters of these series in comparison with DGRE, which indicate a better similarity between ERA5 and DGRE reference observations, so that missing observations could be filled from ERA5 series.

### 3. Historical trends and possible future scenarios

#### 3.1. Observed series and current trends

Northern Tunisia was subject to a significant increase in temperature over the past 50 years, with average increases of 0.5 °C per decade, a total increase of 2.5 °C since 1970, including over the northern regions (Figure 3). Annual rainfall in northern Tunisia is variable but its long-term trend has been stable since 1970 (Figure 3).

#### 3.2. Possible climate future scenarios

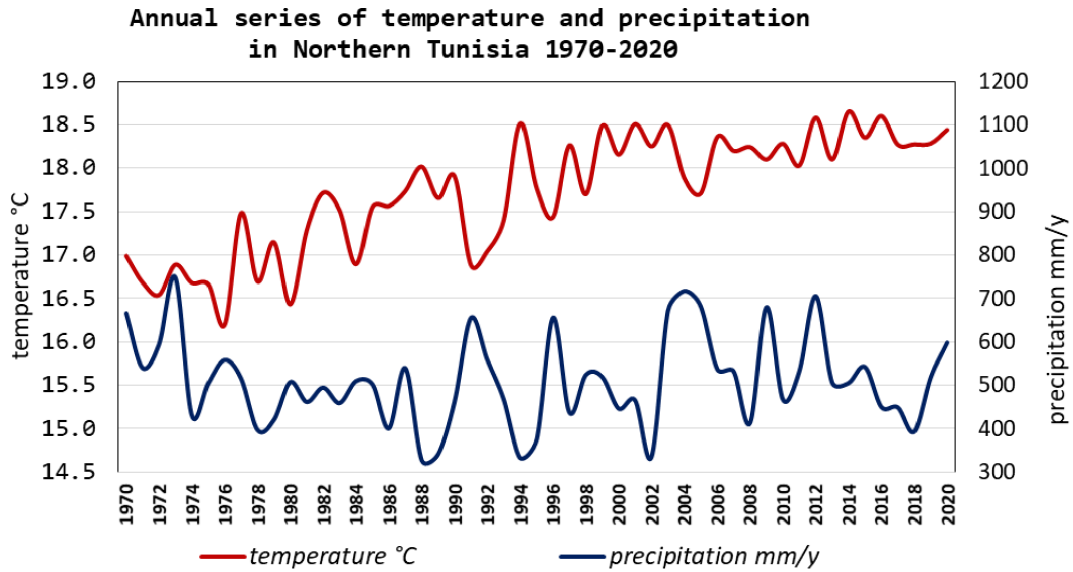
The climate change community has established a scenarios framework to facilitate integrated inter-comparisons of GCMs projections. The new framework adopted in CMIP6 combines the Representative Concentration Pathways (RCPs) defined within the CMIP5 with Shared Socioeconomic Pathways (SSPs) in a Scenario Matrix Architecture [O'Neill et al., 2016, Riahi et al., 2017].

In Figure 4, each cell of the matrix indicates a combination of socioeconomic development pathway (i.e., an SSP) and climate outcome based on a particular forcing pathway. Dark blue cells indicate scenarios that will serve as the basis for climate model projections in Tier 1 of Scenario MIP; light blue cells indicate scenarios in Tier 2 (additional scenarios). For example, the trend scenario SSP2 is compatible with the forcing scenario RCP4.5 (4.5 W/m<sup>2</sup>). Scenario SSP2-4.5 represents the medium part of the range of future forcing pathways and combines intermediate societal vulnerability with an intermediate forcing level as its land use and aerosol pathways are not extreme relative to other SSPs.

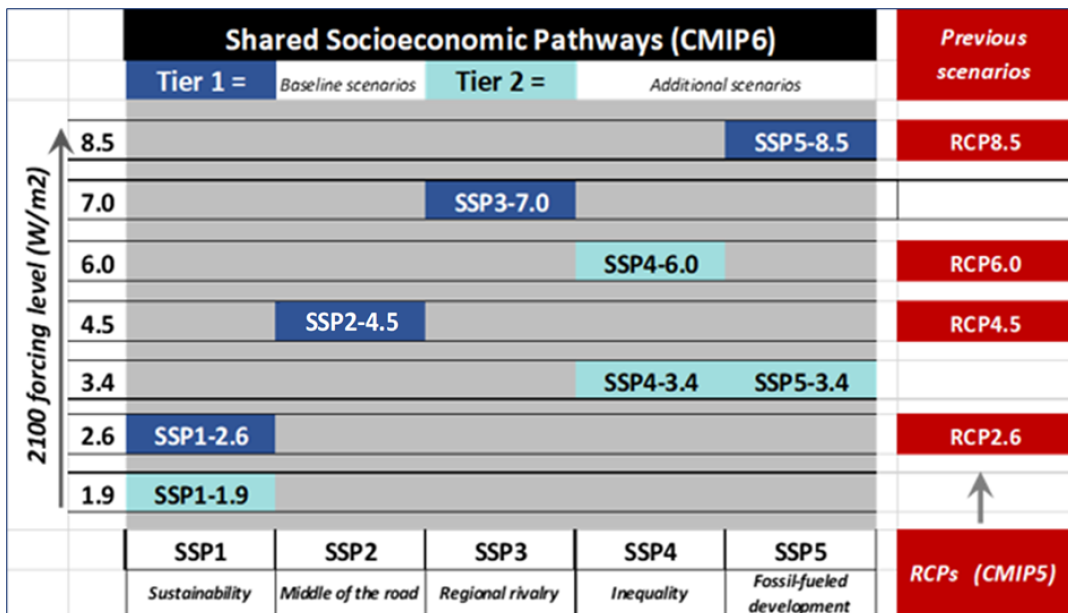
The present research aims to explore the predictability of water resources using climate outputs derived from CMIP6 ESMs' simulations related to scenario SSP2-4.5. Within this scenario, social, economic, and technological trends do not shift markedly from historical patterns, which makes it a dynamics-as-usual scenario. Projections assume a slowdown in economic growth with average growth rates in the second half of the century [Riahi et al., 2017]. We will further investigate this scenario. In doing so, we explore a reference case that should allow an evaluation of the approach and open the way to comparative explorations of various scenarios [Fricko et al., 2017].

#### 3.3. Climate change, climate models and projections

According to global projections, the temperature will continue to increase in Tunisia throughout the end of the century. A significant reduction in annual precipitation, with a trend of increase in the intensity of heavy rainfall events, is expected under high emissions scenarios. This will affect the water resources of the country, with the majority of projections indicating a progressive decrease in runoff and



**Figure 3.** Mean annual temperature (°C) and precipitation (mm/y) in Northern Tunisia (Average values of Northern Tunisia obtained by weighting the governorates values according to their areas); data sources are: DGRE, Prointec-Comete [2019] and WBG [2022].

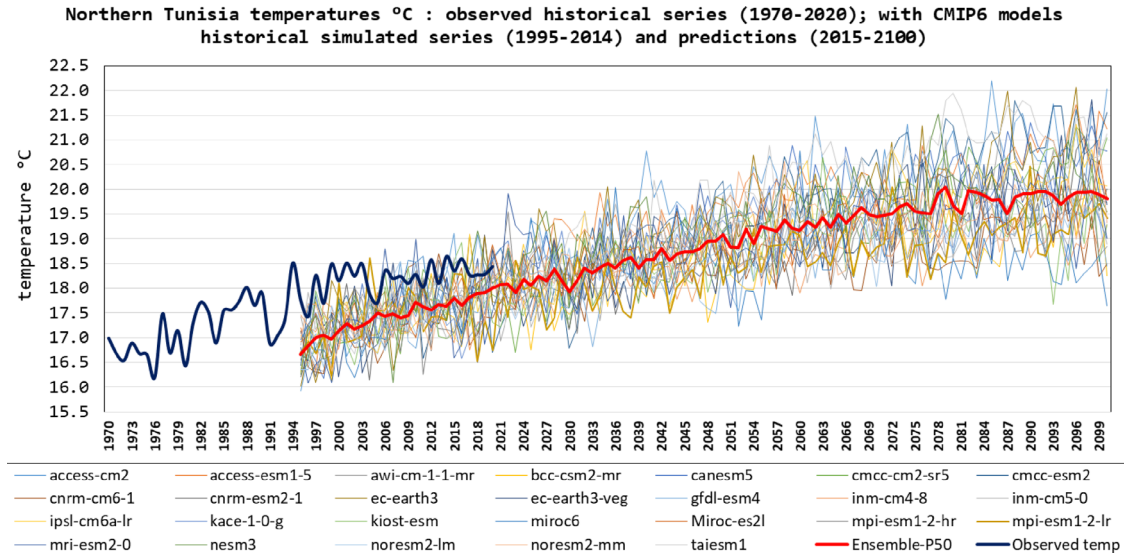


**Figure 4.** SSP-RCP scenario matrix illustrating MIP (Model Intercomparison Project) simulations [adapted from O’Neill et al., 2016].

groundwater recharge [Bargaoui et al., 2014, Döll, 2009, Oueslati et al., 2012, Nasr et al., 2008, Schewe et al., 2014, Slama et al., 2020].

The study carried out by AFD-MA [2021] analyzes the effects of RCP4.5 and RCP8.5 climate scenarios on Tunisia’s water resources and food security. Projections on the 2050 and 2100 horizons indicate a





**Figure 5.** Observed and predicted annual Temperatures on Northern Tunisia for the CMIP6 active Models, Scenario ssp2.45, period 1970–2100 [Data from WBG, 2022].

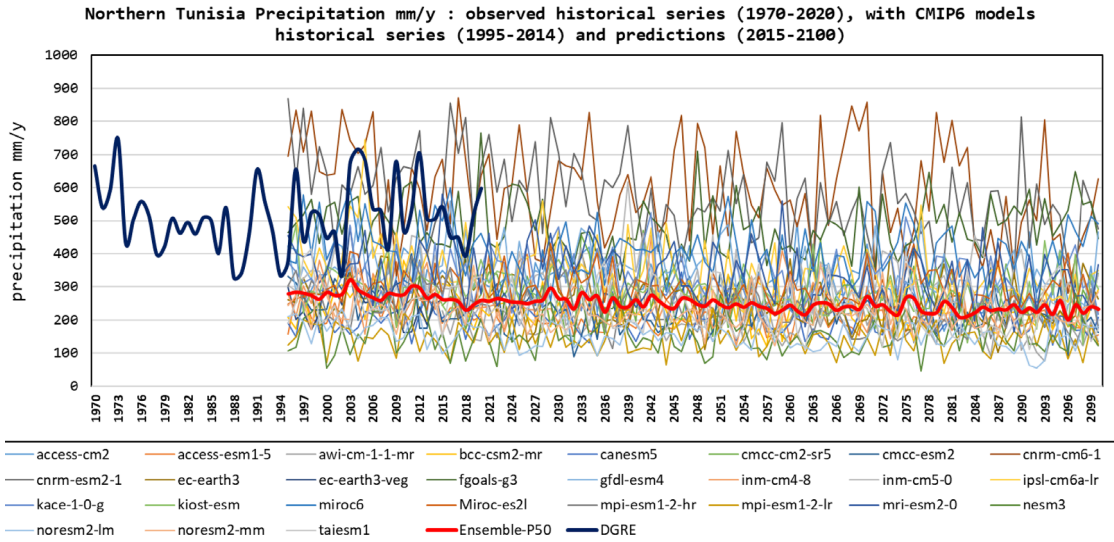
rise of the arid and semi-arid climatic stages towards the North and a quasi-disappearance of the humid stage by 2100. They indicate also a rise in winter temperatures, which affects crop development stages and yields, and increases water stress with a national cereal production decrease between 16% and 38% depending on the scenarios and horizons [AFD-MA, 2021].

The 2021 IPCC Report [IPCC, 2021] assesses results from climate models participating in the CMIP6, which include new and better representations of physical, chemical, and biological processes, as well as higher resolution. Observed warming is within the very likely range of the CMIP6 ensemble. However, some differences from observations remain, for example in regional precipitation patterns. Analysis of models' outputs includes both historical period and future projections, as well as predictability of the climate system on various temporal and spatial scales.

The models used in CCKP-CMIP6 compilation are the following: ACCESS-CM2; ACCESS-ESM1-5; AWI-CM-1-1-MR; BCC-CSM2-MR; CAMS-CSM1-0; CANESM5; CESM2; CMCC\_CM2-SR5; CMCC-ESM2; CNRM-CM6-1; CNRM-ESM2-1; EC-EARTH3; EC-EARTH3-VEG; FGOALS-G3; GFDL\_ESM4; HADGEM3 -GC31-II; INM-CM4-8; INM-CM5-0; IPSL\_CM6A\_LR;

KACE-1-0-g; KIOST-ESM; MIROC-ES2I; MIROC6; MPI\_ESM1-2-HR; MPI-ESM1-2-LR; MRI-ESM2; NESM3; NORESM2-LM; NORESM2-MM; TAIESM1; UKESM1-0-II.

Specific projections of Temperatures and Precipitations for Northern Tunisia: CMIP6 historical simulations are performed over different periods from 1850 to 2014. Figures 5 and 6 present the hydroclimatic outputs respectively temperatures and precipitation, resulting from the simulations relating to the ssp2.45 scenario. These are driven from the 26 models whose outputs on Northern Tunisia are exploitable and include retrospective simulations over the historical reference period (1995–2014) and prospective projections up to the end of the century (2015–2100). The figures also show data on the reference period (1970–2020): the observed temperatures (mostly reanalyzed ERA5 values) and the observed precipitations (from DGRE database). We can see that many models provide series with negative bias vis-a-vis historical temperature and precipitation series, which is well attested by the Ensemble distribution (aggregated prediction obtained by applying weightings to the results of the 26 GCMs) and in particular its median series (P50).



**Figure 6.** Observed and predicted annual Precipitation on Northern Tunisia for the CMIP6 active Models, Scenario ssp2.45, period 1970–2100 [Data from WBG, 2022, DGRE, Pointec-Comete, 2019].

## 4. Modeling climate change and water resources

### 4.1. A model to convert precipitation into actual evapotranspiration, runoff, and groundwater recharge

From projections of temperature ( $T$ ) and precipitation ( $P$ ) over the next decades developed by various global climate models, we propose, with a hydrological model, to quantify the hydrological impacts of Climate Changes on Northern Tunisia and predict the future of blue and green water resources at various horizons of the 21st century.

The hydrological model is a production function type with two reservoirs: a balance reservoir and a transfer one. System inputs are precipitation and potential evapotranspiration; outputs are actual evapotranspiration, infiltration, and runoff. In semi-arid regions such as Tunisia, this type of model is operational, tested, and validated with a daily time step, but given the availability of climatic data series and the lengths of periods to simulate, we had to work on a monthly basis. On the water budget, such a time step extension results in errors that are all the greater as the climate aridity is higher [Besbes, 1978]. This constraint led us to restrict our study area to the limits of Northern Tunisia climatic region. Executed in sequence, the set of equations governing the water

cycle in the model reservoirs (Figure 7) are, all quantities expressed in monthly water heights:

Water balance reservoir:

- (i)  $AET = \text{Min}(R_{t-1} + P; PET)$ ;
- (ii)  $R = \text{Min}((R_{t-1} + P - AET); R_{\text{max}})$

Transfer reservoir:

- (iii)  $F = \text{Max}((R_{t-1} + P - AET - R_{\text{max}}); \theta)$ ;
- (iv)  $I = \text{Min}(F, I_{\text{max}})$ ;
- (v)  $RU = F - I$

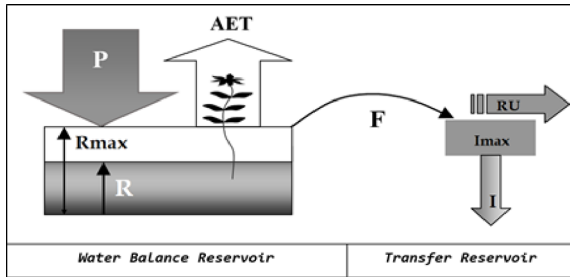
where: AET is the Actual Evapo-Transpiration;  $t$  the time, here the month;  $P$  the precipitation; PET the Potential Evapo-Transpiration;  $R$  the soil reserve;  $R_{\text{max}}$  the maximum soil reserve capacity;  $F$  the total Flow;  $I$  the infiltration (effective infiltration, or groundwater recharge);  $I_{\text{max}}$  the maximum infiltration capacity; RU the runoff. The two parameters  $R_{\text{max}}$  and  $I_{\text{max}}$  need to be calibrated using the observed data.

The Potential evapotranspiration is given, for temperate and Mediterranean regions, by the monthly TURC formula [Alexandris et al., 2008, Bonnet et al., 1970, McMahon et al., 2013]:

$$PET = 0.4 \times (T / (T + 15)) \times (IG + 50).$$

$T$  is the temperature in Celsius degrees.

IG is the global solar radiation in  $\text{cal}/\text{cm}^2/\text{day}$  (over the study region and on annual average, IG varies from  $350 \text{ cal}/\text{cm}^2/\text{day}$  in Bizerte to



**Figure 7.** Schematic diagram of the hydrological model.

410 cal/cm<sup>2</sup>/day in Siliana, with a maximum in July and a minimum in January [ANME (National Agency for Energy Management), 2023]).

The study area is limited to the 11 administrative districts in the northernmost region of the country. For each of these districts, the CCKP [WBG, 2022] provides monthly historical series of temperature and precipitation, which are transformed into the corresponding series of infiltration (groundwater recharge), runoff, and AET using the monthly hydrological model. The monthly outputs of the model are first integrated and converted into annual output series. The model is run separately for each of the governorates considered as lumped isolated watersheds. The general assessment is then carried out on the 11 governorates: individual annual series are integrated on space for the whole districts to build a yearly model over the entire study area. The simulated series are finally compared to observed hydrometric data.

The historical runoff data series are produced and edited by DGRE (Tunisian Ministry of Agriculture) in the Hydrological Yearbooks, which have been published regularly since the early 1980s. [DGRE, 2020]. The Yearbooks, which cover the three major hydrological basins making up the study area (Far North-Ichkeul, Medjerda, Cap Bon-Miliane, Figure 1), report on the measurements taken at the main hydrometric stations with variable frequencies but are most often formatted into daily values, then aggregated into monthly and annual values. These values do not consider the flow exchanges between Algeria and Northern Tunisia, estimated today at 100 Mm<sup>3</sup>/year on average in favor of Tunisia [BPEH, 2019], nor the weak diffuse coastal runoff not monitored with hydrometric stations. In the

region's water budget, these two quantities work in opposite directions. Compared to the mean basin runoff (2.3 km<sup>3</sup>/y), we consider as a first hypothesis that these two quantities (# 4%) are confined to error margins.

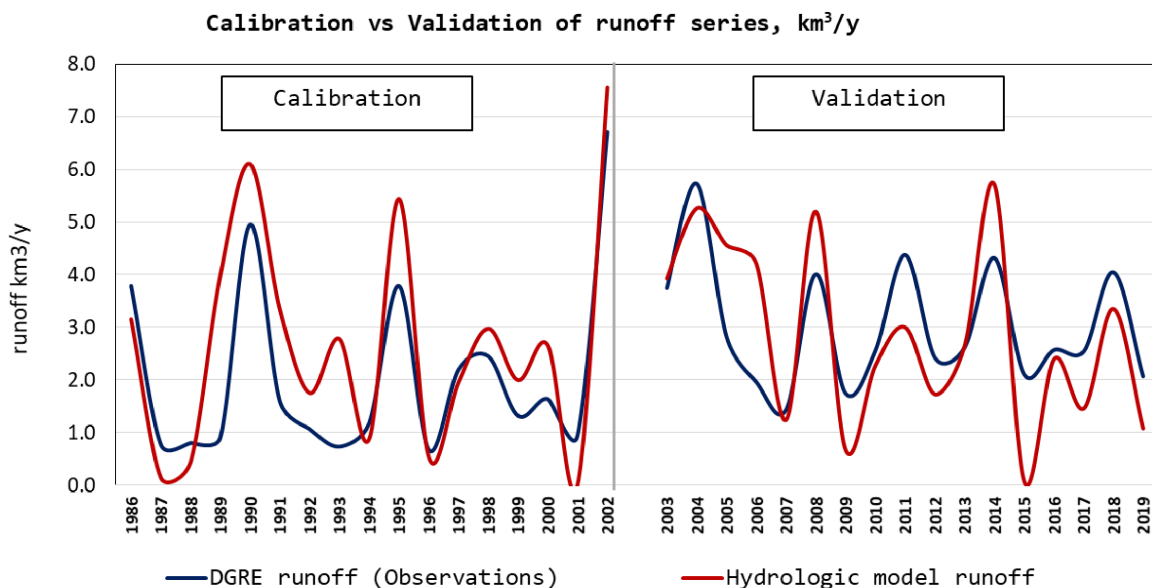
#### 4.2. The hydrological model calibration process

The first step of model calibration consists to define a calibration period for which we search the optimal value of the following two parameters:  $R_{max}$  the maximum soil reserve capacity, and  $I_{max}$  the maximum infiltration capacity. Recall that for each governorate these parameters represent spatially averaged characteristics of the whole district and are unable to be directly compared with field measurements. The second step is validation; it consists to define a validation period for which we run the model with the already calibrated parameters  $R_{max}$  and  $I_{max}$ . The calibration period extends over 17 hydrological years, from 01/09/1985 to 31/08/2002; the validation extends over the 17 following years from 01/09/2002 to 31/08/2019 (Figure 8).

The calibration process aims to reproduce as well as possible the observed annual runoff volumes [DGRE, 2020]. As for the second major term of the hydrological resources, annual groundwater recharge, there are no direct observations or precise measurements. We must therefore refer to the most recent published field estimates [Besbes et al., 2019b].

Calibrating the model parameters:

- (i) The maximum soil reserve capacity  $R_{max}$ : the first calibration tests showed that the monthly time step operation is incompatible with the potential evapotranspiration displayed in the study region, yet one of the wettest in Tunisia. Any significant Soil Reserve Value will automatically feed into the month's AET, leaving virtually nothing for runoff and groundwater recharge. It was only by setting  $R_{max}$  to zero that the model was able to function normally, calculating quite plausible water budgets. Under similar climatic conditions, this type of model works very well daily, the monthly model is supposed to integrate the daily balances of the month. We succeeded in obtaining plausible results directly at the monthly model by introducing the artifice of  $R_{max} = 0$ .



**Figure 8.** Modeled and observed yearly runoff for calibration and validation series ( $\text{km}^3/\text{y}$ ).

- (ii) The maximum infiltration capacity  $I_{\max}$ : the dominant value is 10 mm/month, except in the governorates of Beja and Jendouba where  $I_{\max} = 8$  mm.

We assess the model calibration quality, as well as validation, by estimating: (i) the average relative bias between computed and observed runoff of the series, (ii) the correlation coefficient between modeled and observed annual runoffs; (iii) Furthermore and more specifically, our model estimates also the groundwater recharge: we then calculate the average deviation between computed and field estimated recharge. Finally, we draw a scatter plot of modeled vs. observed yearly runoff, respectively for calibration and validation series.

The hydrologic model calibration is based on a double aggregation: (a) a spatial aggregation of the outputs obtained for the eleven governorates of the study region, (b) a temporal aggregation of the monthly outputs (runoff, groundwater recharge) to build yearly series that can be compared to observed or estimated data. The results are as follows:

- (i) The average relative bias between computed and observed annual runoff for calibration and validation yearly series establishes at respectively 22% and -5%.

- (ii) The correlation coefficients between modeled and observed annual runoffs are respectively 0.86 and 0.74.
- (iii) The average computed annual recharge for the whole series 1985–2019 is  $0.68 \text{ km}^3/\text{y}$ . The average field estimated annual recharge is  $0.67 \text{ km}^3/\text{y}$  [Besbes et al., 2019b]. The average deviation between computed and field-estimated recharge is  $0.01 \text{ km}^3/\text{y}$  that is a relative mean deviation of 1.5%.

#### 4.3. Selection of CMIP6 GCMs compatible with historical reference series

We have seen, Figures 5 and 6, that considering the historical reference period 1995–2014, most models exhibit a negative bias vis-à-vis historical temperature and precipitation series. We must therefore proceed to the choice of the models which are the most suitable to correctly reproduce climatic and hydrologic series for the period 1995–2014, considered as the historical reference period and for which there are simulations by all of the CMIP Models.

The selection process of the most relevant GCMs for predictive simulations follows two steps:

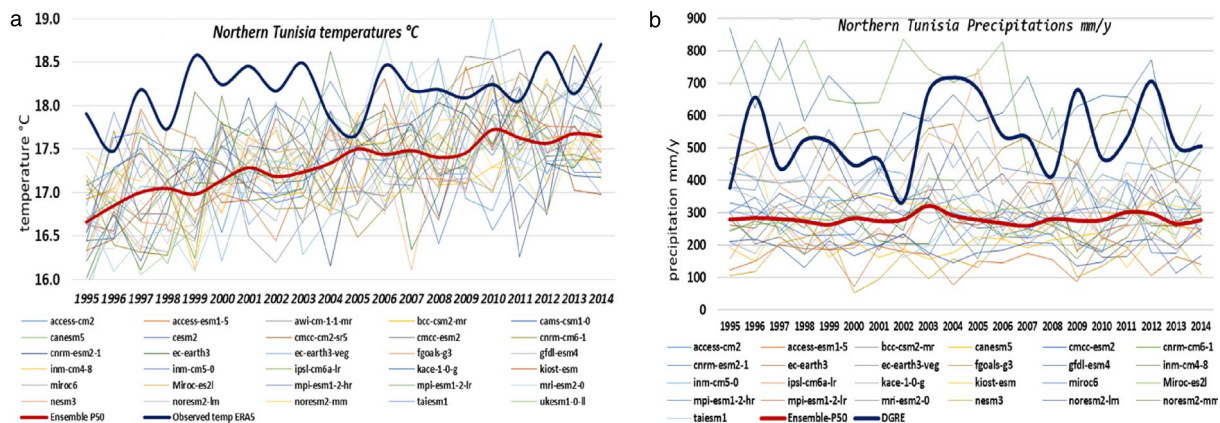
- (1) The First level of selection criteria concerns the predictability of the historical

**Table 2.** Statistical parameters of the precipitation series generated by active GCMs on 1995–2014, compared to the observed reference DGRE series

	Mean mm/y	Standard deviation mm/y	Mean bias/DGRE mm/y	Mean relative bias
DGRE observed series	<b>554.1</b>	<b>115.4</b>		
access-cm2	334.9	69.7	-219.2	-40%
access-esm1-5	251.7	58.7	-302.4	-55%
bcc-csm2-mr	259.9	65.9	-294.2	-53%
canesm5	206.4	44.2	-347.6	-63%
cmcc-esm2	277.1	57.9	-277.0	-50%
cnrm-cm6-1	<b>652.2</b>	<b>141.6</b>	<b>98.1</b>	<b>18%</b>
cnrm-esm2-1	<b>630.0</b>	<b>121.5</b>	<b>75.9</b>	<b>14%</b>
ec-earth3	277.7	62.7	-276.4	-50%
ec-earth3-veg	256.0	74.4	-298.1	-54%
fgoals-g3	<b>504.6</b>	<b>62.6</b>	<b>-49.5</b>	<b>-9%</b>
gfdl-esm4	364.9	72.8	-189.2	-34%
inm-cm4-8	269.3	62.9	-284.7	-51%
inm-cm5-0	339.8	66.6	-214.3	-39%
ipsl-cm6a-lr	<b>421.1</b>	<b>111.6</b>	<b>-133.0</b>	<b>-24%</b>
kace-1-0-g	324.8	71.5	-229.3	-41%
kiost-esm	318.4	69.2	-235.7	-43%
miroc6	<b>451.0</b>	<b>87.1</b>	<b>-103.1</b>	<b>-19%</b>
miroc-es2l	297.8	50.2	-256.3	-46%
mpi-esm1-2-hr	272.7	66.1	-281.4	-51%
mpi-esm1-2-lr	154.3	43.0	-399.8	-72%
mri-esm2-0	359.3	69.8	-194.8	-35%
nesm3	170.5	72.1	-383.6	-69%
noresm2-lm	182.8	34.5	-371.3	-67%
noresm2-mm	253.5	50.9	-300.6	-54%
taiesm1	252.1	62.7	-302.0	-55%
Ensemble-P50	280.4	14.3	-273.7	-49%

temperature and precipitation series, which is evaluated by the degree of conformity of the chronological graphs and by the quality of the statistical parameters of the series. This selection was limited to precipitation (resp. Figures 9a and 9b) with criteria based on the statistical series parameters. Indeed and concerning the temperature, vis-à-vis observed values, the 30 models present series with almost identical biases, and relative

biases of 5% with a standard deviation of 0.06%, which means that there is no significant difference between the series resulting from the models, which would have allowed a selection. As for precipitation (Figure 9b and Table 2), few series come close to the reference one (DGRE). According to criteria that minimize differences with DGRE, it thus appears that five GCMs give precipitation close to the observations, in particular with very



**Figure 9.** Historical reference and predicted series (1995 to 2014) for annual temperature (a) and precipitation (b).

**Table 3.** Statistical parameters of runoff and groundwater recharge series generated by active GCMs on 1995–2014, compared to observed reference DGRE series

		Data DGRE	cnrm-cm6-1	cnrm-esm2-1	fgoals-g3	miroc6	ipsl-cm6a-lr
Runoff	Mean km <sup>3</sup> /y	2.87	3.93	3.64	1.99	1.49	1.30
	Std deviation km <sup>3</sup> /y	1.59	2.55	2.17	1.23	1.45	1.49
	Mean Bias vis-à-vis DGRE		1.06	0.76	-0.88	-1.38	-1.58
	Mean Relative Bias		37%	27%	-31%	-48%	-55%
Recharge	Mean km <sup>3</sup> /y	0.67	0.93	0.91	0.65	0.55	0.46
	Std deviation km <sup>3</sup> /y	0.00	0.32	0.35	0.21	0.28	0.26
	Mean Bias vis-à-vis DGRE		0.26	0.24	-0.02	-0.12	-0.21
	Mean Relative Bias		39%	36%	-3%	-18%	-31%

low average relative biases (Table 2); these are the models cnrm-cm6-1, cnrm-esm2-1, fgoals-g3, ipsl-cm6a-lr, miroc6.

(2) The second level of selection criteria concerns the outputs of the hydrological model and their ability to reproduce the reference historical surface runoff and groundwater recharge. This reproducibility is assessed by the conformity of computed chronological series toward the statistical parameters of the reference series. Regarding groundwater recharge, the average field estimated value is 0.67 km<sup>3</sup>/y [Besbes et al., 2019b]. As for runoff, each of the CMIP6 GCMs generates, over the reference period 1995–2014, rainfall series used as input to the hydrological model to produce the corresponding runoff series, which is compared to the

series observed by DGRE [2020]. The hydrological results (Table 3) confirm in the first analysis those obtained on precipitation (Table 2), namely the five models: cnrm-cm6-1, cnrm-esm2-1, fgoals-g3, ipsl-cm6a-lr, miroc6, which give the best results.

But more precisely, and more particularly with regard to runoff for which there are series of reliable observations, the statistical parameters of Table 3 clearly indicate that the outputs of three models among the five selected still remain far from the reference data, either too optimistic or too pessimistic, and would therefore not guarantee sufficient reliability on the predictions. These are the models: cnrm cm6-1, ipsl-cm6a-lr, and miroc6. On the other hand, the cnrm-esm2-1 and fgoals-g3 models give results very close to the historical reference data, particu-

larly for the runoff that they globally frame, with a respective bias of +0.76 and  $-0.88 \text{ km}^3/\text{y}$ . We do not find this symmetry at the groundwater recharge level, where fgoals-g3 reconstitutes the expected result on its own. However, given that we do not have direct measurements of the groundwater recharge but only an estimate from expertise, we will give greater credit to the models that best predict runoff.

## 5. Predicting blue and green water according to the scenario ssp2-45

### 5.1. Construction of a global climate model specific to Northern Tunisia

Among the CMIP6 models, the cnrm-esm2-1 and the fgoals-g3 are therefore those that best restore the series of historical runoff over Northern Tunisia; but this restitution is imperfect: the cnrm slightly overestimates the sought solution and the fgoals slightly underestimates it. We will consider that a suitable model would be a composition of the two models over the computational grid, a composite model, with a weighting  $\alpha$  such that the composite precipitation Prc is written:  $\text{Prc} = \alpha \times \text{cnrm} + (1 - \alpha) \times \text{fgoals}$ . The same weighting is applied to temperature. To find the weighting that minimizes the difference with observations, we vary  $\alpha$  between 0 and 1 on the rainfall series of the two models (cnrm and fgoals) for the reference historical period 1995–2014. We obtain the results presented in Table 4, where we see that there is no solution that simultaneously minimizes both the deviations in runoff and in recharge, but we can consider that the best compromise would be reached with  $\alpha = 0.5$ . In what follows, we will call this model the “cnrm-fgoal” model.

### 5.2. Projection of blue water resources with the cnrm-fgoal model

The hydrological behavior of Northern Tunisia is simulated for the next successive situations: (i) the known period 1995–2014 considered as the historical reference period with observed precipitation and runoff series, and estimated groundwater recharge, (ii) the forecasting projection perspective covering the period 2015–2100. The three Global Climate Models: cnrm-esm2-1, fgoals-g3 and their composite “cnrm-fgoal”, are used to first simulate the hydrologic history of the basin for the reference period, and

then to predict the hydrological impacts related to the ssp2-45 scenario. Figures 10 and 11 report the obtained results. The explanations and comments on these figures will be included in the summary Table 5. However, concerning both runoff and groundwater recharge, we can already observe that, after a particularly wet period that extends until 2030, the region will experience a series of dry cycles that will last until the end of the 21st century.

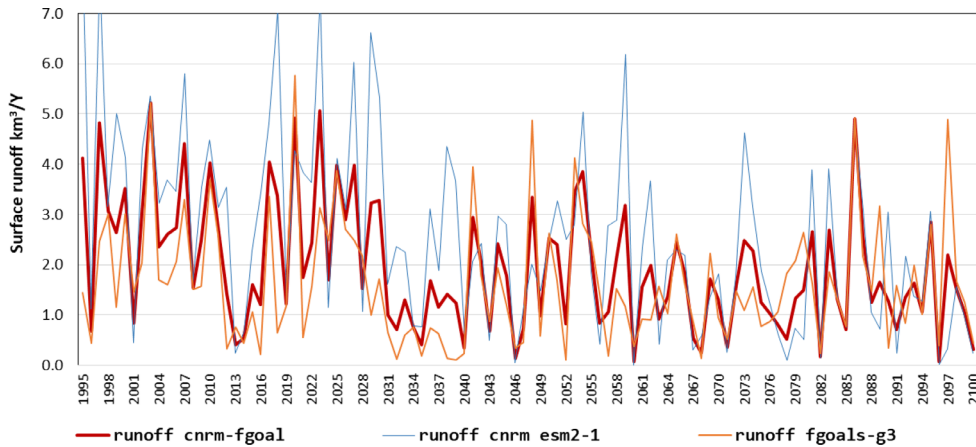
### 5.3. Projection of green water resources

#### 5.3.1. The green water model: formulation and validation

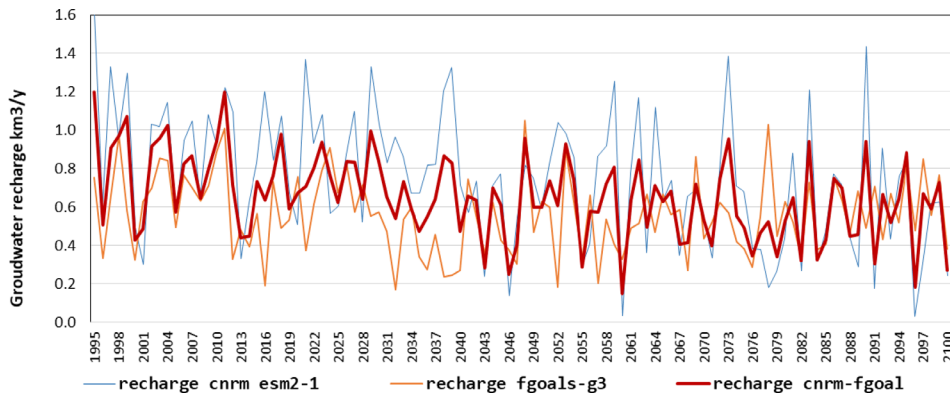
The average Actual Evapotranspiration (AET) carried out by the hydrological model described above represents the ultimate potential of Green Water (GW) resources that can be extracted by plants from the soil. Only part of this potential is productive and can be converted into food production. Indeed: (i) Only part of arable land available for rain-fed crops is cultivated, (ii) Evapotranspiration does not occur throughout the year, especially for annual crops such as cereals whose active period is limited in time, (iii) Cultivated area is not always entirely harvested due to possible crop failure during low rainfall episodes. The Green Water Model (GWM) is formulated on the basis of hydrological model outputs, which provide AET. To determine the productive part of GW resource, we define the rate of conversion  $\beta$  of the annual AET, which represents the maximum green water potential of the whole region under study (GWP), into annual GW actually used by the crops, so that we have  $\text{GW} = \beta \cdot \text{GWP}$ , where  $\text{GWP} = A \cdot \text{AET}$ ,  $A$  is the yearly cultivated area. GW here is assumed to be the green water useful volume, which is proportional to agricultural production.

As for the hydrological model, the GWM is formulated for each governorate and adjusted for cereals cultivated area ( $8640 \text{ km}^2$  on average) based on sub-national data from the Ministry of Agriculture [MARH, 1998–2012] and national data from FAO-STAT [FAO, 2017a,b]. These data sources provide regional statistics on agricultural production from 1985 to 2020. The model outputs by governorate are integrated over the entire study region and adjusted on observations.

Figure 12 shows the relationship between the annual production of cereals (CPr) in Million Tons



**Figure 10.** ssp2-45 Runoff simulation with cnrm-esm2-1, fgoals-g3, and the selected cnrm-fgoal model for historical observation period and projections.



**Figure 11.** ssp2-45 Groundwater recharge simulation with cnrm-esm2-1, fgoals-g3, and the selected cnrm-fgoal model for historical observation period and projections.

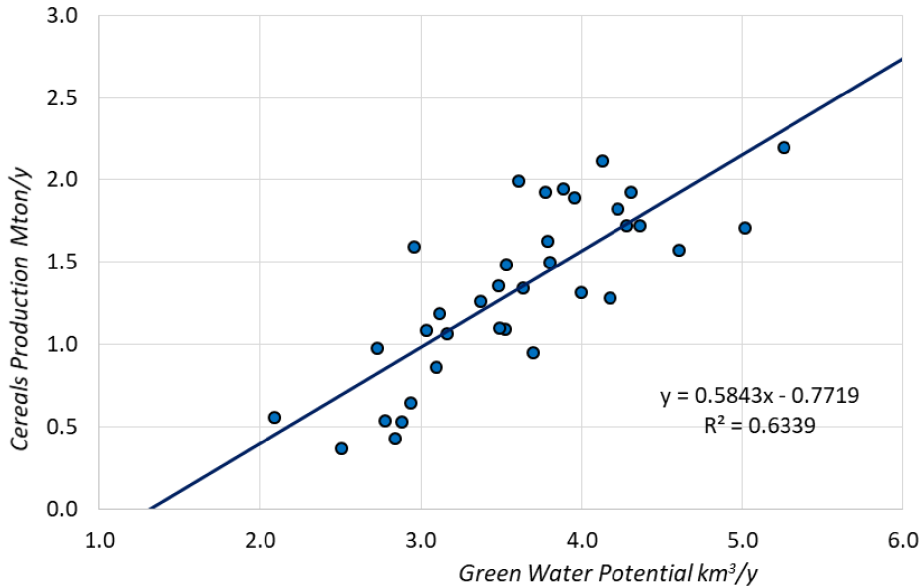
**Table 4.** The weighting effect on runoff and groundwater recharge results, from 100% cnrm-esm2-1 ( $\alpha = 1$ ) to 100% fgoals-3g ( $\alpha = 0$ )

Alpha	1	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0
Bias on runoff km <sup>3</sup> /y	0.76	0.55	0.34	0.14	-0.04	-0.22	-0.37	-0.52	-0.65	-0.77	-0.88
Bias on recharge km <sup>3</sup> /y	0.24	0.23	0.20	0.18	0.16	0.13	0.10	0.10	0.04	0.01	-0.02

according to cereals’ Green Water Potential (GWP) assimilated to the total volume of AET. The slope of the least squares regression line provides the specific increase in production per unit of GWP, which amounts to 0.58 kg/m<sup>3</sup> i.e. a specific volume of GW of 1.71 m<sup>3</sup>/kg. This value compares quite well with the water productivity of cereals in arid countries [Zimmer, 2013].

The specific volume of GW is then used to calculate the equivalent GW of annual cereal production and deduce the corresponding beta conversion rates. These are fitted using linear regression over the period 1985–2019 and presented in Figure 13, where the regression line shows a significant increase in the conversion rates of GWP into useful GW, from 0.5 to 0.7, reflecting crop yield improvement. The conver-





**Figure 12.** Cereal production in Northern Tunisia related to Green Water Potential AET (period 1985–2019).

sion function beta is used to generate retrospective GW amounts (Figure 13) compared to observed cereals production. Figure 13 shows that the GW volumes produced by the model manage to capture the relationship between climatic conditions and cereal production, mainly grown in rain fed conditions. During this historical reference period, GW follows the evolution of the beta factor, indicating significant yield evolution despite the stability of cereal surfaces, [Attiaoui and Boufateh, 2019]. Cereals Yields improvement is not specific to the study area; it is found throughout Tunisia as well as in other Maghreb countries (Algeria and Morocco). Voluntary agricultural policies, as well as a number of incentives and favorable farming practices explain that in the case of wheat, for example, yields have doubled in twenty years [Besbes et al., 2019a].

### 5.3.2. Simulation of the green water future with the selected models:

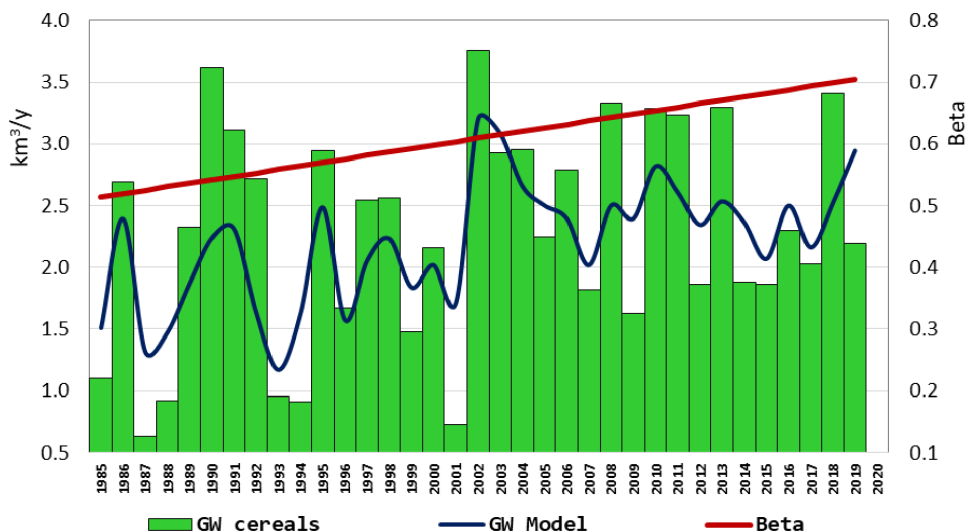
To make the predictive results comparable with the current situation, prospective simulations covering the period 2020–2100 were carried out with initial conditions, those prevailing in the 2010s, including crop yield and cultivated areas. For instance, the average maximum value of beta over 2011–2020 (0.68) will be kept constant throughout the prospective

study period. Thus, this exercise makes it possible to examine the impacts of regional parameters predicted by Global Climatic Models on green water resources. For both blue water and green water, all calculations are carried out at the level of each of the eleven governorates, and then aggregated at the level of the entire region (Figure 14 and Table 5). Calculus uses outputs of the two GCMs cnrm-esm21 and fgoals-g3, as well as their composite cnrm-fgoal previously adjusted as part of the hydrological model.

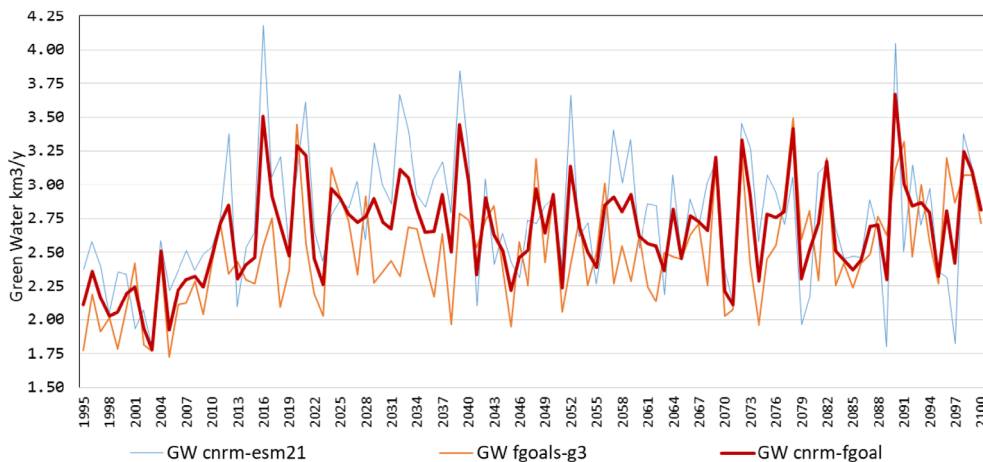
### 5.4. Results of water resources prediction and summary table

Table 5 is a synthetic composition intended to guide the reader. It summarizes precipitation, runoff, recharge, and green water calculated from the composite model “cnrm-fgoal”, coupled to the hydrological and the green water models, implemented over the observation and the projections periods. Here are the main points and findings:

- (i) The decade 2011–2020 represents the historical reference state of the hydro-climatic system; it is the initial state of the projections, calculated with the model by reference to the observed state and to which any variation is related.



**Figure 13.** Checking the Green Water Model over the period 1985–2019, by comparison with cereals’ water footprint (km<sup>3</sup>/y).



**Figure 14.** ssp2-45 useful Green Water simulation with cnrm-esm2-1, fgoals-g3, and cnrm-fgoal for historical observation period (1995–2014) and projections (2015–2100).

- (ii) We have chosen to present the general results of the study in successive sequences of 20 years, i.e. four sequences to represent the period simulated in the projections.
- (iii) A first reading of the table highlights the great fragility of blue water in relation to changes in precipitation and temperature, which is not the case for less sensitive green water. In a first analysis, this difference is probably due to the priority given to AET when modeling the water budget.

This being so, we can note the following main results expressed in Table 5:

- (iv) Precipitation: starting from 16.33 km<sup>3</sup>/y in the 2010s, precipitation decreases to a minimum of around 14 km<sup>3</sup>/y in the 2050s and then slightly rises to 14.7 km<sup>3</sup>/year at the end of the century; this represents an average decline of 9% over the 21st Century.
- (v) Temperatures: during the century, the average annual temperature increased by nearly 2 °C, ranging from 17.56 °C to 19.44 °C.

**Table 5.** Summary table in Northern Tunisia for overall modeling with the GCM “cnrm-fgoal”, the hydrological model and the green water model

Item/period	2011–2020	2021–2040	2041–2060	2061–2080	2080–2100	Summary 2021–2100
Average observed precipitation (Pr) km <sup>3</sup> /y	15.03					
Average precipitation cnrm-fgoal km <sup>3</sup> /y	16.23	15.56	14.53	14.07	14.69	14.71
Difference with initial Pr %	0%	–4%	–11%	–13%	–10%	–9%
Average reference temperature ERA5 °C/y	18.30					
Average temperature by cnrm-fgoal °C/y	17.56	17.95	18.57	19.00	19.44	18.74
Difference with initial temperature %	0%	2%	6%	8%	11%	7%
Average observed runoff km <sup>3</sup> /y	3.08					
Average runoff hydrologic model km <sup>3</sup> /y	2.15	1.99	1.88	1.35	1.60	1.70
Difference with initial runoff %	0%	–7%	–13%	–37%	–26%	–21%
Average estimated recharge km <sup>3</sup> /y	0.67					
Average recharge hydrologic model km <sup>3</sup> /y	0.72	0.72	0.59	0.57	0.57	0.61
Difference with initial recharge %	0%	0%	–18%	–21%	–21%	–15%
Average green water harvested km <sup>3</sup> /y	2.58					
Average green water model km <sup>3</sup> /y	2.67	2.60	2.44	2.45	2.53	2.50
Difference with initial GW %	0%	–3%	–9%	–8%	–5%	–6%
Total actual evapotranspiration AET km <sup>3</sup> /y	13.37	12.85	12.06	12.15	12.52	12.39
Difference with initial AET %	0%	–4%	–10%	–9%	–6%	–7%

- (vi) With regard to runoff, it decreases on the composite model cnrm-fgoal, from 2.15 km<sup>3</sup>/y on average in the 2010s to 1.35 in the 2070s, i.e. a drop of 37%, then increases to form a century average value of 1.70 km<sup>3</sup>/y. i.e. an average drop of 21%.
- (vii) During the century, aquifers recharge went from 0.72 to 0.57 km<sup>3</sup>/y, i.e. a drop of 21%.
- (viii) With regard to Green Water, the predictive simulations for 2100 show evolutions slightly below the initial level, with a decrease of up to –6% on average.

## 6. Discussion

Finally, the following points should be pointed out in the form of a discussion:

- (1) *On the green water model:* over the projected period 2021–2100, the temperature increases steadily. This warming results in higher water requirements for plants: (i) higher evapotranspiration demand, (ii) longer and warmer growth periods. The first factor is already effective in terms of the results obtained; indeed, the decrease in GW is small

compared to that of precipitation. The second factor is not taken into account by our green water model; one can imagine that such an additional withdrawal of green water would lead to even greater decreases in runoff and recharge [Mankin et al., 2019].

- (2) *On the development of scenario ssp2.45:* at the end of the 2060s (Table 5), we observe an important reduction in runoff followed by a rise. Similar results are observed in the simulations obtained from the outputs of other GCMs we analyzed, i.e. cnrm cm6-1 and miroc6. We previously attributed this phenomenon to the priority given to evapotranspiration by the hydrological model. This is partly true, but there is another reason which produces a perhaps even greater effect. Indeed: the SSP2 narrative assumes moderate global population growth with a demographic transition in the second half of the century. Population stabilization combined with sustained income growth would be accompanied by an improvement in agricultural productivity as well as in carbon and energy intensity. On the other hand, the

increasing rigor of climate change mitigation is expected to decrease dependence on fossil fuels, in particular by optimizing the energy mix [Fricko et al., 2017]. All these factors are likely to reduce the effects of human activities that climate models seem to capture.

- (3) *On other studies dealing with long-term water resources prediction in Northern Tunisia:*
- (i) Precipitation forecasts with the “cnrm-foal” model indicate a slight change over the first half of the century, then a huge decrease with a maximum drop around the sixties (−13%). Then follows a slow recovery, bringing the average variation 2021–2100 to (−9%). These values are within orders of magnitude of previous predictions with RCM for Northern Tunisia. Under scenario RCP 4.5: Deidda et al. [2013] and Bird et al. [2016] predict a reduction of around (−17%) for 2040–2070. For the same period, Dakhlaoui et al. [2022] find an average value of (−8%) in the region, increasing to (−18%) for 2070–2100. Terink et al. [2013] announce a (−5% to −10%) for 2040–2050 and the AFD-MA study [AFD-MA, 2021] predicts, for all of Tunisia, a precipitation drop of (−6% to −9%) by 2050 and (−9% to −18%) by 2100.
  - (ii) Concerning runoff and under scenario RCP 4.5, Dakhlaoui et al. [2022] predict runoff reductions of (−10% to −30%) for 2040–2070 and −20% to −38% for the period 2070–2100, figures which are close to those of our present simulations, which expect reductions around −37% by 2070 and −26% by 2090.
- (4) *Consequences of green water predictions for cereal production:* In line with the results of most studies in the Maghreb countries and the Mediterranean basin, rainfall is considered the crucial factor in cereals production [Zimmer, 2013]. This suggests that a precipitation decline would have harmful repercussions on cereals production. According to the policy-free scenario (RCP8.5), the AFD-MA [2021] study predicts a decline of more

than a third of National agricultural production by 2100. Based on statistical approaches, Attiaoui and Boufateh [2019] estimate that nationally, a 1% decrease in rainfall would result in (0.92%) decrease in cereals production in the short run and (1.295%) in the long run. These would be less marked in the more watered cereal-growing regions of the north. The authors further suggest that the increase in temperature would even have a positive effect on cereals production in these colder regions of Tunisia. Our green water model provides comparable orders of magnitude of the impact of climate change on cereals production. The predictive simulations for 2100 show an average decrease of around 6% in green water (e.g., cereals production) for an increase in temperature of 7% and a decrease in precipitation of 9%. It is as if the effects of the increase in temperature and the decrease in precipitation compensate for each other to produce smaller changes in the predicted AETs and thus in crops productions, albeit the conversion rates and cultivated areas, which are considered constant.

In sum, the results of the simulations predict an increase in temperature accompanied by a long-term decrease in precipitation. The implications of these predictions on Green Water are less significant than on Blue Water. This is an essential finding of this research that reinforces the previous analyzes showing the weight and the paramount role of green water in the national water budget [Besbes et al., 2019b]. These findings, their deepening, monitoring and possibly updating constitute essential information to guide the development of adaptation strategies compatible with the future potential of water resources in all their forms and for each region.

## 7. Conclusion

The simulations carried out by the GCMs that participated in the CMIP6 exercise provide temperature and precipitation series for the study area. We considered in this study the IPCC medium scenario SSP2-4-5, designed to prolong current trends. Analysis of the

thirty GCMs used in the research led to the selection of only two GCMs (cnrm-esm2.1, fgoals-g3) whose outputs: (i) are the closest to the observed historical temperature and precipitation series but imperfectly, (ii) applied as inputs to the hydrological model, generate outputs that frame the observed runoff series. The modeling systems' runoff outputs are weighted to fit observations and build a composite best-fit model: the "cnrm-fgoal" weighted at 0.5–0.5 for the prospective hydrological simulations until 2100.

Inputs and outputs of the cnrm-fgoal composite model indicate slight alterations of water resources until the forties, then a huge alteration from the middle of the 21st century and a maximum drop in the sixties, with a significant drop in precipitation (–13%), runoff (–37%), groundwater recharge (–21%) and Green Water (–8%). Subsequently, we note a recovery in all indicators, so that the average variation over most of the 21st century, from 2021 to 2100, will be a decrease of –9% for precipitation, –21% for runoff, –15% for recharge, and –6% for Green Water.

To our initial question, stated as follows: can we predict Water Resources only with GCMs, without downscaling, the results obtained in the present study show that the use of the raw predictions of original climate models for hydrological modeling purposes on large basins is possible. Simulations at these scales appear to provide relative precisions comparable to what has been produced in Northern Tunisia by Regional Models applied to small/moderate basins after inputs bias corrections.

The direct implementation of climate predictions for hydrological studies is fundamentally based on a rigorous selection of suitable climate models for the region, carried out based on an exhaustive analysis of the historical simulations of the climatic models and their comparison with observed data on the same periods. This requires that reliable and complete hydrological data are available on spatiotemporal scales compatible with those of climate simulations.

### Conflicts of interest

Authors have no conflict of interest to declare.

### References

AFD-MA (2021). French Development Agency-Ministry of Agriculture. Tunisia. Contribution to

the elements of the preparatory phase of the National Adaptation Plan for food security. Summary. Dec. 2021.

- Alexandris, S., Stricevic, R., and Petkovic, S. (2008). Comparative analysis of reference evapotranspiration from the surface of rainfed grass in central Serbia, calculated by six empirical methods against the Penman–Monteith formula. *Eur. Water*, 21/22, 17–28.
- Allan, J. A. (1998). Moving water to satisfy uneven global needs: trading water as an alternative to engineering it. *ICID J.*, 47(2), 1–8.
- ANME (National Agency for Energy Management) (2023). Solar Photovoltaic. Solaire Photovoltaïque |ANME. accessed on 21/03/2023.
- Attiaoui, I. and Boufateh, T. (2019). Impacts of climate change on cereal farming in Tunisia: a panel ARDL–PMG approach. *Environ. Sci. Pollut. Res.*, 26, 13334–13345.
- Bargaoui, Z., Trambay, Y., Lawin, E. A., and Servat, E. (2014). Seasonal precipitation variability in regional climate simulations over northern basins of Tunisia. *Int. J. Climatol.*, 34(1), 235–248.
- Besbes, M. (1978). *Estimation of Groundwater Recharge. A Regional Effective Infiltration Model*. Doc.th. Pierre & Marie Curie University, Paris.
- Besbes, M., Chahed, J., and Hamdane, A. (2014). *Sécurité Hydrique de la Tunisie, Gérer l'eau en Conditions de Pénurie*. Ed. L'Harmattan, Paris.
- Besbes, M., Chahed, J., and Hamdane, A. (2019a). Food and water management in Northwest Africa. In *The Oxford Handbook of Food, Water and Society*, page 426. Oxford University Press, New York.
- Besbes, M., Chahed, J., and Hamdane, A. (2019b). *National Water Security: Case Study of an Arid Country: Tunisia*. Springer International Publishing, Cham, Switzerland.
- Besbes, M., Chahed, J., Hamdane, A., and De Marsily, G. (2010). Changing water resources and food supply in arid zones: Tunisia. In Schneider-Madanes, G. and Courel, M. F., editors, *Water and Sustainability in Arid Regions*. Springer, Berlin.
- Bird, D. N., Benabdallah, S., Gouda, N., Hummel, F., Koeberl, J., La Jeunesse, I., Meyer, S., Pretenthaler, F., Soddu, A., and Woess-Gallasch, S. (2016). Modelling climate change impacts on and adaptation strategies for agriculture in Sardinia and Tunisia using AquaCrop and value-at-risk. *Sci. Total Environ.*, 543, 1019–1027.

- Bonnet, M., Delarozière, O., Jusserand, C., and Roux, P. (1970). In *Calcul des bilans d'eau mensuels et annuels par les méthodes de Tornthwaite et de Turc*. BRGM, France. 70SGN107HYD.
- BPEH (2019). *National Water Sector Report*. Office of Planning and Hydraulic Balances. BPEH. The Ministry of Agriculture, Tunisia.
- Bruyère, C. L., Done, J. M., Holland, G. J., and Fredrick, S. (2014). Bias corrections of global models for regional climate simulations of high-impact weather. *Clim. Dyn.*, 43, 1847–1856.
- Chahed, J., Hamdane, A., and Besbes, M. (2008). A comprehensive water balance of Tunisia: blue water, green water and virtual water. *Water Int.*, 33(4), 415–424.
- Collins, W. J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., et al. (2011). Development and evaluation of an Earth-System model—HadGEM2. *Geosci. Model. Dev.*, 4(4), 1051–1075.
- Cos, J., Doblas-Reyes, F., Jury, M., Marcos, R., Bretonnière, P. A., and Samsó, M. (2022). The Mediterranean climate change hotspot in the CMIP5 and CMIP6 projections. *Earth Syst. Dyn.*, 13(1), 321–340.
- Dakhlaoui, H., Hakala, K., and Seibert, J. (2022). Hydrological impacts of projected climate change on Northern Tunisian headwater catchments—an ensemble approach addressing uncertainties. In *Climate Change in the Mediterranean and Middle Eastern Region*, pages 499–519. Springer, Cham.
- de Marsily, G. (2008). Eau, changements climatiques, alimentation et évolution démographique. *Rev. Sci. Eau./J. Water Sci.*, 21(2), 111–128.
- de Marsily, G. (2020). Will we soon run out of water? *Ann. Nutr. Metab.*, 76(1), 10–16.
- de Marsily, G. and Abarca-del Rio, R. (2016). Water and food in the twenty-first century. In *Remote Sensing and Water Resources*, pages 313–337. Springer, Cham.
- Deidda, R., Marrocu, M., Caroletti, G., Pusceddu, G., Langousis, A., Lucarini, V., Puliga, M., and Speranza, A. (2013). Regional climate models' performance in representing precipitation and temperature over selected Mediterranean areas. *Hydrol. Earth Syst. Sci.*, 17(12), 5041–5059.
- DGRE (1980–2020). *Hydrological Yearbooks of Tunisia*. General Directorate of Water Resources (DGRE), The Ministry of Agriculture, Tunisia.
- DGRE, Prointec-Comete (2019). *Carte des Ressources en eau de la Tunisie. CRET. Phase III: Elaboration de la carte des écoulements Superficiels*. Direction Générale des Ressources en eau (DGRE), Ministère de l'agriculture, Tunis.
- Döll, P. (2009). Vulnerability to the impact of climate change on renewable groundwater resources: a global-scale assessment. *Environ. Res. Lett.*, 4(3), article no. 035006.
- FAO (2017a). AQUASTAT: FAO's global water information system, the Land and Water Division. <http://www.fao.org/nr/water/aquastat/main/index.stm>, consulted on 11 Oct 2022.
- FAO (2017b). FAOSTAT: food and agriculture data. Statistics Division. <http://www.fao.org/faostat/en/#data,k> consulted on 11 Oct 2022.
- Farsani, I., Farzaneh, M. R., Besalatpour, A. A., Salehi, M. H., and Faramarzi, M. (2019). Assessment of the impact of climate change on spatiotemporal variability of blue and green water resources under CMIP3 and CMIP5 models in a highly mountainous watershed. *Theor. Appl. Climatol.*, 136, 169–184.
- Fathalli, B., Pohl, B., Castel, T., and Safi, M. J. (2019). Errors and uncertainties in regional climate simulations of rainfall variability over Tunisia: a multi-model and multi-member approach. *Clim. Dyn.*, 52(1), 335–361.
- Foughali, A., Trambly, Y., Bargaoui, Z., Carreau, J., and Ruelland, D. (2015). Hydrological modeling in Northern Tunisia with regional climate model outputs: performance evaluation and bias-correction in present climate conditions. *Climate*, 3(3), 459–473.
- Fricko, O., Havlik, P., Rogelj, J., Klimont, Z., Gusti, M., Johnson, N., et al. (2017). The marker quantification of the Shared Socioeconomic Pathway 2: a middle-of-the-road scenario for the 21st century. *Glob. Environ. Change*, 42, 251–267.
- Hamed, M. M., Nashwan, M. S., and Shahid, S. (2022). A novel selection method of CMIP6 GCMs for robust climate projection. *Int. J. Climatol.*, 42(8), 4258–4272.
- Harris, I., Osborn, T. J., Jones, P., et al. (2020). Version 4 of the CRU TS monthly high-resolution gridded; multivariate climate dataset. *Sci. Data*, 7, article no. 109.
- Herger, N., Abramowitz, G., Knutti, R., Angélil, O., Lehmann, K., and Sanderson, B. M. (2018). Select-

- ing a climate model subset to optimise key ensemble properties. *Earth Syst. Dyn.*, 9, 135–151.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., and Muñoz Sabater, J. (2020). The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.*, 146(730), 1999–2049.
- Hoekstra, A. Y. (2003). *Virtual Water Trade: Proceedings of the International Expert Meeting on Virtual Water Trade, Delft, The Netherlands, 12–13 December 2002*. Value of Water Research Report, Series No. 12. IHE, Delft, The Netherlands.
- Hughes, J. P. and Guttorp, P. (1994). A class of stochastic models for relating synoptic atmospheric patterns to regional hydrologic phenomena. *Water Resour. Res.*, 30(5), 1535–1546.
- INM (2022). *Banque de Données Climatologiques. INM | Institut National de la Météorologie (météo.tn)*. Institut National de la Météorologie, Tunis.
- IPCC (2021). In Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., editors, *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK and New York, NY, USA.
- Jiang, Q., Li, W., Fan, Z., He, X., Sun, W., Chen, S., Wen, J., Gao, J., and Wang, J. (2021). Evaluation of the ERA5 reanalysis precipitation dataset over Chinese Mainland. *J. Hydrol.*, 595, article no. 125660.
- Jiao, D., Xu, N., Yang, F., et al. (2021). Evaluation of spatial-temporal variation performance of ERA5 precipitation data in China. *Sci. Rep.*, 11, article no. 17956.
- Kim, S., Eghdamirad, S., Sharma, A., and Kim, J. H. (2020). Quantification of uncertainty in projections of extreme daily precipitation. *Earth Space Sci.*, 7(8), article no. e2019EA001052.
- King, L., Nasr, Z., Almohamad, H., and Maag, C. C. (2007). le Climat. In *MARH, GTZ, Gopa, Exaconsult: Stratégie nationale d'adaptation de l'agriculture tunisienne et des écosystèmes aux changements climatiques*. Janvier, Tunis. Ch. 7.2.
- Knutti, R., Sedláček, J., Sanderson, B. M., Lorenz, R., Fischer, E. M., and Eyring, V. (2017). A climate model projection weighting scheme accounting for performance and interdependence. *Geophys. Res. Lett.*, 44, 1909–1918.
- Laurent, A., Fennel, K., and Kuhn, A. (2021). An observation-based evaluation and ranking of historical Earth system model simulations in the northwest North Atlantic Ocean. *Biogeosciences*, 18(5), 1803–1822.
- Li, X., Tan, L., Li, Y., Qi, J., Feng, P., Li, B., Liu, D. L., Zhang, X., Marek, G. W., Zhang, Y., Liu, H., Srinivasan, R., and Chen, Y. (2022). Effects of global climate change on the hydrological cycle and crop growth under heavily irrigated management—A comparison between CMIP5 and CMIP6. *Comput. Electron. Agric.*, 202, article no. 107408.
- Mandal, S., Breach, P. A., and Simonovic, S. P. (2016). Uncertainty in precipitation projection under changing climate conditions: a regional case study. *Am. J. Clim. Change*, 5(1), 116–132.
- Mankin, J. S., Seager, R., Smerdon, J. E., Cook, B. I., and Williams, A. P. (2019). Mid-latitude freshwater availability reduced by projected vegetation responses to climate change. *Nat. Geosci.*, 12(12), 983–988.
- Maraun, D. and Widmann, M. (2018). Cross-validation of bias-corrected climate simulations is misleading. *Hydrol. Earth Syst. Sci.*, 22(9), 4867–4873.
- MARH (1998–2012). *Annuaire des statistiques agricoles*. Ministère de l'Agriculture et des Ressources Hydrauliques, Tunisia.
- McMahon, T. A., Peel, M. C., Lowe, L., Srikanthan, R., and McVicar, T. R. (2013). Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis. *Hydrol. Earth Syst. Sci.*, 17, 1331–1363.
- McSweeney, C. F., Jones, R. G., Lee, R. W., and Rowell, D. P. (2015). Selecting CMIP5 GCMs for downscaling over multiple regions. *Clim. Dyn.*, 44, 3237–3260.
- Nasr, Z., Almohammed, H., Gafrej Lahache, R., Maag, C., and King, L. (2008). Drought Modelling under climate change in Tunisia during the 2020 and 2050 periods. *Option Méditerran. Séries A*, 80, 365–369.
- NOAA (2022). Global Historical Climatology Network monthly (GHcNm). [https://www.ncei.noaa.gov/data/ghcnm/v4beta/doc/ghcn-m\\_v4\\_prpc\\_inventory.txt](https://www.ncei.noaa.gov/data/ghcnm/v4beta/doc/ghcn-m_v4_prpc_inventory.txt).
- Oki, T., Sato, M., Kawamura, A., Miyake, M., Kanae, S.,

- and Musiak, K. (2003). Virtual water trade to Japan and in the world. In Hoekstra, A. Y., editor, *Value of Water Research Report, Series No.12*. IHE, Delft, The Netherlands.
- O'Neill, B. C., Tebaldi, C., Van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., et al. (2016). The scenario model intercomparison project (ScenarioMIP) for CMIP6. *Geosci. Model Dev.*, 9(9), 3461–3482.
- Oueslati, I., Lili-Chabaane, Z., Shabou, M., Zribi, M., Ben Issa, N., Chakroun, H., Galafassi, D., Rathwell, K., Hoff, H., and Pizzigalli, C. (2012). Methodology to Analyse the actual and the future effect of water scarcity on the available water resources in Merguellil watershed. *Geophys. Res. Abstr.*, 14, article no. EGU2012-9366-1.
- Ramirez-Villegas, J., Challinor, A. J., Thornton, P. K., and Jarvis, A. (2013). Implications of regional improvement in global climate models for agricultural impact research. *Environ. Res. Lett.*, 8(2), article no. 024018.
- Renault, D. and Wallender, W. W. (2000). Nutritional water productivity and diets: from crop per drop, towards nutrition per drop. *Agric. Water Manage.*, 45, 275–296.
- Riahi, K., Van Vuuren, D. P., Kriegler, E., Edmonds, J., O'Neill, B. C., Fujimori, S., and Tavoni, M. (2017). The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: an overview. *Glob. Environ. Change*, 42, 153–168.
- Schewe, J., Heinke, J., Gerten, D., Haddeland, I., Arnell, N. W., Clark, D. B., Dankers, R., Eisner, S., Fekete, B., Colón-González, F. J., Gosling, S. N., Kim, H., Liu, X., Masaki, Y., Portmann, F., Satoh, Y., Stacke, T., Tang, Q., Wada, Y., Wisser, D., Albrecht, T., Frieler, K., Piontek, F., Warszawski, L., and Kabat, P. (2014). Multi-model assessment of water scarcity under climate change. *Proc. Natl. Acad. Sci. USA*, 111(9), 3245–3250.
- Shokouhifar, Y., Lotfirad, M., Esmaeili-Gisavandani, H., and Adib, A. (2022). Evaluation of climate change effects on flood frequency in arid and semi-arid basins. *Water Suppl.*, 22(8), 6740–6755.
- Slama, E., Gargouri-Ellouze, E., and Bouhlila, R. (2020). Impact of rainfall structure and climate change on soil and groundwater salinization. *Clim. Change*, 163(1), 395–413.
- Somot, S., Ruti, P., Ahrens, B., Coppola, E., Jordà, G., Sannino, G., and Solmon, F. (2018). Editorial for the Med-CORDEX special issue. *Clim. Dyn.*, 51(3), 771–777.
- Switanek, M., Maraun, D., and Bevacqua, E. (2022). Stochastic downscaling of gridded precipitation to spatially coherent subgrid precipitation fields using a transformed Gaussian model. *Int. J. Climatol.*, 42(12), 6126–6147.
- Tarek, M., Brissette, F. P., and Arsenault, R. (2020). Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. *Hydrol. Earth Syst. Sci.*, 24, 2527–2544.
- Terink, W., Immerzeel, W. W., and Droogers, P. (2013). Climate change projections of precipitation and reference evapotranspiration for the Middle East and Northern Africa until 2050. *Int. J. Climatol.*, 33(14), 3055–3072.
- Wang, Z., Zhan, C., Ning, L., and Guo, H. (2021). Evaluation of global terrestrial evapotranspiration in CMIP6 models. *Theor. Appl. Climatol.*, 143(1), 521–531.
- Watanabe, S., Hajima, T., Sudo, K., Nagashima, T., Takemura, T., Okajima, H., Nozawa, T., Kawase, H., Abe, M., Yokohata, T., Ise, T., Sato, H., Kato, E., Takata, K., Emori, S., and Kawamiya, M. (2011). MIROC-ESM 2010: model description and basic results of CMIP5-20c3m experiments. *Geosci. Model Dev.*, 4(4), 845–872.
- WBG (2022). The World Bank Group, Climate Change Knowledge Portal. <https://climateknowledgeportal.worldbank.org/download-data>.
- WMO (2022). Climate Explorer: Starting point. KNMI Climate Explorer <https://climexp.knmi.nl/>.
- Zhang, M. Z., Xu, Z., Han, Y., and Guo, W. (2022). Evaluation of CMIP6 models toward dynamical downscaling over 14 CORDEX domains. *Clim. Dyn.*, pages 1–15.
- Zimmer, D. (2013). *L'empreinte eau. Les faces cachées d'une ressource vitale*. Charles Léopold Meyer, Paris.



# **Karsts & Submarine Springs**





Research article

Geo-hydrological Data & Models

# Coastal karst aquifers and submarine springs: what future for their water resources?

Perrine Fleury<sup>✉,a</sup>, Séverin Pistre<sup>✉,b</sup> and Michel Bakalowicz<sup>✉,\*,b</sup>

<sup>a</sup> BRGM direction régionale Occitanie, Montpellier, France

<sup>b</sup> HSM, Univ. Montpellier, CNRS, IMT, IRD, Montpellier, France

*E-mails:* p.fleury@brgm.fr (P. Fleury), severin.pistre@umontpellier.fr (S. Pistre),

michel.bakalowicz@gmail.com (M. Bakalowicz)

**Abstract.** This paper is intended to be a reminder of a series of scientific works whose origin is largely attributable to G. de Marsily, without attempting to make an exhaustive overview of submarine springs and coastal karst aquifers. The main results show that coastal karst aquifers with their submarine outlets and the possibility of natural seawater intrusion are particularly fragile and must be studied and managed with appropriate tools. For the coming decades, the evolution of the climate as results from the IPCC scenarios leads us to question the effects of the increase of sea level on this coastal resource, and the effects of frequent and severe droughts, added to an increase in withdrawals from aquifers caused by the coastalisation of the population.

**Keywords.** Groundwater, Resource, Management, Coastal karst, Submarine spring, Mediterranean.

*Manuscript received 31 March 2022, revised 22 September 2022, accepted 30 September 2022.*

## 1. Introduction

In the early 2000s, advances in reconnaissance and offshore work made it possible to consider tapping water from submarine springs to offset the decreasing availability of water in coastal regions where development projects and new population settlements were flourishing. Specialized companies sought to develop devices to capture this unconventional resource. In 1999, on a television program, an underwater contractor had just conducted an experiment to try to show that the capture of water emerging from a submarine spring was feasible. The spectacular demonstration consisted of placing an “umbrella” over the Mortola spring located in Italy, close to the

French border, at 36 m below sea level (bsl). The entrepreneur, founder of a French company dedicated to this objective, turned to G. de Marsily, who was associated with studies on karstic springs of Mediterranean coastal aquifers [Arfib et al., 2002, Arfib and de Marsily, 2004], in order to obtain scientific and technical support for the understanding of hydrogeology and thus ensure the success of his project.

Since the 1960s, local studies have focused on some of the most spectacular coastal or submarine springs [Kuscer and Kuscer, 1962, Moulard et al., 1965], sometimes with a view to exploiting the resource. The best example is that of the Port-Miou spring (calanque of Cassis, close to Marseille, South of France), on which the Société des Eaux de Marseille had undertaken studies and work to try to lower the salinity of its water with a view to using it [for a synthesis, see Tardieu and Potié, 2015]. The sub-

\* Corresponding author.

marine outlet of this large karstic conduit, emerging 12 m bsl at the bottom of the Port-Miou cove has long been known. It was formally identified by de Marsigli [1725] in the early 18th century. Considered as the precursor of oceanography [Toulmond, 2006], de Marsigli associated this important and continuous discharge of underground fresh water to a conduit draining the Provence hinterland; but above all, he was able to bring this phenomenon closer to what he had previously observed in Lika, where karstic plateaus are the origin of submarine discharge of fresh water along the Croatian coast of the Adriatic Sea.

It was not possible to respond to the request of the company without addressing questions that had not yet been resolved by these dispersed studies. In particular, it appeared necessary to characterize these submarine outlets and to improve the understanding of the functioning and structure inherited from eustatic variations of the aquifers feeding these unconventional discharges, in order to build a typology of the springs, with a view to their eventual capture. In addition, an exhaustive inventory of existing knowledge and data was essential to serve as a basis for a market study and commercial prospecting with a view to establishing these catchments [Fleury, 2005, Fleury *et al.*, 2007b, 2008].

The development of a prototype (Figure 1) to capture the Mortola spring, near Menton [Fleury *et al.*, 2007a] showed the difficulties of production, due mainly to two factors which are very difficult to control: the evaluation of the flow at the spring, and the sealing of the device. They are both dependent on the karstic nature of the spring and its aquifer, favoring a certain variability of the flow and the existence of leaks linked to the karstic conduit system. Subsequent studies have confirmed that this problem of saline intrusion is inherent to the karstic nature of the reservoir characterized by conduits and fractures connected to the marine environment, sometimes at great depth.

The global inventory of submarine springs [Fleury, 2005] highlighted several key points:

- Submarine springs are a very special case of Submarine Groundwater Discharge [SGD, Zektser *et al.*, 1983], since they are localized, concentrated, fresh or brackish water sources; they are therefore associated with a specific morphological and biological envi-



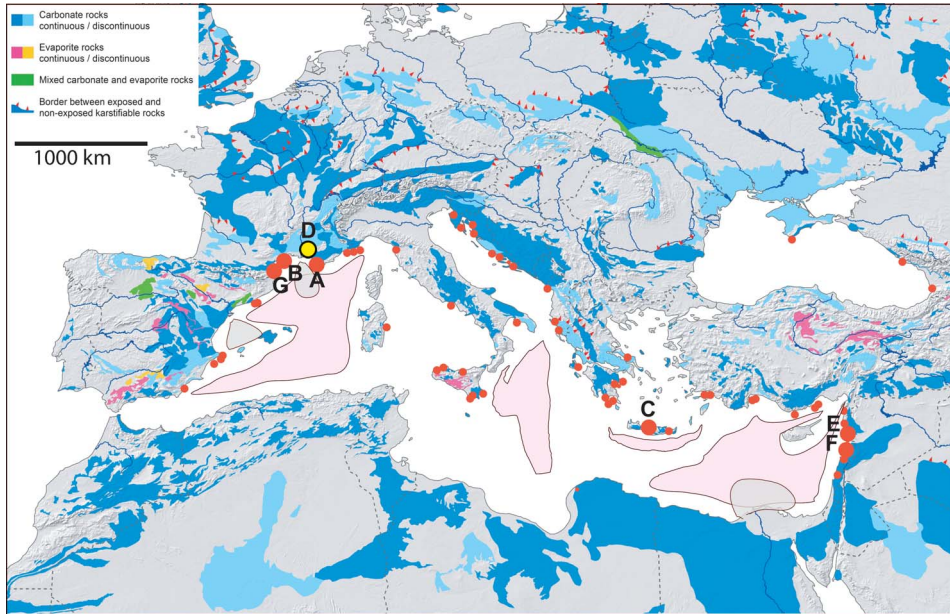
**Figure 1.** Ghislain de Marsily in front of the prototype of the catchment device of the Mortola spring, before its installation, in the port of Menton, France, in 2003 (Photo by M. Bakalowicz).

ronment. Hence, they could be relevant water supplies.

- The vast majority are associated with karst aquifers. The rest are linked to highly fractured volcanic aquifers.
- Most of the known springs are located in the Mediterranean basin, related to Jurassic and Cretaceous limestone in large areas (Figure 2).

These primary works allowed the establishment of collaborations that led to a European project for the study of non-conventional water resources, mainly those of the Mediterranean coastal karst aquifers (MEDITATE project, 2004–2007, <https://cordis.europa.eu/project/id/509112>). Sites in Turkey, Syria and Lebanon, identified as corresponding to submarine springs of great interest for supplying population needs, have been studied by different complementary approaches.

In this paper, the main contributions of these investigations will be summarized. We will then show the bottlenecks that make these littoral karst aquifers still very difficult, if not impossible, to exploit in the framework of an integrated regional management. Finally, we will share our thoughts on the future of these resources in the face of the challenges posed by climate change and population growth in these re-



**Figure 2.** Map of karst areas around the Mediterranean, taken from WOKAM map [Goldscheider *et al.*, 2020] with the location of the main known karst submarine springs with red dots [adapted from Fleury, 2005]. The three main remaining hypersaline basins at the Messinian low stage are shown in pink with the main post-Messinian fluvial fans in grey (adapted from the maps from [https://fr.wikipedia.org/wiki/Crise\\_de\\_salinit%C3%A9\\_messinienne](https://fr.wikipedia.org/wiki/Crise_de_salinit%C3%A9_messinienne)). The main coastal and submarine springs cited in the text are represented with larger red dots: A: Mortola, France; B: Port-Miou, France; C: Almyros of Heraklion, Crete, Greece; D: Fontaine de Vaucluse, France; E: Gulf of Bassieh, Syria; F: Gulf of Chekka, Lebanon; G: Fontestramar, France.

gions.

## 2. Main results

The comprehensive bibliography carried out on submarine springs, as well as on many coastal springs, and Mediterranean coastal karsts [Fleury, 2005, Fleury *et al.*, 2007b] revealed that groundwater from all these outlets is brackish, for at least a part of the year, and therefore unfit for drinking water. Originally, either a very local marine intrusion into the conduit or a leaching of saline formations into the aquifer was invoked. The “inversac” phenomenon was first described and interpreted by de Serres [1849], and then by Gèze [1938], in a channel between the Ambressac coastal karstic spring and the marine lagoon of Thau, near Montpellier, South of France. He described two inverse currents driven by the difference in density, one of fresh water at the surface towards the sea, the other at the bottom towards the spring, interpreted as an intrusion of

marine water into the aquifer. Inversacs are well documented also at the La Vise submarine spring, in the Thau lagoon, near Montpellier, France [Ladouche *et al.*, 2018]. This is a mechanism typical of conduit flow, particularly sensitive to fluctuating water head relationships, between fluids of different density.

The same phenomenon of flow stratification driven by density difference occurs inside the main conduit at Port-Miou, where initial work led to the building of a partial dam in the bottom of the karst conduit to prevent intrusion of the incoming marine water flow, in order to be able to capture freshwater [Tardieu and Potié, 2015]. This construction allowed the interruption of direct seawater intrusion through the base of the conduit.

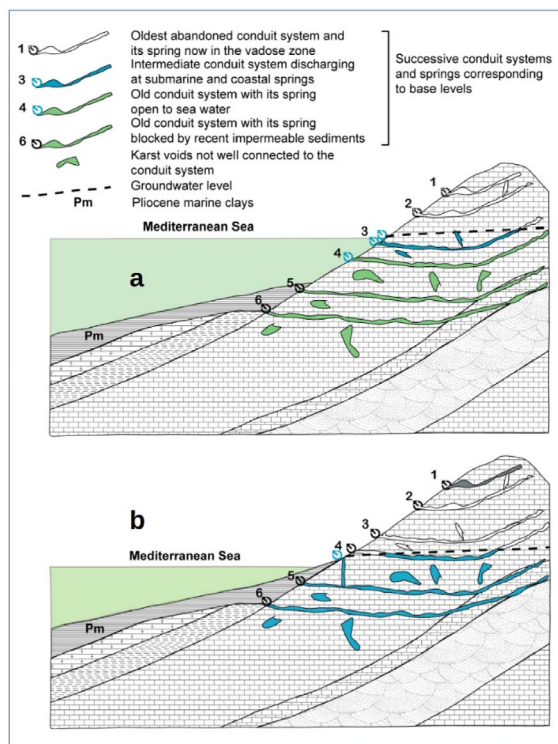
Unfortunately, salinity only decreased slightly, showing that most of the salinization does not come from this localized conduit, but probably from a more diffuse intrusion in the aquifer at depth. Chemical and isotopic analyses ruled out contamination by dissolution of gypsum–salt formations, such as

those in the Triassic [Blavoux *et al.*, 2004]. More recent studies on the Almyros of Heraklion (Crete, Greece) concluded to a marine intrusion in the aquifer at around 600 m bsl [Arfib and de Marsily, 2004]. The intrusion was supposed to occur through a conduit-matrix system, implying the existence of karst conduits at a depth incompatible with the lowering of the sea level during Quaternary glaciations (120 to 140 m bsl).

Only the development of karst features at depth due to a considerable lowering of the sea level can account for both the remarkable occurrence of karst submarine springs (KSMS), brackish coastal springs along the Mediterranean coast and the salinization of carbonate aquifers by marine intrusion at great depth. The Messinian Salinity Crisis [MCS,  $-7.2$  to  $-5.3$  Ma; Rouchy and Saint-Martin, 1992, Rouchy *et al.*, 2006] is the major geological event that explains this situation [Fleury *et al.*, 2007a]. With a maximum lowering of the sea level from 1500 to 2500 m during a few hundred thousand years, the valleys of the main rivers were deepened to reach the new sea level, in a few residual basins (Figure 2). At the same time, karstification was able to develop in all the carbonate massifs of the basin.

Numerous cave diving explorations have been carried out in recent years in flooded karstic conduits connected to the outlets (Fontestramar and Port-Miou, South of France). Particularly impressive and perilous, they have reached record depths (286 m bsl reached by X. Méniscus in 2015 at Fontestramar, <https://o-dive.com/record-du-monde-pour-xavier-meniscus/>). The objective of these surveys was to map the coastal karst network in order to identify the conduits responsible for the salt intrusion and eventually consider their closure. Given the complexity and the multitude of karst voids, some large ones connected to the sea which induces both diffuse and localized saline intrusion, these experiments are unfortunately doomed to failure.

The research conducted in particular within the framework of the MEDITATE project on different types of Mediterranean coastal karst systems has made it possible to propose a typology, serving as a basis for a specific strategy of exploitation and protection of the groundwater resource [Bakalowicz, 2018]. Karst systems located along the Mediterranean coastline during the MCS appear to be unique among karst areas of the world. Huge conduit sys-



**Figure 3.** The two main types of post-Messinian coastal karst systems in the Mediterranean basin. (a) Plugged, and (b) unplugged.

tems developed in a relatively short geological time, in carbonate formations, several hundred to several thousand meters thick. They developed over large recharge areas, extending to all coastal carbonate massifs. They ceased to function abruptly as a result of the Zanclean sea level rise (5.33 Ma). At many places mainly related to river valleys, sedimentation of marine blue clays occurred, covered by Pliocene delta sediments. The Zanclean transgression created two main hydrogeological situations of post-Messinian karst systems (Figure 3).

### 2.1. *Plugged karst systems*

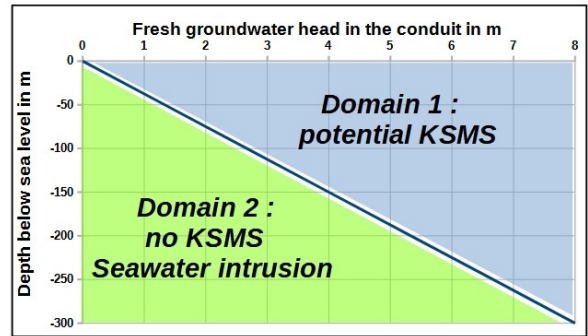
The systems that discharged into the main valleys during the MSC became plugged with marine clays during the Zanclean transgression (Figure 3a). They were totally filled up with fresh groundwater which discharged generally at the summit of Pliocene formations. Their characteristics can be compared to

those of basins and grabens, i.e., a large extent that provides abundant resources and a large storage capacity [El Hakim and Bakalowicz, 2007]. In the lower Rhone River valley, France, Fontaine de Vaucluse is a typical example of these post-Messinian karst systems. These are the Vaucluse-type springs, overflow of the phreatic zone after the plugging of the original springs. However, their conduit systems are typically too deep to be intersected by wells, except sometimes near the spring as in the case of the Lez spring, France [Maréchal *et al.*, 2013].

In some cases, marine clays or other formations such as basalts may have plugged coastal aquifers, thus confining groundwater and preventing seawater intrusion. This is the situation at the El Sin spring and Banyias karst aquifer in the Gulf of Bassieh, Syria [Abou Zakhem and Hafez, 2007], and at the Gradole, and Jadro springs, Croatia [Mijatovic, 2007]. These aquifers are exploited for water supply thanks to wells placed 100 m from the sea shore at Banyias, or pumping stations at the springs in Croatia. These karst systems are the most suitable for active management, using the significant reserves as a buffer to allow pumping of the groundwater at rates close to their mean discharge during the low-flow season, as is done at the Lez spring [Maréchal *et al.*, 2013].

## 2.2. Unplugged coastal karst systems

However, most carbonate formations along the coasts were either not covered with sediments or the sediments may have eroded after MSC. Consequently, surface karst features such as sinkholes which are now flooded are generally connected to conduit systems that still function inland in the recharge area (Figure 3b). Because of their karstic characteristics, most unplugged system KSMS are subject to seasonal variations of flow rate and quality. As long as the fresh-water head in the conduits is higher than the seawater head, the karst systems discharge at submarine and coastal springs corresponding to these flooded karst features. Numerous coastal and submarine springs are known along the Mediterranean coast. During periods of drying up, the piezometric level decreases. At the freshwater–saltwater interface, due to density differences, a reversal of the hydraulic gradient can occur. Groundwater head may be lower than the seawater head (Figure 4, green

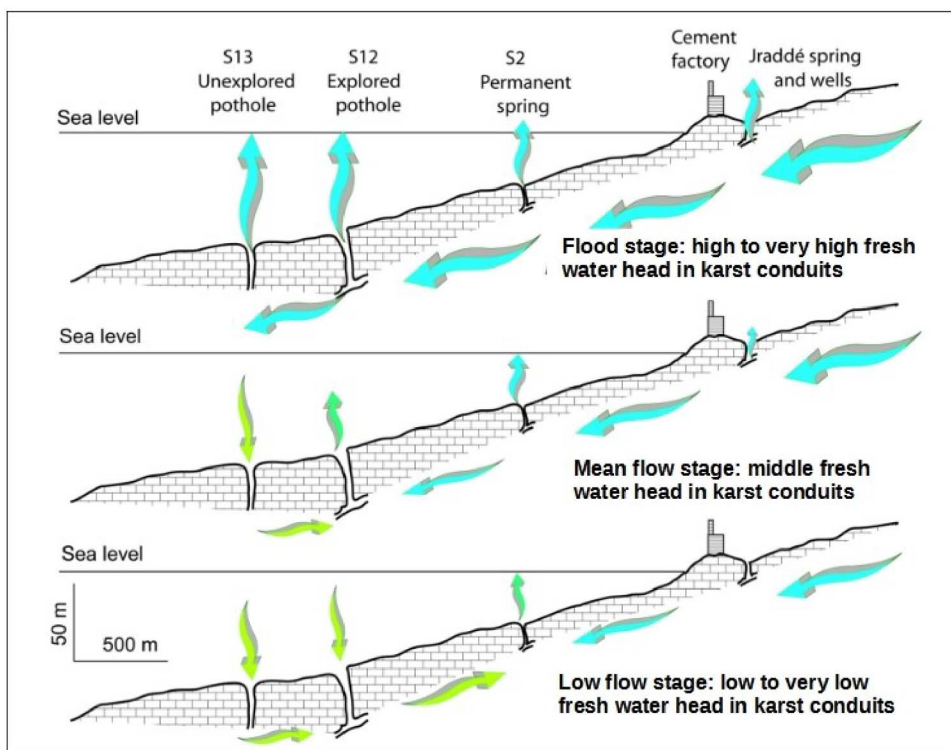


**Figure 4.** Sea-water vs fresh groundwater in coastal aquifers. Diagram showing the depth below sea level at which a KSMS may discharge, depending on the fresh water head in the conduit discharging at the spring. The blue zone is the most frequent domain of submarine springs discharging fresh or brackish groundwater. The green zone is the domain of common sea water inflow in coastal aquifers, through karst flooded features.

domain). The conduits may then swallow seawater which salinizes groundwater at least into the conduits, up to KSMS and even to coastal springs which then discharge brackish water. The salinity varies according to (1) the sea level, and (2) the fresh groundwater level in the conduit system, as shown by the data from the Chekka karst system, Lebanon [El-Hajj, 2008, Bakalowicz and El-Hajj, 2012], the best documented example to this day (Figure 5).

## 3. Technical and scientific obstacles

The first main questioning is related to the assessment of their flow rate and seasonal variations. The economic value of capturing a KSMS mainly depends on the potentially exploited volume of water and its quality. Various authors have used different methods to assess it, including stable isotopes, remote sensing and visual estimates from the dimension of the mark at the sea surface. Generally, it is almost impossible to make a direct and continuous flow measurement, except during low stage for systems with a unique spring. The specific layout of the Port-Miou spring made the monitoring of the salinity and the flow rate inside the conduit relatively easy, which was around



**Figure 5.** Scheme illustrating the functioning of the Chekka coastal discharge system, on a cross section along the Nahr el Asfour River paleo-valley, North Lebanon [modified from El-Hajj, 2008]. The deepest flooded karst features swallow seawater during the low flow stage. The permanent spring S2 discharges brackish water. Blue arrows: fresh water flow; green arrows: seawater flow; blue green arrows: brackish water.

1.5 m<sup>3</sup>/s at the beginning of the low flow stage, following the last recharge [April to May 2006; Arfib and Douchet, 2011, Cavalera, 2007].

Direct measurements by means of flow meters in pipes catching the output discharge were performed at the Mortola and Chekka KSMS at several times. Indirect assessments were attempted by delineating the recharge area, i.e., the extension of the karst system from geology and tracing tests, and the calculation of the annual water balance, the input flow at swallow holes. Both methods show that previous estimates of instantaneous as well as mean annual discharges were exaggerated by 5 to 10 times in the few studied cases. This overestimation is related to the acceleration of the flow of water from the aquifer when it reaches the marine environment due to the difference in density. This phenomenon creates a narrow column of fresh or brackish water, whose velocity is driven by the difference of density with sea-

water, not at all by the natural flow at the outlet. The divers roughly assessed the discharge from either the flow velocity above the spring or even from the extent of the surface anomaly created by the upwards flow.

Modeling of the plume above the spring has been attempted to analyze flows, including return currents carrying sediment to the bottom [Buongiorno Nardelli *et al.*, 2017]. These could be used to estimate flow from accurate in-situ measurements. A Submarine Underwater Vehicle had been developed to explore these plumes and to determine all the parameters needed for such modeling [Bakalowicz *et al.*, 2003a,b]; however, the project was not completed.

Several KSMS may have a mean annual discharge of several m<sup>3</sup>/s which makes them interesting potential water resources in coastal areas that often lack water. However, their hydrological functioning is typically karstic, characterized by strong seasonal, even



daily, variations of discharge and water quality, as occurs at Chekka Bay submarine springs, Lebanon [El-Hajj, 2008]. The recharge area of the Chekka karst system covers about 250 km<sup>2</sup> of North Lebanon. Its main KSMS discharges at 23 m bsl 0.1 m<sup>3</sup>/s during low flow stage, of brackish water at 35% of seawater, decreasing to 0.030 m<sup>3</sup>/s at 50 to 80% of seawater at the end of the dry season, while the first estimate suggested 2 m<sup>3</sup>/s of fresh water [Moulard *et al.*, 1965]. During winter floods, the system discharges at several KSMS, with the most important being found at 110 m bsl. The water balance gives an average of 2.15 m<sup>3</sup>/s, with a high reliability as the hydrogeological catchment is well controlled, while Moulard *et al.* [1965] estimated it to be at least 25 m<sup>3</sup>/s.

Several tools are presently very efficient for identifying SGD from coastal karst aquifers [Jou-Claus *et al.*, 2021, Schubert *et al.*, 2014]. Despite there being no reliable method for assessing their discharge and consequently their actual groundwater resource, especially in the Mediterranean basin, some risky estimation claims that “*Submarine groundwater discharge can dominate inflows in restricted seas or embayments. For example, karst comprises 60% of the shoreline of the Mediterranean and is estimated to contribute 75% of its freshwater input – much of this input must be via SGD.*” [UNESCO, 2004]. However, in the same publication [UNESCO, 2004, p. 19], the authors write that: “*Recent attempts to quantify the amount of total underground fluid flow into Florida Bay Submarine groundwater discharge by modeling vary by 3–4 orders of magnitude!*”.

Fresh groundwater discharge from coastal karst aquifers can also be calculated directly in water budgets. Bayari *et al.* [2011] estimated from water balance that the coastal karst aquifers of the southwestern Mediterranean coast of Turkey should discharge an average of about 360 L/s per km of carbonate shoreline, 120 km long. The data collected by the WOKAM project [Goldscheider *et al.*, 2020] should allow an assessment of much better reliability.

In all these unplugged coastal karst systems, the conduits are open to the sea at different depths, which makes easy natural localized seawater intrusion, especially during the low flow stage. All submarine and many coastal springs discharge brackish water, at least during the dry season. It was shown that seawater intrusion may occur some hundred meters below sea level [Arfib and Douchet, 2011, Bakalowicz

and El-Hajj, 2012].

Various tests have been done to capture submarine springs using, for instance, particular bells to cover the outlet as it was done by Stefanon [1972] in the Gulf of Taranto, Italy. The system that drives fresh or brackish water to the sea surface reduces the water head over the spring, which increases the seawater head in the conduit, enhances seawater intrusion, and raises the salinity of the captured water. Capturing “fresh” water from karst submarine springs thus appears to be an impossible challenge. As a result, these submarine springs are not exploitable, not only because of the low discharge during the dry season but also because of brackish water. Brackish water can be desalinated [Ahdab and Lienhard, 2021, Breznik and Steinman, 2011] for supply water, which consumes less energy than desalinating seawater. However, the process implies that the salinity should be stable, a condition which is not fulfilled at KSMS. Therefore, it seems much better to tap groundwater from inland wells, and monitor salinity and flow at submarine springs, wells or coastal springs for reducing seawater intrusion and preventing it by reducing pumping rates in wells [Bakalowicz, 2018]. KSMS then should be considered as interesting tools for groundwater resources management in coastal areas.

#### 4. Biodiversity and exchanges between sea and groundwater from karst aquifers

The investigations of nutrient inputs to the ocean from continental waters, including SGD [Santos *et al.*, 2021], have been conducted for about 20 years to assess their contribution to marine biodiversity. The concentrated discharge of karst ground water from submarine springs is obviously being investigated [Bejannin *et al.*, 2020, Tamborski *et al.*, 2020]. It is considered negligible in quantity, but important in quality [Luijendijk *et al.*, 2020]. However, the volumes of freshwater discharged directly to the sea are incomparably smaller than those from rivers, such as the Nile, the Rhone and the Po Rivers. In addition, most of the Mediterranean KSMS is subject to a very high seasonal variability of flow associated with that of their salinity: on the whole, the high flows of the rainy season and snowmelt (October to April) bring fresh water, rather rich in nutrients of continental origin, into a sea agitated by storms, while the rest

of the year the waters are rather brackish, with a low land contribution. Thus, a limited overall contribution to marine biodiversity is to be expected. Reliable data are still lacking relating nutrients and fauna richness around these underwater sources.

In fact, from a biodiversity standpoint, the input of seawater into coastal karst aquifers likely has a major impact on biodiversity. But it seems that the part played by the MCS is not yet fully considered by scientists. For a long time, the hydrobiologists have shown [Rouch and Danielopol, 1987] that the underground aquatic fauna has its origin in the surface fauna, which takes refuge in the aquifers or colonizes them. Studies showed that in France at least, the large Mediterranean karstic aquifers contain an abundant, diversified and originally hypogean aquatic fauna, compared with those of other regions [Deharveng *et al.*, 2009]. Thus, one aspect of the impact of MCS on Mediterranean coastal karst aquifers, currently ignored, is that of the introduction of marine species into the aquifers during the Zanclean transgression (5.3 Ma). Moreover, since this major abrupt event [Lofi *et al.*, 2005], all unplugged karst systems are subjected to permanent exchanges of marine water and fauna, which allow an active colonization of the subsurface environment by epigeous organisms. This is a research field that deserves to be developed, in order to give responses to fundamental questions about adaptation and evolution in fresh groundwater [Mammola *et al.*, 2020].

## 5. Global changes and coastal karst aquifers

Unlike porous and fissured aquifers, the mechanism of marine intrusion in karst aquifers cannot be described spatially by a simple physical law. Indeed, the existence of drainage structures inherited from a long and complex geological history creates a heterogeneity of groundwater flows such that seawater can naturally penetrate the aquifer very deep inland. As shown by the phenomenon of *inversac* and the salinization of the aquifer, the coastal karst aquifers, with their conduits open to the sea at variable depths, put into a very fragile equilibrium relationship the fresh water flowing from the continent and the marine water that repels it by density effect. This fragile equilibrium is prone to a natural decrease of the

freshwater head, due to the reduction of groundwater recharge, consequence of a reduction of the effective rainfall and/or of soil sealing, and to the increase of the marine water head by the rise of the sea level. In addition, anthropogenic forcing, of which pumping is the main factor, obviously enhances considerably the natural effects. The following examples can give an idea of the seawater intrusion issue in karstic aquifers.

The Burren karst in Ireland, subject to low anthropogenic forcing, has been studied to characterize the hydrodynamic functioning of a karst aquifer in relation to seawater movement [Perriquet *et al.*, 2014]. The salinity and groundwater levels were measured in boreholes and also directly in the phreatic zone to assess the extent of the intrusion of the seawater wedge into the aquifer as a function of both karst recharge and tidal movements. The marked spatial disparity of the seawater wedge was analyzed as a function of both the hydrodynamic and the structural properties of the karst aquifer. Results showed that the extent of the seawater wedge depends not only on the intrinsic properties of the aquifer, but also on the relative influence of recharge and tide effects on groundwater levels, which are opposite. Recharge in the Burren region throughout the year is sufficiently high to prevent seawater intrusion more than about one kilometer from the shoreline. A strong tidal amplitude appears to drive the sudden seawater intrusion observed in the nearshore aquifer, while the position of the groundwater level appears to influence the intensity of the increase in salinity.

This study highlights the importance of monitoring for characterizing the salinization phenomenon in coastal karst aquifers. The same approach was applied at two coastal karsts in the south of France: the karst aquifer of Aumelas discharging into the Thau Lagoon [Pétre *et al.*, 2020] and the Eastern Corbières aquifer, both located to the West of Montpellier. These aquifers are instrumented with salinity, temperature and groundwater level data loggers to assess the seawater intrusion, as part of a research project conducted by the French Geological Survey (BRGM). The outlets, namely the Vise submarine spring in the Thau lagoon, at 30 m bsl, and Fontestramar coastal spring, at +4 m above sea level (asl) are monitored, as well as boreholes and surface flow upstream swallow holes. The study on the Thau karst

is currently ongoing.

Over the last 50 years, the Thau karst aquifer was subject to seven brackish water intrusions via the submarine spring of La Vise. These events, called *inversacs*, took place in 1967, 1983, 1993, 2008, 2010, 2014 and 2021–2022 [Ladouche *et al.*, 2018]. The last episode was particularly long, lasting 14 months. However, their duration varies from 1 month to almost a year and a half. Over the period from 1967 to 2007, only three *inversacs* occurred in 40 years, which represents a low decadal occurrence, while since 2008 four *inversacs* occurred in only 14 years. The frequency of occurrence increased significantly, related to an increase in groundwater pumping, associated with the increase in population and to irrigation, and probably also to an increase in drought periods linked to climate change.

The karst of the Eastern Corbières area offers an important groundwater resource, with an average flow of 2.5 m<sup>3</sup>/s, i.e., nearly 80 million m<sup>3</sup>/year, but the chloride concentration can reach 2.6 g/L, in a seasonal salinity range from 0.3 to 3.5 g/L [Dörfliger *et al.*, 2008]. Furthermore, the episodes of intense rainfall seem to be intensifying. In this context, in catchments with a karst component, the understanding of the relationship between surface flows and karst aquifers is becoming essential, in order to be able to forecast floods and thus anticipate their effects.

Large-scale changes impacting groundwater resources were already observed around the Mediterranean. Thus, a regional study carried out in Morocco [Ahmed *et al.*, 2021] has shown a drop in groundwater levels in several basins, which led to the disappearance of the water table in areas with shallow aquifers. One of the main factors is related to the change in agricultural practices from rain-fed to agriculture irrigated by groundwater. This sudden shift occurred at the expense of careful planning for the sustainability of groundwater resources. The consequential effects are an increasing risk of contamination by seawater intrusion, an increasing desertification, increasing costs of groundwater pumping and deterioration of ecosystems. These threats are expected to continue and even increase in the future if groundwater levels continue to decline. In order to counteract these threats, Morocco began to develop a policy to control groundwater extraction and use. This policy aims to promote efficient and economic

use of groundwater resources, particularly in threatened aquifers, and to strengthen regulatory and institutional enforcement.

The assessment of the relative effects of climate change and anthropogenic forcing was also approached by coupling different model simulations under different climate change scenarios with hydrological models. This was done by Sivelle *et al.* [2021] on a Mediterranean karst system in the South of France. They coupled the results with scenarios of future groundwater extraction for drinking water supply (no abstraction, present-day abstraction, +50% abstraction and +100% abstraction at horizon 2100). They show that climate change will have the main effect; the extraction scenarios will be a secondary but non-negligible factor.

These regional resource management policies are developing around the world. Another example can be taken from an important karst aquifer in Florida. In this context, Silver Springs, a large karst spring system with an average flow rate of over 20 m<sup>3</sup>/s, draining the Floridan aquifer, was studied [Klammler *et al.*, 2020]. The recharge area, approximately 2300 km<sup>2</sup>, is underlain by an aquifer system containing salt water at a depth of about 500 m, which is connected to the ocean through a permeable formation. It is located near the center line of the Florida peninsula, approximately 80 km from both the Atlantic and Gulf coasts. The recharge–storage–discharge dynamics of freshwater in this aquifer were characterized and a parsimonious model was developed. This work gives opportunity for regional and long-term management of groundwater resources.

The sudden modification of salinity conditions near the submarine springs and more widely in the coastal lagoons that they feed should lead to a modification and a rapid reduction of biodiversity with loss of species. Stenohaline or euryhaline organisms, for strictly halophilic organisms, are likely to disappear, with serious consequences for the fish farming and aquaculture activities linked to these lagoons, including socio-economic disturbances for the local populations, who are dependent on these activities [Rodríguez Estrella, 2000]. Finally, saline intrusion far from the coast may affect the catchments for drinking water supply as well as those used for agricultural operations in the coastal areas [Romero Macias *et al.*, 2000].

These complex coastal karst systems are characterized by a fragile state, very sensitive to seasonal variations in recharge, but also to sea level. In the context of climate change, an increase in the frequency and duration of drought periods is predicted, which will intensify seawater intrusion, due to the lowering of the hydraulic head in the conduits. Moreover, the sea level rise will magnify this effect, so that the quality of their groundwater will be deeply modified, at least in the parts closest to the coast.

It is clear that, because of their extreme sensitivity to variations in the equilibrium of heads between fresh groundwater and seawater, coastal karst aquifers play a watchdog role, likely to alert very quickly to unfavorable evolutions of the groundwater resource due to overexploitation and/or a decrease in recharge. Regular or early salinization during the year as well as the frequency of flow inversion, as in the case of the Vise spring, constitutes an alert concerning an unfavorable evolution of the groundwater resource. Under these conditions, coastal karst springs must be considered as observatories for monitoring coastal groundwater resources and saline intrusion. The implementation of coastal water resource observatories with monitoring of electrical conductivity is therefore becoming essential.

## 6. Summary

Coastal karst aquifers are beginning to be well identified on a global scale, by their location, their coastal and submarine springs and their general functioning. The vast majority are located in the Mediterranean basin, where their origin and their particular functioning are inherited from a major geological event, the Messinian salinity crisis (−7.2 to −5.3 Ma), responsible for the development of karst several hundred meters below the current sea level. Outside the Mediterranean basin, karst could only develop a few tens of meters below the current level, during the eustatic crises of the Quaternary.

Two types of coastal karst aquifers have been identified:

- those protected from marine intrusion by impermeable post-Messinian formations (marine blue clays, basalt), called “plugged karst systems”;
- those whose discharge zone, often including surface karst phenomena, now flooded, directly in contact with marine waters, called “unplugged karst systems”.

The second type of coastal karst aquifer, very common along the Mediterranean coast, is the origin of the numerous seasonal and permanent karst submarine springs (KSMS). These aquifers are very difficult to exploit, because the slightest variations in hydraulic head between fresh groundwater in the conduits and seawater facilitate natural saline intrusion, sometimes far from the coast, and at significant depths. This is why pumping operations require permanent observation of all recharge and withdrawal data, by setting up water head and salinity monitoring networks, possibly associated with artificial intelligence tools. Furthermore, direct tapping of these KSMS is not generally feasible, precisely because it contributes to modifying the delicate load balance between the sea and the karst conduits, thus facilitating the seawater intrusion. In any case, a good hydrogeological knowledge of these systems must precede the definition of the monitoring network and the management plan of these coastal aquifers.

The submarine discharge of coastal karst aquifers certainly impacts marine biodiversity by carrying nutrients and dissolved matter from the continent. But these impacts are limited compared to the much larger effects from rivers. However, the biodiversity of numerous lagoons fed by karst sources are largely modified and reduced in favor of halophilic species, but not well documented. In fact, the input of seawater into coastal karst aquifers likely has a major impact on groundwater biodiversity, by introducing marine species which then colonize aquifers, and adapt progressively to the constraints of subterranean conditions. The role of the MCS in the colonization of coastal karst aquifers would require careful attention.

Population movements towards the coasts create very strong pressures on these fragile resources, including increased groundwater withdrawals, to which will be added those related to climate change, such as the rise in sea level, the different seasonal distribution of rainfall and the increase in evapotranspiration. Coastal karst aquifers offer a freshwater resource that is certainly considerable, but needs to be evaluated, for the entire Mediterranean region. In order for them to really contribute to the maintenance

of human activities in the best possible conditions, it becomes necessary, and even urgent, to study them in detail in order to protect them from uncontrolled intensive withdrawals, which could rapidly cause their general and irreversible salinization. The increase in groundwater withdrawals linked to the coastalisation of populations and activities, the reduction in aquifer recharge, and the rise in sea level will inevitably and rapidly have consequences on the salinization of coastal karst resources. Under these conditions, some karst systems will have to be abandoned and the water supply policy for coastal cities will have to be rethought.

### Conflicts of interest

Authors have no conflict of interest to declare.

### Acknowledgement

We are particularly grateful to Mrs Madeleine Wheatley who revised the English writing.

### References

- Abou Zakhem, B. and Hafez, R. (2007). Environmental isotope study of seawater intrusion in the coastal aquifer (Syria). *Environ. Geol.*, 51, 1329–1339.
- Ahdab, Y. D. and Lienhard, J. H. (2021). Desalination of brackish groundwater to improve water quality and water supply. Chap. 41. In Mukherjee, A., Scanlon, B. R., Aureli, A., Langan, S., Guo, H., and McKenzie, A. A., editors, *Global Groundwater Source, Scarcity, Sustainability, Security, and Solutions*, pages 559–575. Elsevier, Amsterdam.
- Ahmed, M., Aqnouy, M., and El Messari, J.-S. (2021). Sustainability of Morocco's groundwater resources in response to natural and anthropogenic forces. *J. Hydrol.*, 603(Part A), article no. 126866.
- Arfib, B. and de Marsily, G. (2004). Modeling the salinity of an inland coastal brackish karstic spring with a conduit-matrix model. *Water Resour. Res.*, 40(11), article no. W11506. 1–10.
- Arfib, B., de Marsily, G., and Ganoulis, J. (2002). Les sources karstiques côtières en Méditerranée: étude des mécanismes de pollution saline de l'Almyros d'Héraklion (Crète), observations et modélisation. *Bull. Soc. Géol. Fr.*, 173(3), 245–253.
- Arfib, B. and Douchet, M. (2011). État des connaissances hydrogéologiques et spéléo-plongée sur les rivières souterraines sous-marines de Port-Miou et du Bestouan (Cassis, France). In *Proceedings of the 9th Conference on Limestone Hydrogeology*, pages 25–28, Besançon, France.
- Bakalowicz, M. (2018). Coastal karst groundwater in the Mediterranean: a resource to be preferably exploited onshore, not from karst submarine springs. *Geosciences*, 8, article no. 258. 1–12.
- Bakalowicz, M. and El-Hajj, A. (2012). Les sources karstiques sous-marines de Méditerranée sont-elles des ressources non conventionnelles exploitables? In *18<sup>èmes</sup> journées techniques du Comité Français d'Hydrogéologie, CFH-AIH. Ressources et gestion des aquifères littoraux, Cassis*, pages 91–98.
- Bakalowicz, M., Fleury, P., Dörfliker, N., and Seidel, J. L. (2003a). Coastal karst aquifers in Mediterranean regions. 1. A valuable ground water resource in complex aquifers. In *TIAC'03*, volume 1, pages 125–128. IGME, Alicante.
- Bakalowicz, M., Fleury, P., Jouvencel, B., Promé, J. J., Becker, P., Carlin, T., Dörfliker, N., Seidel, J. L., and Sergent, P. (2003b). Coastal karst aquifers in Mediterranean regions. 2. A methodology for exploring, exploiting and monitoring submarine springs. In *TIAC'03*, volume 1, pages 673–680. IGME, Alicante.
- Bayari, S. C., Ozyurt, N., Oztan, M., Bastanlar, Y., Varinlioglu, G., Koyuncu, H., Ulkenli, H., and Hamarat, S. (2011). Submarine and coastal karstic groundwater discharges along the southwestern Mediterranean coast of Turkey. *Hydrogeol. J.*, 19, 399–414.
- Bejannin, S., Tamborski, J. J., van Beek, P., Souhaut, M., Stieglitz, T., Radakovitch, O., Claude, C., Conan, P., Pujo-Pay, M., Crispi, O., Le Roy, E., and Estourne, C. (2020). Nutrient fluxes associated with submarine groundwater discharge from karstic coastal aquifers (Côte Bleue, French Mediterranean Coastline). *Front. Environ. Sci.*, 7, article no. 205.
- Blavoux, B., Gilli, E., and Rousset, C. (2004). Alimentation et origine de la salinité de la source sous-marine de Port-Miou (Marseille-Cassis). Principale émergence d'un réseau karstique hérité du Messinien. *C. R. Geosci.*, 336(6), 523–533.
- Breznik, M. and Steinman, F. (2011). Desalination of coastal karst springs by hydro-geologic, hydro-technical and adaptable methods. In Schorr,

- M., editor, *Desalination, Trends and Technologies*, pages 41–70. IntechOpen, London.
- Buongiorno Nardelli, B., Budillon, F., Watteaux, R., Ciccone, F., Conforti, A., De Falco, G., Di Martino, G., Innangi, S., Tonielli, R., and Iudicone, D. (2017). Pockmark morphology and turbulent buoyant plumes at a submarine spring. *Cont. Shelf Res.*, 148, 19–36.
- Cavalera, T. (2007). *Etude du fonctionnement et du bassin d'alimentation de la source sous-marine de Port Miou (Cassis, Bouches-du-rhone). Approche multicritère*. PhD thesis, Université de Provence. 403 pages.
- de Marsigli, L. F. (1725). *Histoire physique de la mer*. Aux depens de la Compagnie. Translation of: Brieve ristretto del Saggio fisico intorno alla storia del mare, 1711, Venice, e-book: <https://play.google.com/store/books/details?id=RpsSxRY4bRcC&rdid=book-RpsSxRY4bRcC&rdot=1>.
- de Serres, M. (1849). In *De la source de l'Abysse et de la Fontaine d'Embressac, dans les environs de Balaruc (Hérault)*, Mémoires de la section des Sciences, pages 223–234. Académie des Sciences et Lettres de Montpellier, Montpellier.
- Deharveng, L., Stoch, F., Gibert, J., et al. (2009). Groundwater biodiversity in Europe. *Freshw. Biol.*, 54, 709–726.
- Dörfliger, N., Fleury, P., and Ladouche, B. (2008). Inverse modelling approach to allogenic karst system characterisation. *Ground Water*, 47(3), 414–426.
- El-Hajj, A. (2008). *Laquifère carbonaté karstique de Chekka (Liban) et ses exutoires sous-marins. Caractéristiques hydrogéologiques, fonctionnement et modélisation*. PhD thesis, Univ. Montpellier and USJ Beyrouth. 261 pages.
- El Hakim, M. and Bakalowicz, M. (2007). Significance and origin of very large regulating power of some karst aquifers in the Middle East. Implication on karst aquifer classification. *J. Hydrol.*, 333, 329–339.
- Fleury, P. (2005). *Sources sous-marines et aquifères côtiers méditerranéens. Fonctionnement et caractérisation*. PhD thesis, Univ. Paris 6. 286 pages.
- Fleury, P., Bakalowicz, M., and Becker, P. (2007a). Caractérisation d'un système karstique à exutoire sous-marin, exemple de la Mortola (Italie). *C. R. Geosci.*, 339, 407–417.
- Fleury, P., Bakalowicz, M., and de Marsily, G. (2007b). Submarine springs and coastal karst aquifers: a review. *J. Hydrol.*, 339, 79–92.
- Fleury, P., Bakalowicz, M., de Marsily, G., and Cortes, J. M. (2008). Functioning of a coastal karstic system with a submarine outlet, in southern Spain. *Hydrogeol. J.*, 16(1), 75–85.
- Gèze, B. (1938). Les sources mystérieuses des monts de la Gardiole (Hérault). *Géographie*, 69(4), 193–208.
- Goldscheider, N., Chen, Z., Auler, A. S., Bakalowicz, M., Broda, S., Drew, D., Hartmann, J., Jiang, G., Moosdorf, N., Stevanovic, Z., and Veni, G. (2020). Global distribution of carbonate rocks and karst water resources. *Hydrogeol. J.*, 28(5), 1–17.
- Jou-Claus, S., Folch, A., and Garcia-Orellana, J. (2021). Applicability of Landsat 8 thermal infrared sensor for identifying submarine groundwater discharge springs in the Mediterranean Sea basin. *Hydrol. Earth Syst. Sci.*, 25, 4789–4805.
- Klammler, H., Jawitz, J.-W., Annable, M.-D., Yaquian, J.-A., Hatfield, K., and Burger, P. (2020). Decadal scale recharge-discharge time lags from aquifer freshwater-saltwater interactions. *J. Hydrol.*, 582, article no. 124514.
- Kuscer, I. and Kuscer, D. (1962). Observation on brackish karst sources and sea swallow-holes on the Yugoslav coast. In *Mémoires A. I. H., Réunion d'Athènes*, volume 5, pages 344–353. BRGM, Orléans.
- Ladouche, B., Lamotte, C., Hemelsdael, R., Petre, M.-A., Dewandel, B., Leonardi, V., Seidel, J.-L., and Séranne, M. (2018). DEM'Eaux Thau - Synthèse et valorisation préliminaire des données sur l'hydrosystème de Thau (34) - Livrable L1. Rapport final. BRGM/RP-68483-FR, 313 p.
- Lofi, J., Gorini, C., Berné, S., Clauzon, G., Tadeu Dos Reis, A., Ryan, W. B. F., and Steckler, M. S. (2005). Erosional processes and paleo-environmental changes in the Western Gulf of Lions (SW France) during the Messinian Salinity Crisis. *Mar. Geol.*, 217(1–2), 1–30.
- Luijendijk, E., Gleeson, T., and Moosdorf, N. (2020). Fresh groundwater discharge insignificant for the world's oceans but important for coastal ecosystems. *Nat. Commun.*, 11, article no. 1260.
- Mammola, S., Amorim, I. R., Bichuette, M. E., Borges, P. A. V., Cheeptham, N., Cooper, S. J. B., Culver, D. C., Deharveng, L., Eme, D., Lopes Ferreira, R., Fiser, C., Fiser, Z., Fong, D. W., Griebler, C., Jeffery, W. R., Jugovic, J., Kowalko, J. E., Lilley, T. M.,

- Malard, F., Manenti, R., Martínez, A., Meierhofer, M. B., Niemiller, M. L., Northup, D. E., Pellegrini, T. G., Pipan, T., Protas, M., Reboleira, A. S. P. S., Venarsky, M. P., Wynne, J. J., Zagamajster, M., and Cardoso, P. (2020). Fundamental research questions in subterranean biology. *Biol. Rev.*, 95(6), 1855–1872.
- Maréchal, J. C., Vestier, A., Jourde, H., and Dörfliger, N. (2013). L'hydrosystème du Lez: une gestion active pour un karst à enjeux. *Karstologia*, 62, 1–6.
- Mijatovic, B. (2007). The groundwater discharge in the Mediterranean karst coastal zones and freshwater tapping: set problems and adopted solutions. Case studies. *Environ. Geol.*, 51, 737–742.
- Moulard, L., Mijatovic, B., Kareh, R., and Massaad, B. (1965). Exploitation d'une nappe karstique captive à exutoires sous-marins. Problèmes posés et solution adoptée côte libanaise. In *Conference on Ground Water Resources*, pages 237–250. UNESCO, Paris.
- Perriquet, M., Léonardi, V., Tiernan, H., and Jourde, H. (2014). Saltwater wedge variation in a non-anthropogenic coastal karst aquifer influenced by a strong tidal range (Burren, Ireland). *J. Hydrol.*, 519(Part B), 2350–2365.
- Pétre, M. A., Ladouche, B., Seidel, J. L., Hemelsdaël, R., de Montety, V., Batiot-Guilhe, C., and Lamotte, C. (2020). Hydraulic and geochemical impact of occasional saltwater intrusions through a submarine spring in a karst and thermal aquifer (Balaruc peninsula near Montpellier, France). *Hydrol. Earth Syst. Sci.*, 24(11), 5655–5672.
- Rodríguez Estrella, T. (2000). Transformaciones físicas, químicas y biológicas, por la acción de las aguas del trasvase Tajo-Segura, en el unita hidrogeológica del Campo de Cartagena y en su vecino Mar Menor (Murcia, España). *Hydrogéologie*, 2, 23–37.
- Romero Macias, E., Gonzalez Martinez, A., and Garrido Morillo, R. (2000). Risks of detrital aquifer contamination in the coastal area of Huelva Province (Western Andalusia, Spain). *Hydrogéologie*, 3, 39–46.
- Rouch, R. and Danielopol, D. L. (1987). L'origine de la faune aquatique souterraine, entre le paradigme du refuge et le modèle de la colonisation active. *Stylogia*, 3(4), 345–372.
- Rouchy, J. M. and Saint-Martin, J. P. (1992). Late Miocene events in the Mediterranean as recorded by carbonate-evaporite relations. *Geology*, 20, 629–632.
- Rouchy, J. M., Suc, J. P., Ferrandini, J., and Ferrandini, M. (2006). The Messinian Salinity Crisis revisited. *Sediment. Geol.*, 188–189, 1–8.
- Santos, I. R., Chen, X., Lecher, A. L., Sawyer, A. H., Moosdorf, N., Rodellas, V., Tamborski, J., Cho, H.-M., Dimova, N. T., Sugimoto, R., Bonaglia, S., Li, H., Hajati, M. C., and Li, L. (2021). Submarine groundwater discharge impacts on coastal nutrient biogeochemistry. *Nat. Rev. Earth Environ.*, 2(5), 307–323.
- Schubert, M., Scholten, J., Schmidt, A., Comanducci, J. E., Pham, M. K., Mallast, U., and Knoeller, K. (2014). Submarine groundwater discharge at a single spot location: evaluation of different detection approaches. *Water*, 6, 584–601.
- Sivelle, V., Jourde, H., Bittner, D., Mazzilli, N., and Trambly, Y. (2021). Assessment of the relative impacts of climate changes and anthropogenic forcing on spring discharge of a Mediterranean karst system. *J. Hydrol.*, 598, article no. 126396.
- Stefanon, A. (1972). Capture and exploration of submarine spring. In *Proceedings of the Oceanology International Conference*, volume 72, pages 427–430, Brighton, England.
- Tamborski, J., van Beek, P., Conan, P., Pujo-Pay, M., Odobel, C., Ghiglione, J. F., Seidel, J. L., Arfib, B., Diego Feliu, M., Garcia-Orellana, J., Szafran, A., and Souhaut, M. (2020). Submarine karstic springs as a source of nutrients and bioactive trace metals for the oligotrophic Northwest Mediterranean Sea. *Sci. Total Environ.*, 732, article no. 139106.
- Tardieu, B. and Potié, L. (2015). Freshwater submarine springs: role of a dam in submerged karst. Investigations, measurement and works in subterranean rivers of Cassis, France, from 1964 to 2013. In Lollino, G., Arattano, M., Rinaldi, M., Giustolisi, O., Marechal, J. C., and Grant, G., editors, *Engineering Geology for Society and Territory*, volume 3. Springer, Cham.
- Toulmond, A. (2006). *Brève histoire illustrée de l'océanographie : les scientifiques sur les chemins des mers*. Publ. Station de Roscoff, Roscoff, <https://www.sb-roscoff.fr/sites/www.sb-roscoff.fr/files/documents/station-biologique-roscoff-breve-histoire-illustree-de-l-oceanographie-3078.pdf>.
- UNESCO (2004). *Submarine Groundwater Discharge. Management Implications, Measurements and Effects*. IHP-VI, Series on Groundwater No. 5, IOC

Manuals and Guides No. 44. UNESCO, Paris.  
Zektser, I. S., Dzhamalov, R. G., and Safronova, T. I.  
(1983). Role of submarine groundwater discharge

in the water balance of Australia. *IAHS-AISH Publ.*,  
1, 209–219.





---

Research article

Geo-hydrological Data & Models

# Assessing the salinization mechanisms of coastal brackish springs

Esteban Sanz<sup>a</sup>, Jesús Carrera<sup>\*,\*,b,c</sup>, Carlos Ayora<sup>\*,b,c</sup>, Alfredo Barón<sup>d</sup> and Concha González<sup>d</sup>

<sup>a</sup> Vår Energi ASA, Stavanger, Norway

<sup>b</sup> Institute of Environmental Assessment and Water Research, IDAEA-CSIC, Barcelona, Spain

<sup>c</sup> Associated Unit: Hydrogeology group (UPC-CSIC), Spain

<sup>d</sup> Water Authority, Balearic Government, Palma de Mallorca, Spain

*E-mails:* esteban.sanz@live.com (E. Sanz), [jesus.carrera.ramirez@gmail.com](mailto:jesus.carrera.ramirez@gmail.com), [jcrgeo@idaea.csic.es](mailto:jcrgeo@idaea.csic.es) (J. Carrera), [cayora1@gmail.com](mailto:cayora1@gmail.com) (C. Ayora), [baron.a@telefonica.net](mailto:baron.a@telefonica.net) (A. Barón)

**Abstract.** Brackish springs have been extensively reported in Mediterranean coastal carbonate formations. The phenomenon is puzzling because these springs may discharge high flow rates with significant salinities at elevations of several meters above sea level. Although these springs have been studied for millennia (since the ancient Greeks), controversy persists. In essence, they are typically assumed to consist of a karstic conduit bringing fresh groundwater that meets a saltwater conduit at a branching point, where the brackish mixture flows up to the spring mouth. Here, we analyze the hydraulics of the system for two simple cases, depending on whether seawater flow occurs through an open conduit or a porous aquifer. To this end, we derive the equations governing variable-density turbulent flow through the conduits. These turn out to be similar to the traditional ones, except that (1) Bernoulli's head (energy per unit weight) is substituted by the energy per unit volume, and (2) this energy is not a potential, as a flow path dependent term needs to be added. We solve these equations to obtain the characteristic functions relating fresh to seawater discharge depending on spring concentration. These functions are specific to every brackish spring and representative of the karst system and salinization mechanism. They allow us to show that the spring salinity is mostly controlled by the weight of the water column flowing towards the spring mouth (for low flow rates) and by energy dissipation (for high flow rates). When concentration and flow rate data are available, it is possible to characterize: (1) the resistance of the upwards conduit from the response at high flow rate; (2) the depth of the branching point from the concentration at low discharges; and (3) the hydraulic resistance of the seawater conduit from the slope of concentration vs. flow rate at low spring discharges. These patterns are compared to field data from three different springs, which yields insights on the conceptual model governing every particular spring.

**Keywords.** Brackish spring, Salinization, Conduit branching, S'Almadrava, Modeling, Coastal springs.

*Manuscript received 18 April 2022, revised 31 January 2023, accepted 21 February 2023.*

---

\* Corresponding author.

## 1. Introduction

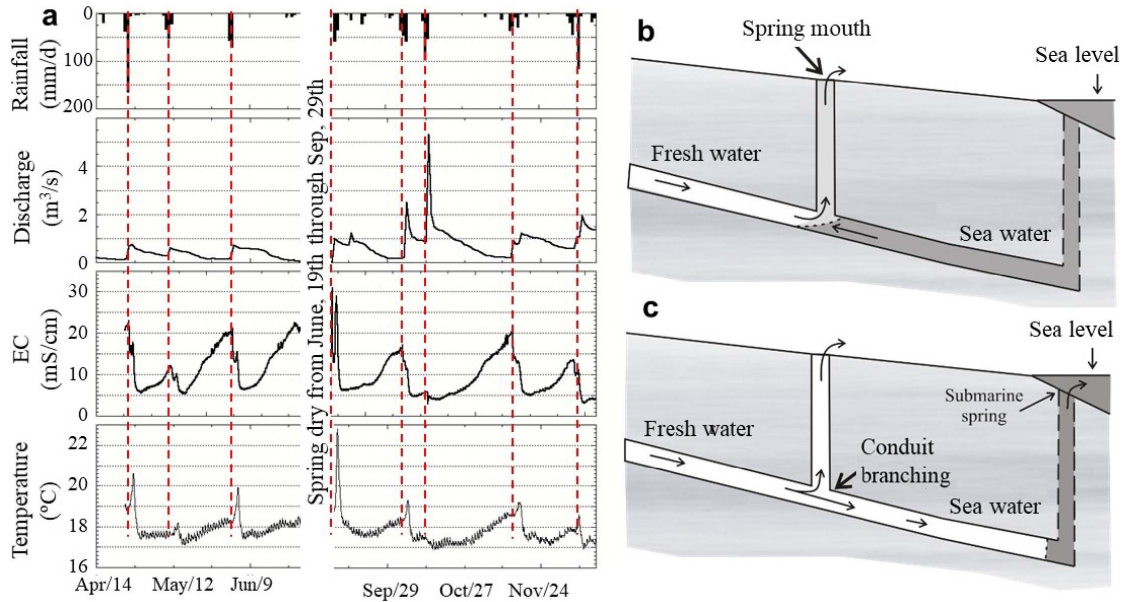
Ghislain de Marsily is best known for his contributions to the identification of aquifer heterogeneity [Ackerer *et al.*, 2023], his book [de Marsily, 1986], or his recent works on the global relevance of water [e.g., de Marsily, 2009], to list a few of his contributions. Less known is his work on seawater intrusion and specifically on the puzzle of coastal springs that discharge water containing a large fraction of seawater above sea level [Arfib and de Marsily, 2004, Arfib *et al.*, 2007], which motivates this paper.

Advances in science have overturned the ancient Greeks belief that seawater, consisting of water and earth (two of the five basic elements), flowed inland where its earthen component slowly returned to the ground (its natural state) giving rise to fresh groundwater. It is believed that this concept originated in the Island of Cephallonia (Greece) as a result of the simultaneous observation of a continuous flow of seawater into a sink on the west side of the island, and frequent clouds over the island's mountains [Stringfield and LeGrand, 1969]. The fact that seawater could be seen entering the ground, together with the observation that water got fresher (and lighter) uphill, suggested that by the time it reached the mountain top, it had turned so light as to become air (clouds), "clearly" showing that the hydrological cycle was just the opposite of what we now know. The particular phenomenon of seawater inflow in Cephallonia triggered a scientific controversy that would last for many centuries. A number of imaginative but often physically unacceptable hypotheses had been advanced to account for the phenomenon [Fouqué, 1867, Wiebel, 1873, in Breznik, 1973, Crosby and Crosby, 1896 and Fuller, 1907] until Maurin and Zoetl [1967] demonstrated with a colouring test that the inflowing seawater returned at brackish karst springs located on the opposite side of the island at a distance of 15 km.

Brackish springs are relatively frequent in coastal carbonate formations and their existence has been extensively reported, especially in the Mediterranean area [Payne *et al.*, 1978, reviews of Breznik, 1973, Leontiadis *et al.*, 1988, Fleury *et al.*, 2007a, Acikel and Ekmekci, 2018]. In fact, more than 300 springs have been identified in the coast of former Yugoslavia alone [Bonacci and Roje-Bonacci, 1997]. They essentially consist of inland or submarine karst outlets

discharging waters with flow-dependent salinity. They may discharge high flow rates with significant salinities (presumably coming from the sea) at elevations of several meters above sea level, which reveals the especial complexity of their hydraulic and salinization mechanisms (see, e.g., records of rainfall, flow rate, salinity, and temperature at S'Almadrava spring, which discharges at 8 m.a.s.l. in Majorca, Spain, Figure 1a). Although these phenomena have been studied for many years, controversy persists.

The occurrence of a well-developed deep karst system in a seawater intrusion zone appears to be the key factor in the formation of brackish springs. This karst could be formed by the combination of tectonic activity and variations in the sea level over geologic time scales [Fleury *et al.*, 2007b]. Moreover, mixing of fresh and seawater may lead to carbonate dissolution [Rezaei *et al.*, 2005], further broadening the conduit section at depth and modifying the seawater intrusion pattern. Under these conditions, a plausible scenario consists of a karst conduit open directly to the sea, allowing mixing at a deep conduit branching [Breznik, 1973]. The conduit branching is the junction of three conduits: one from inland with freshwater, another one connected to the sea, and the third one with brackish water leading to the spring mouth (Figure 1b). Such conduit branching has never been directly observed, and the uncertainties associated with its geometry and dimensions have led authors to propose different physical mechanisms to explain salinization at heights above sea level. Proposed salinization mechanisms can be divided into two groups [Stringfield and LeGrand, 1969]: salinization due to hydrodynamic effects and salinization due to the greater density of seawater. The first group involves physical conditions of flow in tubular openings or other channels which serve as Venturi tubes. This requires a narrowing of the freshwater conduit at the intersection with the seawater one, allowing the freshwater flowing through to cause a depression, of the order of  $v^2/2g$  that sucks in seawater. This mechanism has been proposed for springs showing a rising or steady salinity with increasing discharge flow, such as the Slanac spring (Croatia) [Breznik, 1973] and the Makaria spring (Greece) [Maramathas *et al.*, 2006]. However, considering the salinity and elevation of the springs, these situations would require very high groundwater velocities and particular conduit networking morphologies that would be dif-



**Figure 1.** Left, field measurements of rainfall at the recharge area, discharge, electrical conductivity (EC) and temperature of S'Almadrava spring at 8 m.a.s.l in Majorca (Spain) from April, 14th through December, 26th, 1996 year (note that the spring was dry during the summer). Right, conduit branching conceptual model for brackish springs discharging a fraction of seawater above sea level. When the freshwater flow is small (a), the density of mixed water in the vertical branch is smaller than that of seawater and brackish water discharges at the spring mouth. When the freshwater flow is large (b), it intrudes into the conduit connected to the sea, thus creating a submarine spring on the sea floor. Flow is turbulent if all conduits are open (T-T, for Turbulent–Turbulent, model).

difficult to find in nature or to maintain without erosion, not to mention that mixing fresh and seawater leads to aggressive conditions [i.e., limestone dissolution, Sanford and Konikov, 1989, Rezaei *et al.*, 2005, Sanz *et al.*, 2011]. The Venturi effect was also conjectured by McCullough and Land [1992] to explain the inflow of seawater into submarine karst conduits in response to tides and waves. However, the inflow is not in phase with sea level fluctuations, which makes the Venturi effect unlikely. Instead, a hydromechanical effect [Guarracino *et al.*, 2012] explains why the maximum flow rate coincides with the mid-point of the rising tide.

The second group of proposed mechanisms assumes that seawater is sucked at the conduit branching when the hydrostatic pressure in the upwards brackish water conduit is less than that in the seawater tube (that is, when the weight of the brackish mixture up to the spring mouth is less than the weight of seawater up to the sea level). This situ-

ation becomes possible if the depth of the branching point is large enough to balance the elevation of the spring mouth with the reduced density of brackish water as compared to seawater. The required depth of the branching point is of several hundred meters, which adds mystery to these peculiar springs. Bakalowicz [2014] attributes deep karst conduits to the Messinian salinity crisis some 5.6 million years ago, when the Mediterranean Sea became isolated from the Atlantic Ocean. This caused the sea-level to drop by at least 1500 m [Ryan, 1976], thus lowering the base discharge level and the elevation of karst outlets at large depths below current sea level. In fact, not only the low base level, but also the increased salinity contribute to the development of deep karst conduits. Furthermore, as discussed above, mixing of fresh and salt water leads to calcite subsaturation even when both end members are in equilibrium with calcite and the effect increases with the salinity contrast. Enhanced dissolution coupled to the large

drops in sea level elevation during the Messinian salinity crisis may explain why these springs are well documented in Mediterranean shores, but rare elsewhere. Kohout [1965, in Kohout *et al.*, 1977] describe saline springs in Florida. Their observations have been attributed to a thermal convection effect (labelled “Kohout effect”): cold (i.e., dense) seawater enters the aquifer, where it warms up, thus becoming lighter, so that it can rise to a significant height. The relevance of this effect increases with the temperature contrast between the cold and hot waters and with the depth of the cold sea. Both factors are large in Florida [Hughes *et al.*, 2007], but relatively small in the Mediterranean Sea, where we have estimated that the Kohout effect could account for only some 25 cm rise above sea level.

A singular feature of Mediterranean brackish springs is the variation salinity with spring discharge, which has been attributed to variations of freshwater pressure with changes in freshwater flow from the aquifer [Gjurasin, 1943, and Kuscer, 1950, in Breznik, 1973]. This mechanism has been widely accepted for springs that display an inverse relation of flow discharge with respect to salinity, i.e., discharging freshwater during high flows and, below a certain flow-limiting value, increasingly saline waters with decreasing discharge. Examples include Almyros of Heraklion spring (Crete, Greece) [Breznik, 1973] Blaz spring (Croatia) [Bonacci and Roje-Bonacci, 1997], Pantan [Breznik, 1973] and S'Almadrava spring [Sanz *et al.*, 2003].

In some cases, geological constraints suggest that seawater contamination of freshwater flowing through a karst conduit would not be necessarily related to a conduit network connected directly to the sea, but rather to a diffusive seawater intrusion from the porous matrix surrounding the conduit [Arfib and de Marsily, 2004]. This situation may occur when a freshwater conduit crosses a saline-intruded fissured matrix zone. Arfib and Ganoulis [2004] performed laboratory experiments demonstrating that a considerable mass exchange may happen under these conditions. Vertical conduits, such as the spring conduit in Figure 1, connected to deep horizontal conduits, such as the fresh and sea water conduits, have been described by Lofi *et al.* [2012], who link these conduits to karstic activity during the Messinian crisis. While the model of Figure 1 makes geologic sense, it is hard to ascertain whether the

sea water conduit is open to the sea or connected through a permeable porous medium. We conjecture that some insight into the most dominant mechanism (open conduit or diffusive connection) in a specific case can be gained from modeling.

There have been attempts to apply numerical models to brackish springs. They have been motivated by the fact that the discharge of those springs represents a precious resource in areas with limited water resources. Modeling proved to be an important tool to test different options of spring development proposed in the last decades [Breznik, 1973, 1988, Leontiadis *et al.*, 1988, Bonacci and Roje-Bonacci, 1997, Cardoso, 1997, Sanz *et al.*, 2003]. Some authors have employed nonlinear analysis or inverse modeling to calculate unit hydrographs and impulse responses from rainfall data in the recharge area [Lambrakis *et al.*, 2000, Pinault *et al.*, 2004]. This type of models, extensively used for freshwater karst springs, use box reservoirs to represent the relationship between an input and an output signal. However, these methods do not consider the physical processes and mechanisms controlling the spring functioning. Maramathas *et al.* [2003] adopted a different approach based on the mass and energy balance on a hydrodynamic analog, which included three reservoirs flowing from tubes lying adjacent to the spring. Two reservoirs represent the karstic system (two karst subsystems with different depletion period), and the third one emulates the sea. This model assumes the existence of a conduit branching with a conduit open directly to the sea (although this is not simulated), and computes discharge and chloride concentration of the spring using rainfall data as model input. Although the model was successfully applied to the Almyros spring of Heraklion, mixing at the conduit branching and the variable-density turbulent flow in conduits were not considered. In fact, the equations governing variable-density turbulent flow have never been addressed before. We believe that these equations are crucial for a full understanding of the physics of spring functioning. In contrast to this work, Arfib and de Marsily [2004] applied a different conceptual model to the same spring. They assumed the existence of a single freshwater conduit surrounded by a saline porous matrix, where salinization of freshwater in the conduit is a consequence of saline flux from the matrix. This approach considered constant-density turbulent flow in the

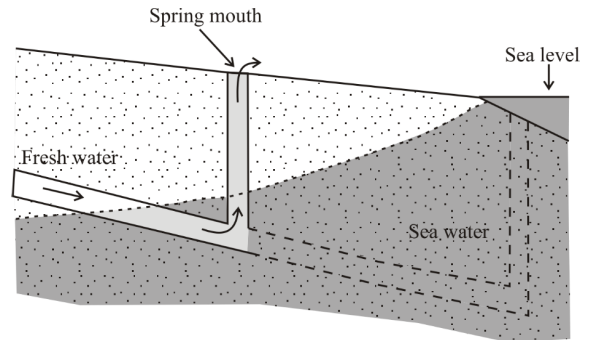
conduits and mass exchange ratio driven by the head difference between the conduit and the matrix. Surprisingly, these two studies at Almyros spring provided similar results even though they were based on rather different conceptual models and approaches. Therefore, a detailed description of these conceptual models for brackish springs and a direct comparison among them should be addressed.

The objectives of this study are (1) to derive the equations governing density-dependent turbulent flow in open conduits, (2) to describe the salinization mechanisms of inland brackish springs presenting a connection with the sea through an open karst conduit or a diffusive seawater connection, and (3) to compare the spring discharge and concentration response for these two salinization mechanisms.

## 2. Theory

### 2.1. Conceptual models of brackish springs

Any conceptual model aimed at explaining brackish springs must include a well developed deep karst system and identify the dominant mechanism of seawater contamination, i.e., the way in which seawater intrudes into the aquifer and mixes with freshwater. Here we analyze two conceptual models that synthesize the features of the most plausible natural systems (Figures 1 and 2). Both conceptual models can explain inverse relations of concentration and discharge of inland springs but apply different salinization mechanisms. The first conceptual model, Turbulent–Turbulent (T–T), assumes that groundwater flows only through a network of conduits and that the seawater contamination occurs at a deep branching point through a conduit of undefined shape connected to the sea (Figure 1). In this environment, ground water follows the hydraulic laws for pipes, resulting on laminar or turbulent flow depending on the velocity and properties of the fluid, and the shape and extent of the conduit section [Chadwick and Morfett, 1998]. This problem can be simplified into a mass and energy balance at the conduit branching. Mixing of fresh and saltwater occurs when the interface is placed at the conduit branching (Figure 1a). If the freshwater flow is high enough, freshwater will intrude the seawards conduit and create a submarine spring of, eventually, freshwater (Figure 1b). By contrast, if freshwater flow is very low (or eventually



**Figure 2.** Scheme of the Turbulent–Porous (T–P) conceptual model for brackish springs that considers salinization from seawater by diffuse exchange between the conduit and the surrounding porous matrix. Dotted areas represent the porous/fissured matrix. Water salinity in the conduit and the matrix increases from white to dark grey.

zero), seawater will intrude into the freshwater conduit thus increasing the seawater contamination in the aquifer.

The second conceptual model for spring salinization, which we term Turbulent–Porous (T–P), is a simplification of the model described by Arfib and de Marsily [2004]. They assumed that the freshwater conduit is embedded in a fissured matrix intruded by seawater. For the sake of simplicity, we treat this fissured matrix as a conduit where water flows according to Darcy's Law (Figure 2).

Given the complexity of karst systems, the schemes of Figures 1 and 2 should be viewed as modeling simplifications. Combinations of multiple branching points contaminated at different depths are likely to occur in natural systems. Nevertheless, the simplicity of these two models should facilitate understanding of the physics controlling the phenomenon, which is the aim of our work.

### 2.2. Governing equations

The conceptual models considered in this study consist both of a mixing place (i.e., conduit branching) in which a conduit connected to the sea and another conduit from the aquifer inland, join a third vertical conduit with mixed water leading to the spring mouth. Depending on the conceptual model, the

seawater conduit is assumed to be an open conduit (T-T case, Figure 1) or a porous/fissured medium (T-P case, Figure 2). In both cases flow is governed by mass and either momentum or energy conservation. We have found no equations for variable density pipe flow in the literature. Therefore we derived them below.

We consider a conduit with open area  $A(l)$  ( $A = A'\phi$ , where  $A'$  is the total area and  $\phi$  porosity, in the case of a porous medium), where  $l$  is the length along the conduit axis. Fluid mass conservation is expressed as:

$$\frac{\partial \rho A}{\partial t} = -\frac{\partial \rho Av}{\partial l} \quad (1)$$

where  $\rho(l, t)$  is density and  $v(l, t)$  is velocity.

Momentum conservation can be expressed in Lagrangian coordinates, in which case it expresses Newton's second law, or in Eulerian coordinates. We adopt the second option, which implies that the variation of momentum ( $\partial \rho Av / \partial t$ ) is equal to the inflow minus outflow of momentum per unit length of conduit ( $-\partial \rho Av^2 / \partial l$ ) plus the forces acting on the fluid (pressure, weight, and conduit resistance), which, expressed per unit length equals  $-\partial PA / \partial l - \rho Ag \cos \theta - f_p$ , where  $\cos \theta = \partial z / \partial l$ ,  $z$  is elevation, and  $f_p$  represents the component along  $l$  of the forces exerted by the conduit walls over the fluid. That is,

$$\frac{\partial \rho Av}{\partial t} = -\frac{\partial \rho Av^2}{\partial l} - \frac{\partial PA}{\partial l} - \rho Ag \cos \theta - f_p. \quad (2)$$

This equation can be used, together with Equation (1), for solving the problem. It can also be written in terms of energy conservation, which may be more intuitive. To this end, we expand the time derivative in Equation (2), use Equation (1) to eliminate the resulting  $\partial \rho A / \partial t$  and perform some minor algebraic manipulations to obtain

$$\rho A \frac{\partial v}{\partial t} = -\left( \rho A \frac{\partial v^2 / 2}{\partial l} + \frac{\partial PA}{\partial l} + \rho Ag \cos \theta + f_p \right). \quad (3)$$

Multiplying Equation (3) by  $v$  and adding Equation (1) multiplied by  $v^2 / 2$ , yields the energy conservation equation for variable density flow in a pipe:

$$\frac{\partial \rho Av^2 / 2}{\partial t} + \frac{\partial \rho Av^3 / 2}{\partial l} + v \frac{\partial PA}{\partial l} + \rho Av g \cos \theta + f_p v = 0. \quad (4)$$

This equation expresses energy conservation explicitly, but it is still not easy to apply because of the

pressure forces exerted by the conduit walls. The normal component of these cancels if  $A$  is assumed constant. Moreover, measurements are rarely more frequent than hourly so that pressure waves can be ignored and the fluid assumed to be incompressible. Therefore, the flow rate ( $Q = Av$ ) is constant along the conduit. With these simplifications, Equation (4) becomes

$$-\frac{1}{Q} \frac{\partial \rho Av^2 / 2}{\partial t} = \frac{\partial}{\partial l} \left( \rho \frac{v^2}{2} + P + \rho g z \right) + \frac{f_p}{A}. \quad (5)$$

Equation (5) resembles Bernoulli's equation except for the fact that  $\rho$  varies in space and time in response to variations in the proportion of seawater. As a result,  $\rho$  cannot be factored out and energy conservation cannot be written in terms of head (energy per unit weight of fluid), but in terms of energy per unit volume. This can be viewed as  $e = \rho gh$ . Another difficulty is that  $e$  is no longer a potential. That is, flow is no longer given solely by differences in local  $e$ , but it depends also on the values of  $\rho$  along the path. To illustrate this result, assume steady-state and integrate Equation (5) along the conduit, between points 1 and 2, which yields:

$$(\rho_2 - \rho_1) \frac{v^2}{2} + P_2 - P_1 + \bar{\rho} g (z_2 - z_1) + \frac{\bar{f}_p L}{A} = 0 \quad (6)$$

where the overbar stands for spatial average and  $L$  is the conduit length  $\bar{\rho}$  is not the average along the conduit length, but along the vertical direction:  $\bar{\rho} = \int \rho dz / (z_2 - z_1)$ . If  $\rho$  is constant, then Equation (6) is the traditional expression of energy conservation ( $e_2$  equals  $e_1$  minus dissipated energy). But the fact that  $e$  depends on the average density along the (vertical component of the) path hinders its use as state variable. Instead, we will use  $H = \rho v^2 / 2 + P$  (energy per unit volume relative to the point elevation) as state variable. Note that, if density is constant, then the traditional Bernoulli head would be  $h = e / \rho g = z + H / \rho g$ .

The expression of  $\bar{f}_p$  depends on whether the conduit is open (turbulent flow) or porous (Darcy flow). In the case of the open conduit, we have used Manning's equation:

$$\frac{\bar{f}_p}{A} = \frac{n^2 \bar{\rho} g v^2}{R_H^{4/3}} \quad (7)$$

where  $n$  is Manning's coefficient and  $R_H$  is the hydraulic radius of the conduit (ratio of  $A$  to wet

perimeter). We have assumed constant cross-sectional area  $A$  (and also  $n$ ). If variable, we would have had to integrate Equation (6), which usually leads, in constant density practice, to an increased  $\overline{f_p}$ , still proportional to  $Q^2$ . In the case of porous medium, Newton's Law of viscosity suggest that the shear stress will be proportional to velocity:

$$\overline{f_p} = \frac{A\mu v}{k} \quad (8)$$

where  $\mu$  is viscosity (assumed constant and equal to  $1 \times 10^{-3}$  m·s/kg) and  $k$  intrinsic permeability. This, together with neglecting  $v^2$  (of the order of  $\mu$ Pa for normal groundwater velocities), yields Darcy's law:

$$v = \frac{k}{\mu} \left( \frac{P_1 - P_2}{L} + \frac{\overline{\rho}g(z_1 - z_2)}{L} \right). \quad (9)$$

The fact that  $\rho$  varies in space and time forces us to solve the salt mass conservation equation, which we have written (neglecting dispersion) as:

$$\frac{\partial c}{\partial t} = -v \frac{\partial c}{\partial l} \quad (10)$$

where  $c$  is salt mass fraction. Density depends on concentration as:

$$\rho = \rho_f \exp(\alpha(c - c_f)) \quad (11)$$

where  $\rho_f$  is the density of fresh water (1000 kg/m<sup>3</sup>) and  $\alpha$  equals 0.69167 for mixtures of fresh and seawater ( $c_s = 3.57 \times 10^{-2}$ ). It should be noted that having assumed the area to be constant and the fluid to be incompressible, the fluid mass conservation is equivalent to solute mass conservation. This can be easily checked by plugging Equation (11) into Equation (1), which leads to Equation (10).

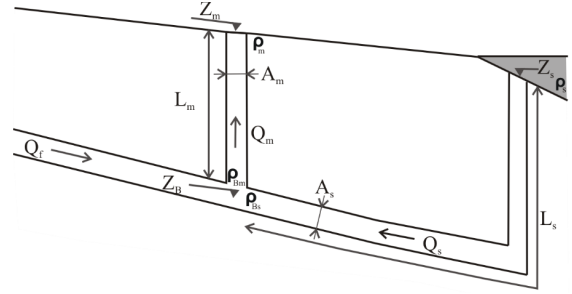
Solution of these equations requires specifying boundary conditions and continuity at inner nodes (branching point). Pressure is specified as equal to atmospheric pressure (zero relative pressure) at the spring mouth. Energy (per unit volume),  $H_s = \rho_s v_s^2/2 + P_s$ , is specified at the seawater entry point, which can be derived from

$$\rho_s g e_s = \rho_s g z_s + \frac{\rho_s v_s^2}{2} + P_s = 0 \quad (12)$$

where the subscript  $s$  stands for seawater entry point (see Figure 3 for symbols). Continuity at the branching conduit is established in terms of fluid mass, solute mass and energy. That is

$$Q_m \rho_m = Q_f \rho_f + Q_s \rho_s \quad (13)$$

$$Q_m \rho_m c_m = Q_f \rho_f c_f + Q_s \rho_s c_s \quad (14)$$



**Figure 3.** Scheme of the conceptual model for brackish springs including the definition of symbols and notations used in the text. See also Nomenclature list.

and

$$H_{Bm} = H_{Bs} \quad (15)$$

where subscripts  $f$ ,  $s$  and  $m$  stand for fresh, sea and mixed waters, respectively,  $H_{Bm}$  is the energy at the branching point obtained from the mixed water conduit, and  $H_{Bs}$  is that obtained from the seawater conduit. These result from result from Equation (6), which yields:

$$\begin{aligned} H_{Bm} &= P_{Bm} + \rho_{Bm} \frac{v_m^2}{2} \\ &= \overline{\rho} g L_m \left( 1 + \frac{n_m^2 v_m^2}{R_{Hm}^{4/3}} \right) + \rho_{m\text{-spring}} \frac{v_m^2}{2} \end{aligned} \quad (16a)$$

where we have assumed that the conduit is vertical  $L_m = z_{\text{spring}} - z_B$ . In the steady state case,  $\overline{\rho} = \rho_{m\text{-spring}} = \rho_m$ , so that  $H_{Bm}$  simplifies to:

$$H_{Bm} = \rho_m g (L_m + b_m Q_m^2) \quad (16b)$$

with  $b_m = (1 + 2L_m n_m^2 / R_{Hm}^{4/3}) / 2A_m^2$ . Similarly, applying now Equation (6) between the sea entry and the branching points (assuming directly for simplicity that the seawater branch is filled with seawater), yields

$$H_{Bs} = P_{Bs} + \rho_s \frac{v_s^2}{2} = \rho_s g (-z_B - a_s Q_s^2) \quad (17)$$

with  $a_s = L_s n_s^2 / R_{Hs}^{4/3} A_s^2$ . For Darcy flow,  $H_{Bs}$  becomes

$$H_{Bs} = P_{Bs} + \rho_s \frac{v_s^2}{2} = \rho_s g \left( -z_B - \frac{\mu Q_s}{k A_s} \right). \quad (18)$$

In analysing these equations, note that  $z_B$  is negative.

### 3. Methods

#### 3.1. TURBOCODE solver

The above equations are solved using an iterative algorithm (called TURBOCODE) programmed in FORTRAN. Basically, input data include the description of the system (elevation, area, etc.) and the freshwater flow rate. Output consists of the seawater flow rate,  $Q_s$ , and derived variables such as mixed water flow rate and concentration ( $Q_m$  and  $C_m$ , respectively). The procedure is as follows:

- (1) Initialization. Read  $Q_f(t)$ ,  $z_{spring}$ ,  $z_s$ ,  $z_B$ ,  $L_m$ ,  $L_s$ ,  $n_s$  (or  $k_s$ ),  $n_m$  (or  $k_m$ ),  $A_s$ ,  $A_m$ . Time loop. For each time step perform the following operations.
- (2) Assume initial value for  $Q_s$  [ $Q_s(t) = Q_s(t-1)$ ;  $Q_s(0) = Q_f(0)$ ].
- (3) Using  $Q_f$  and Equations (13), (14) and (11) obtain  $Q_m$ ,  $c_m$  and  $\rho_m$ , respectively.
- (4) Solve transport using Equation (10) by means of a particle tracking method. Once the spatial distribution of concentrations is known, compute  $\bar{\rho}$ . Repeat for the seawater and mixed water conduits.
- (5) Compute  $H_{Bm}$  and  $H_{Bs}$  using Equations (16) and (17), or Equation (18).
- (6) If  $H_{Bm} \approx H_{Bs}$ , go to next time step. If  $H_{Bm} > H_{Bs}$ , reduce  $Q_s$  (otherwise, increase  $Q_s$ ), and go to step 3.

#### 3.2. Model setting

Simulations are first performed under steady-flow conditions of constant input freshwater flow to facilitate the understanding of the physics of the problem without memory effects. Transient simulations with a time-dependent freshwater input flow are used later to reproduce the actual functioning of brackish springs. The parameter values used in the simulations are listed in Table 1. They do not respond to any particular brackish spring but are selected partly from S'Almadrava spring (Mallorca, Spain) and partially from literature values of other springs of this type. The elevation of the spring mouth is +8 m.a.s.l. for all the simulations.

### 4. Results and discussion

#### 4.1. An energy balance analysis of Freshwater-Seawater mixing at the conduit branching

For a given freshwater flow rate entering the conduit branching, mixing can be viewed as resulting from a mass and energy (or momentum) balance problem. To understand this balance we first calculate the energy at the conduit branching measured from the seawater side ( $H_{Bs}$ ) as a function of seawater flow ( $Q_s$ ), after losses along the seawater conduit (Equations (17) or (18)). This energy is compared to  $H_{Bm}$  (Equation (16)), energy needed to bring the resulting mixed flow to the spring mouth (that is, the energy at the conduit branching measured from the mixed water side). The problem is solved when both energies become identical (Equation (15)). In fact, this is the way TURBOCODE operates. We consider steady-flow conditions, i.e., a constant density in every conduit, for discussion simplicity.

Results for the T-P conceptual model are shown in Figures 4a and 4b, for two series of small and large  $Q_f$  values, respectively.  $H_{Bs}$  decreases linearly with seawater flow rate because energy loss at the seawater conduit obeys Darcy's law (recall that energy at the sea intake is constant and equal to the hydraulic pressure of seawater). This linear relationship is specific to the seawater conduit and does not depend on how much freshwater is flowing into the system. By contrast, the energy necessary to push a column of mixed water towards the spring mouth,  $H_{Bm}$ , depends on both freshwater ( $Q_f$ ) and seawater ( $Q_s$ ) flow rates at the conduit branching. For any given  $Q_f$ , this energy increases with  $Q_s$  because both friction losses and density of the mixture (and thus the weight of the water column) increase with the mixing ratio. For a low  $Q_f$  (e.g., 0.25 m<sup>3</sup>/s, Figure 4a) energy losses in the conduit are small and the dominant term controlling the overall energy is the weight of the water column (Equation (5)), which decreases with decreasing  $Q_s$  (i.e., with the density of the mixture). For slightly larger  $Q_f$  (e.g., 1.0 m<sup>3</sup>/s, Figure 4a) we find that less energy is necessary to push up the same  $Q_s$  (Figure 4a). This reflects again the fact that energy is most sensitive to the density of the mixed water in the vertical column, which decreases with an increasing  $Q_f/Q_s$  ratio. In fact,  $H_{Bm}$  is virtually insensitive to  $Q_f$  for null  $Q_s$ . This can be attributed to the fact that, for small  $Q_f$ , friction losses in the vertical conduit are



**Table 1.** Parameter values used in the steady-flow and transient simulations, for Turbulent–Porous (T–P) and Turbulent–Turbulent (T–T) conceptual models

	Steady-flow T–P model	Steady-flow T–T model	Transient T–T model
$k_s$ (m <sup>2</sup> )	10 <sup>-6</sup>	-	-
$L_s$ (m)	2200	2200	2200
$A_s$ (m <sup>2</sup> )	35	0.5	0.5
$n_s$ (m <sup>-1/3</sup> ·s <sup>-1</sup> )	-	1.5 × 10 <sup>-2</sup>	1.5 × 10 <sup>-2</sup>
$z_s$ (m)	-	-	700
$z_B$ (m)	540	540	540
$A_m$ (m <sup>2</sup> )	2	2	2
$n_m$ (m <sup>-1/3</sup> ·s <sup>-1</sup> )	1.5 × 10 <sup>-2</sup>	1.5 × 10 <sup>-2</sup>	1.5 × 10 <sup>-2</sup>

See text or Nomenclature list for definition of parameter symbols.

also small and the energy required to push the mixed water up depends only on the weight of the water column, which is insensitive to  $Q_f$  when  $Q_s$  is near zero (Figure 4a).

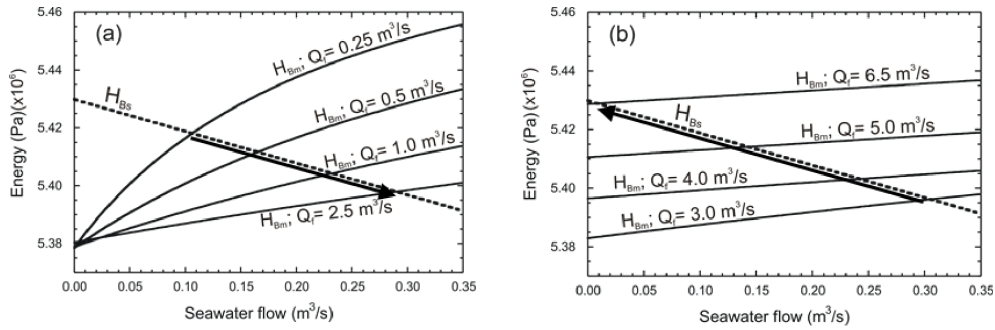
The situation changes for large values of  $Q_f$  (3.0 to 6.5 m<sup>3</sup>/s, Figure 4b) as the energy loss associated with flow resistance in the upwards conduit become increasingly important. In fact, above a critical  $Q_f$  (around 2.5 m<sup>3</sup>/s for this example) friction losses become more important than density effects. As resistance increases quadratically with the flow rate (Equation (7)), so does the energy required to allow the mixed water to ascend. It should be noted that the increase in energy needed to compensate the resistance of the wall of the conduit is not very sensitive to  $Q_s$  because for such high values of  $Q_f$  the proportion of  $Q_s$  in the total spring discharge is minor.

Every point in which two energy lines ( $H_{Bm}$  and  $H_{Bs}$ ) cross in Figure 4 represents a solution of the problem for steady flow regimes, i.e., the energy equilibrium point at the conduit branching for a specific freshwater–seawater mixing ratio (Equation (15)). In summary, for increasing  $Q_f$  values, equilibrium occurs with increasing  $Q_s$  below the freshwater critical value (Figure 4a), but with decreasing  $Q_s$  above it (Figure 4b). The overall trend is represented as a freshwater–seawater curve in Figure 5a. The freshwater–seawater curve is particular for every brackish spring and it is representative of the dimensions of the karst system and the physics of

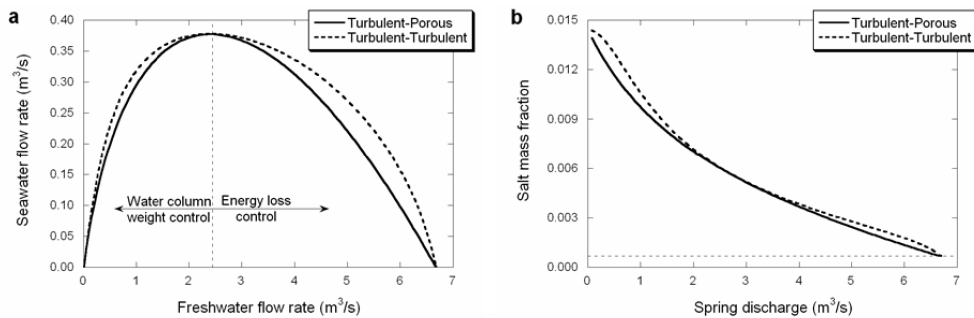
seawater contamination at depth. This critical  $Q_f$  value (around 2.5 m<sup>3</sup>/s for this example, Figure 5a) separates the flow conditions in which the system is primarily controlled either by the weight of the water column or by friction losses. Note that for very large  $Q_f$  values (higher than 6.7 m<sup>3</sup>/s, in this example) the energy equilibrium at the conduit branching is reached for energies higher than the hydraulic pressure of seawater at the conduit outlet to the sea, i.e., for negative  $Q_s$ . When this happens, freshwater flows through the conduit branching towards the sea and a submarine spring is created. This situation is addressed in detail in Section 4.3.

Figure 5b displays the relationship of salt mass fraction and the spring discharge. The relation reproduces the freshwater–seawater curve discussed above with a decrease on the salinity of the discharge with the flow rate.

For the T–T conceptual model, the behaviour of the mixed water conduit is identical to the T–P case because differences are restricted to the seawater conduit. Still, the T–T case also display two ranges of  $Q_f$  depending on whether the solution is controlled by the weight of the mixed water column, or by the energy loss in the conduit (Figure 5a). Differences arise in terms of the energy necessary for a particular seawater flow rate to occur ( $H_{Bs}$ , Equation (17)). The energy loss in the seawater conduit is no longer linear, but quadratic with  $Q_s$ , and so does the decrease of  $H_{Bs}$ . As a consequence, the energy equilibrium at the conduit branching is



**Figure 4.** Energy per unit volume computed at the branching point, as seen from the seawater conduit (dotted line,  $H_{BS}$ ) and the mixed water conduit (single line,  $H_{Bm}$ ), vs. seawater flow rate for (a) low ( $0.0\text{--}2.5\text{ m}^3/\text{s}$ ) and (b) high ( $3.0\text{--}6.5\text{ m}^3/\text{s}$ ) freshwater flow rates. Results correspond to steady-flow simulations of Turbulent–Porous conceptual model (Table 1).



**Figure 5.** Comparison of simulation results for steady-flow simulations of Turbulent–Turbulent and Turbulent–Porous conceptual models (Table 1). (a) Curves of freshwater–seawater ratio at the conduit branching. The ranges of freshwater flow rate for which the freshwater–seawater ratios are dominated by the weight of the water column or by the energy loss due to resistance are indicated. (b) Relation of salt mass fraction and spring discharge. Horizontal dashed line in the plots on the right marks the pure freshwater salt mass fraction in the simulations.

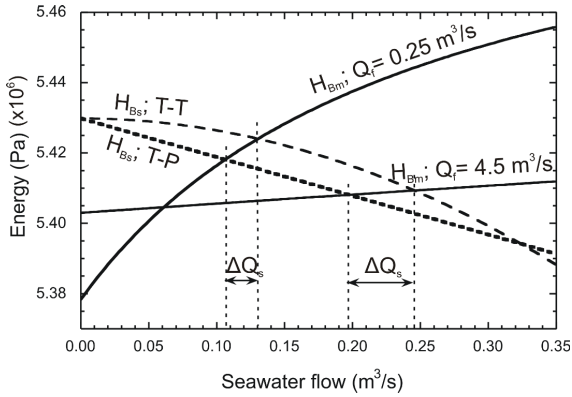
reached for different freshwater–seawater mixing ratios on each conceptual model. Figure 6 illustrates these differences for a low and a high  $Q_f$  in our examples ( $0.25$  and  $4.5\text{ m}^3/\text{s}$ ). Because the examples chosen to illustrate the two conceptual models are designed to hold the same freshwater–seawater ratio at the  $Q_f$  critical value, any equilibrium point for lower or higher freshwater flow rate yields a lower freshwater–seawater ratio at the conduit branching for the T–P conceptual model than for the T–T case (Figure 5a).

The relation of salt mass fraction and spring discharge in the T–T model differs from that obtained for the T–P case near the extreme values of spring discharge (Figure 5b). Thus, the salinity of the spring

for lower and higher flow rates on the T–T model increases slightly with respect to that of the T–P case to finally end at the same point. The relationship of the curve at extreme values of spring discharge may therefore be used as a distinctive feature of T–P or T–T conceptual models. This issue is discussed in detail later.

#### 4.2. Solution for the first extreme point (zero seawater and freshwater flow rates)

We discuss here the two points with zero seawater flow in the freshwater–seawater curve. Figure 5a shows that  $Q_s$  tends to zero when  $Q_f$  approaches zero and for a particular high value of  $Q_f$ . Figure 5a shows



**Figure 6.** Relationship of energy per unit volume computed at the branching point, as seen from the seawater conduit ( $H_{Bs}$ ) and the mixed water conduit ( $H_{Bm}$ ) vs. seawater flow rate for 0.25 and 4.5  $\text{m}^3/\text{s}$  of freshwater flow rate. Results correspond to steady-flow simulations of both Turbulent–Porous and Turbulent-turbulent conceptual models (Table 1). Single vertical dashed lines mark the energy equilibrium points. Arrows indicate the difference in seawater flow rate for a particular freshwater flow rate, between the T–T and T–P conceptual models.

that when the freshwater inflow from the aquifer decreases to zero, the seawater intrusion at the conduit branching also decreases but in a lesser degree. As a consequence, the density of the mixed water flowing up towards the spring mouth increases as does the weight of the water column connected to the spring mouth. Eventually, energy loss and kinetic terms become negligible. Ignoring all velocity dependent terms in Equations (16) and (17) or (18) yields

$$\rho_m L_m = -\rho_s z_B \quad (19)$$

which simply expresses that the weight (pressure) of seawater at the branching point equals the weight of the mixed water column. This expression allows us to derive the elevation of the branching point from the elevation of the spring mouth and the concentration for extremely low flow. In fact, using Equation (11) for  $\rho_m$  and  $\rho_s$ , and  $L_m = z_{\text{spring}} - z_B$  yields

$$z_B = \frac{z_{\text{spring}}}{1 - \exp[\alpha(c_s - c_m)]} \quad (20)$$

where  $z_B$  is negative because  $c_m < c_s$ . It is also worth pointing that, in the T–T case, all ignored terms

depend on  $Q^2$ . Therefore, concentration at the spring mouth should tend to a constant value as  $Q$  tends to zero. That is

$$\left. \frac{\partial c_m}{\partial Q} \right|_{Q=0} = 0 \quad (21)$$

which explains why  $c_m$  becomes constant for small  $Q$  in the T–T graph of Figure 5b.

On the contrary, for the T–P case the concentration at the spring mouth decreases linearly with  $Q$  for small  $Q$  values (Figure 5b). In fact, equating  $H_{Bm}$  (Equation (16)) and  $H_{Bs}$  (Equation (18)) and taking derivatives yields

$$\left. \frac{\partial c_m}{\partial Q} \right|_{Q=0} = -\frac{\mu}{L_m A_s k} \frac{(c_s - c_f)}{\alpha(c_m - c_f)} \quad (22)$$

which suggests that hydraulic resistance of porous conduit could be derived from the slope of the  $c_m$  vs.  $Q$  graph.

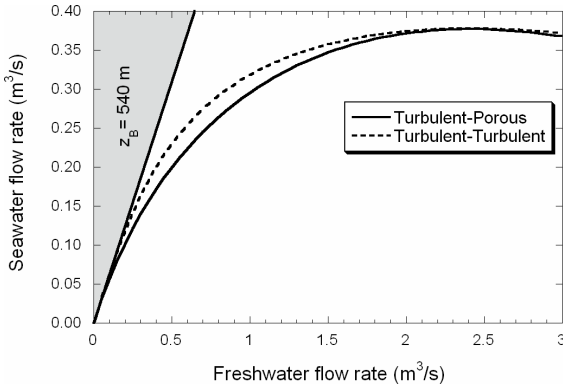
That the behaviour near the origin does not depend much on hydraulic parameters is also shown in the freshwater vs. saltwater discharge curve. Taking derivatives of the mass balance equations (13) and (14) with respect to the  $Q_f$  yields

$$\left. \frac{\partial Q_s}{\partial Q_f} \right|_{Q_f=0} = \frac{\rho_f c_m}{\rho_s (c_s - c_m)} \quad (23)$$

which depends solely on  $c_m$ . Since  $c_m$  at the origin only depends on  $z_B$ , so does the slope of the  $Q_s$ – $Q_f$  curve. This can indeed be seen in Figure 5a. In fact, this slope defines a straight line above which it is not possible to find any valid  $Q_s$  value.

In summary, the behaviour of the  $Q_s$ – $Q_f$  and  $c_m$ – $Q$  curves at the origin is informative of the depth of the branching point, which can be derived either from  $c_m$  or from the slope of the  $Q_s$ – $Q_f$  curve. It is also informative of the type of seawater conduit: turbulent if the slope of the  $c_m$ – $Q$  curve is zero, or darcian, otherwise. In the latter case, the slope of the  $c_m$ – $Q$  curve at the origin allows one to derive the hydraulic resistance of the seawater conduit.

For our example, representing a brackish spring with a spring mouth at +8 m.a.s.l. and a conduit branching at –540 m.a.s.l. (Table 1), we obtain a maximum potential salt mass fraction at the spring mouth of  $1.4 \times 10^{-2}$  (equivalent to 38% of seawater) (Equation (20)). The grey area in Figure 7 illustrates all freshwater–seawater combinations representing a salt mass fraction over the maximum potential calculated, i.e., combinations out of the calculated



**Figure 7.** Curve of freshwater–seawater ratio at the conduit branching for steady-flow simulations of Turbulent–Turbulent and Turbulent–Porous conceptual models (Table 1). The grey area mark freshwater–seawater ratios out of the potential range for the spring mouth elevation and conduit branching depth of the brackish spring simulated.

potential range. In other words, it shows freshwater–seawater mixing ratios that would never be measured at the spring discharge of our modeled brackish spring. As expected, for either conceptual model considered, the solution fits the maximum potential freshwater–seawater ratio for very low  $Q_f$  values. As the  $Q_f$  increases the solution separates from that potential ratio given that the energy loss becomes more significant. It should be pointed out that the dependence of  $Q_s$  on  $Q_f$  is more linear in the T–T case than in the T–P case. This is attributed to the fact that  $H_{Bs}$  remains essentially constant for low  $Q_s$  (Figure 6). This implies that salinity at the spring mouth will tend to be constant value for a range of very low flow rates (say, below  $0.2 \text{ m}^3/\text{s}$  in our example) in the T–T case, while it will grow steadily to the same value in the T–P case. It relevant to point out that this quasi linear behaviour does not require a Venturi effect, as suggested by Maramathas *et al.* [2006] for the Makaria spring.

#### 4.3. Second extreme point ( $Q_s = 0, Q_f > 0$ ) and submarine freshwater spring

The second situation in which we have zero  $Q_s$  occurs for a particular high value of  $Q_f$  ( $6.7 \text{ m}^3/\text{s}$  in

our example, Figure 5a). As discussed before, the energy loss at the mixed conduit increases quadratically with the mixed water flow, so that the seawater flow reduces to minimize the concentration of the mixed water, and therefore the weight of the mixed water column. This explanation implies that the resulting flow rate is independent of the seawater conduit behaviour. In fact, using again the energy continuity condition (Equation (15) with (16b) for  $H_{Bm}$  and either (17) or (18) for  $H_{Bs}$ ) together with  $Q_s = 0$ , yields

$$Q_f = \sqrt{\frac{-\rho_s z_B - \rho_f L_m}{\rho_f b_m}}. \quad (24)$$

Simple inspection of this equation shows that  $Q_f$  at this point solely depends of the elevation of spring and branching point and turbulent losses in the mixed conduit. Energy continuity, together with fluid and solute mass conservation can be used to compute the derivatives of  $C_m$  with respect to  $Q_m$  and  $Q_s$  with respect to  $Q_f$ . The derivation is a bit tedious and the results less informative than at the origin. The only relevant result is that the slope of both curves is far larger in the T–T case, than in the T–P case. This finding is related to the fact that head losses in the seawater conduit are proportional to  $Q_s^2$  in the T–T case, so that they are negligible for small  $Q_s$ . They are proportional to  $Q_s$  in the T–P case, so that they are sizeable even for small  $Q_s$ .

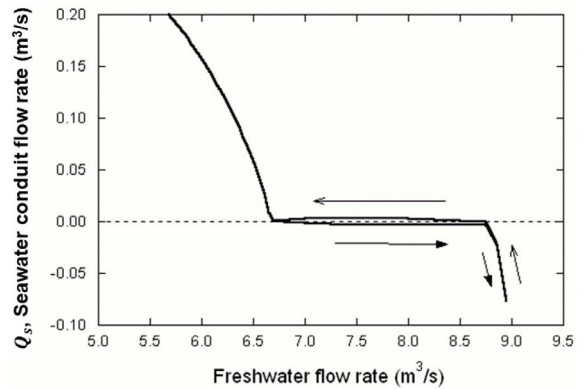
For even higher  $Q_f$ , the functioning of the spring is inverted. The energy necessary at the conduit branching to exceed the energy loss in the mixed water conduit is too high to be maintained. As a consequence, freshwater starts to flow towards the sea reducing the mixed water flow upwards. The freshwater intruding into the conduit connected to the sea promotes a negative  $Q_s$  (i.e., towards the sea) and a submarine spring is created. Concentration at the spring mouth becomes pure freshwater while the concentration in the submarine spring mouth is initially that of pure seawater. When that happens the functioning of the system changes radically and becomes dependent on the magnitude of the freshwater intrusion into the conduit connected to the sea. In fact, the magnitude of the freshwater intrusion will determine the salinity of the submarine spring, which may even become pure freshwater in an extreme situation. It is important to note that the minimum  $Q_f$  value for which a negative  $Q_s$  is produced is essential for a complete characterization of the

spring functioning since it defines the flow ranges for which the system behaves as an inland or a submarine spring. It is also important to note that although any freshwater–seawater curve calculated for a particular brackish spring has a theoretical  $Q_f$  before the  $Q_s$  is inverted, the actual chances for that high  $Q_f$  to occur in the aquifer may be very low and that situation could never happen in reality. Measuring pure freshwater at the spring mouth during high flow discharge periods may be an indicator that freshwater is intruding into the seawater conduit thus forming a submarine spring.

To obtain a more realistic solution to the problem involving submarine springs, equations must be solved in transient-flow mode. In order to do that, a time-dependent freshwater inflow function is designed with an increase from 5.0 to 9.0  $\text{m}^3/\text{s}$  followed by a decrease back to 5.0  $\text{m}^3/\text{s}$ , with a variation rate of 0.5  $\text{m}^3/\text{s}$  every 24 h. This freshwater function is introduced in a T–T simulation type with a sea opening at  $-700$  m.a.s.l. (Table 1). Note that for the T–T conceptual model the depth of the sea outlet needs to be defined (Equation (17)), since the average density in the conduit connecting with the sea may no longer be equal to that of the seawater. Results are shown in Figure 8.

Under these flow circumstances, the controlling parameters are the depth of the sea outlet and the average density in the conduit connected to the sea (which depends on the length of this conduit and the magnitude of the freshwater intrusion into this conduit). If the depth of the sea outlet is equal or less than that of the conduit branching, the pressure at the sea outlet is therefore less than that at the conduit branching, and freshwater is able to easily intrude into the conduit. The submarine discharge will become fresh very quickly, only depending on the length of the conduit.

By contrast, if the sea outlet is deeper than the conduit branching (as it happens in our example, Figure 8), the velocity of the freshwater intruding into the seawater conduit will be very small because pressure at the branching point needs to increase to overcome the double effect of having a higher pressure at the sea opening, and the reducing density of the water filling the conduit, initially assumed to be pure seawater. In Figure 8, freshwater flow,  $Q_f$ , increases from 6.7 to 8.75  $\text{m}^3/\text{s}$ , with a negligible flow in the seawater conduit. As  $Q_f$  increases



**Figure 8.** Curve of freshwater–seawater ratio at the conduit branching for transient simulation of Turbulent–Turbulent conceptual model (Table 1). Arrows indicate the order at which the values are generated when increasing freshwater flow rates (filled arrows) and when decreasing freshwater flow rates (simple arrows). The negative seawater flow situation applies briefly while there is seawater within the seawater conduit.

further, pressure at the branching becomes sufficient to overcome the reduced density of freshwater filling the sea conduit and the submarine spring discharges pure freshwater (freshwater flow above 8.75  $\text{m}^3/\text{s}$ , Figure 8). When  $Q_f$  decreases again (e.g., after an important rainfall event) the behavior of the system inverts and the fluid in the conduit becomes positive again (freshwater flow reduction from 8.75 to 6.7  $\text{m}^3/\text{s}$ , Figure 8). Note that for this range of flow rates, although groundwater flows back towards the conduit branching, the mixed water moving upwards to the spring is still pure freshwater. Only when the seawater front reaches the conduit branching (at 6.7  $\text{m}^3/\text{s}$  approx., in our example) will the spring discharge become salinized again.

#### 4.4. Sensitivity analysis

A set of simulations are conducted to illustrate the sensitivity of the spring response with respect to the parameters used in the simulation. First of all, we identify the controlling parameters for each conceptual model on steady-flow conditions from Equations (16) and (17) or (18). Then a perturbation of  $\pm 20\%$  is applied separately over the base value of every

**Table 2.** Values of the controlling parameters used in the base simulations for the sensitivity analysis, for Turbulent–Porous (T–P) and Turbulent–Turbulent (T–T) conceptual models. These values are perturbed by  $\pm 20\%$  to complete the analysis

T–P model		T–T model	
Parameter	Value	Parameter	Value
$A_m$ (m <sup>2</sup> )	2.0	$A_m$ (m <sup>2</sup> )	2.0
$n_m$	$1.5 \times 10^{-2}$	$n_m$	$1.5 \times 10^{-2}$
$z_B$ (m)	540.0	$z_B$ (m)	540.0
$L_s/k_s A_s$ (m <sup>-3</sup> )	$6.3 \times 10^{+7}$	$n_s \sqrt{L_s}/A_s$ (m <sup>-2.5</sup> )	1.41

See text or Nomenclature list for definition of parameter symbols.

parameter in order to quantify the relative sensitivity of the solution to that parameter (see parameter unperturbed base values in Table 2). The analysis is performed over steady-flow simulations since it allows a better isolation of the effect that each parameter has on the solution. The sensitivity analysis is discussed separately for the two conceptual models. Results are presented in terms of freshwater–seawater curves and relation of salt mass fraction with spring discharge. These representations are especially useful because they can be compared with the most commonly available field measurements in brackish springs, which are spring flow rates and water concentration with time.

#### 4.4.1. Turbulent–Porous conceptual model

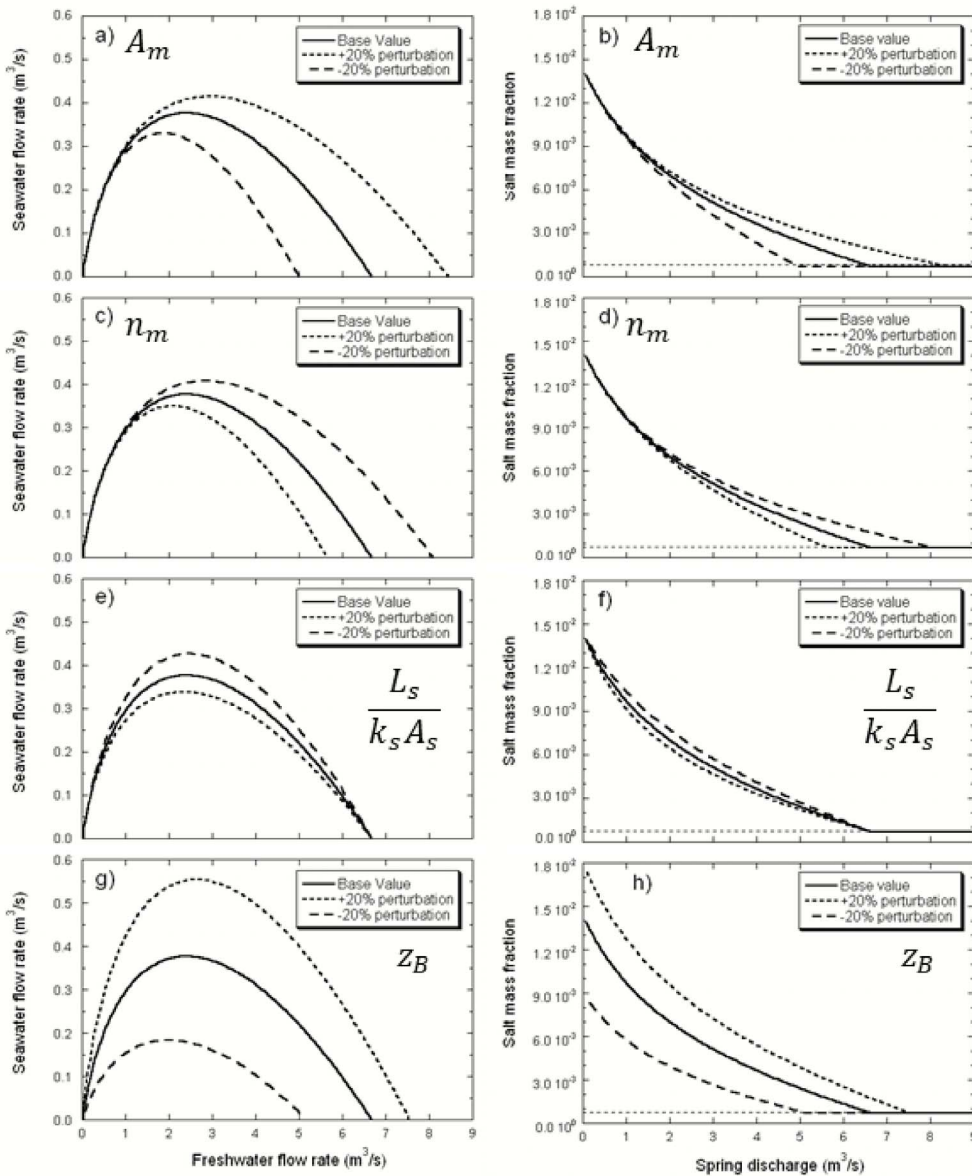
The solution for a T–P problem depends only on four parameters:  $A_m$ ,  $n_m$ ,  $L_s/k_s A_s$  and  $Z_B$ . Although this analysis is performed for steady-flow conditions, these four parameters (plus the history of  $Q_f(t)$ ) would control a transient-flow simulation as well. Figure 9 displays the results ( $Q_s$  vs.  $Q_f$ , and  $c_m$  vs.  $Q_m$ ) for the base case and perturbed parameters.

The parameters  $A_m$  and  $n_m$  control the resistance of the vertical conduit to the water flow (Equation (16)) and therefore show a similar—but opposite—influence on the spring response (Figure 9a–d). As the energy loss in the vertical conduit is a quadratic function of the water flow, the effect of these parameters is negligible for low  $Q_f$  or spring discharges (recall that for low  $Q_f$  the solution was primarily dominated by the weight of the water column, Figure 5a). However, their influence rapidly increases with  $Q_f$  (Figure 9a,c). As a consequence, the expected spring concentration will increase with

decreasing flow resistance, although this effect can be observed only for high spring discharges (Figure 9b,d).  $A_m$  not only controls the energy loss but also the velocity of the groundwater in the conduit, and therefore for transient-flow situations presents a relatively stronger influence on the spring response than  $n_m$ .

The solutions display a negative sensitivity with respect to the parameter  $L_s/k_s A_s$  (Figure 9e,f), i.e., an increase in the resistance to flow in the seawater conduit reduces the proportion of seawater entering the conduit branching. Obviously, the solution does not change near the two zero- $Q_s$  situations, which is also consistent with the equations governing the functioning mechanism of the spring. As discussed above, when  $Q_f$  approaches zero, the solution only depends on the depth of the conduit branching (Equation (18)) and, accordingly, the solution shows no sensitivity to any other parameter than  $z_B$  (Figure 9g,h). Further, for  $L_s/k_s A_s$  the solution shows no sensitivity at the other zero- $Q_s$  point either because the energy equilibrium for this situation to occur depends on the depth of the conduit branching and on the resistance in the vertical conduit (Figure 9, Equations (16) and (18)).

Finally, the depth of the conduit branching,  $z_B$ , exerts an influence on the spring concentration for any discharge (Figure 9h). As discussed above, the proportion of seawater entering the conduit branching increases with  $z_B$  because the pressure resulting from the weight of the seawater column increases accordingly (Figure 9g). Consequently, the spring salt mass fraction for any spring discharge is higher for deeper conduit branchings (Figure 9h). This parameter influences the energy curves of both the



**Figure 9.** Sensitivity analysis  $Q_s$  vs.  $Q_f$  curve (left column) and  $c_m$  vs.  $Q_m$  (right column), with respect to every controlling parameter of the Turbulent–Porous conceptual model (T-P):  $A_m$  (cross sectional area of the mixed water conduit),  $n_m$  (Manning’s coefficient of the mixed water conduit),  $L_s/k_s A_s$  (ratio of length, permeability and cross section area, of the seawater conduit) and  $z_B$  (depth of the branching point). The horizontal dashed line in the plots on the right marks the pure freshwater salt mass fraction in the simulations.

seawater and the mixed water conduit (Equations (16) and (18)). The slope of the freshwater–seawater curve when the freshwater flow approaches zero, is also dependent on the parameter value

(recall Equation (19), Figure 9g). Results also show that shallow conduit branchings will facilitate the formation of submarine springs because the minimum  $Q_f$  necessary to inverse  $Q_s$  is smaller.

An overall analysis of Figure 9 suggests that a spring with measures in a broad range of discharges may allow us to characterize: (1) the resistance of the upwards conduit (but not  $n_m$  and  $A_m L_m$  separately) from the response at high flow rate; (2) the depth of the branching point ( $z_B$ ) from the concentration at low discharges; and (3) the hydraulic resistance of the seawater conduit from the slope of  $c_m$  vs.  $Q_m$  at low spring discharges.

#### 4.4.2. Turbulent–Turbulent conceptual model

The solution for a T–T problem depends on four parameters:  $A_m$ ,  $n_m$ ,  $n_s \sqrt{L_s}/A_s$  and  $Z_B$ . Figure 10 shows the results of the sensitivity of the solution to these parameters. Results for this case are generally analogous to those of the T–P case and similar discussion could be applied. The most striking feature is that the behaviour at the two zero  $Q_s$  points is virtually insensitive to flow resistance parameters at the seawater conduit. This is consistent with Equations (21) and (24). The main difference with the T–P case is that results are less sensitive to  $z_B$  and that the abrupt change in  $Q_s$  near the critical  $Q_f$  remains nearly vertical for all parameter combinations (reflecting that energy losses in the seawater conduit are proportional to  $Q_s^2$ , so that they drop more significantly to zero when  $Q_s$  tend to zero than in the T–P case). The other relevant feature is that, for all parameter sets the maximum  $Q_s$  is flatter in the T–T case than in the T–P case.

In summary, the sensitivity analysis confirms that the main features to distinguish the T–T and T–P models are: (1) near the origin: the slope of the  $c_m$  vs.  $Q_m$  curve (near zero in the T–T case), (2) near the critical  $Q_f$  point, the slope of the  $Q_s$  vs.  $Q_f$  curve (nearly vertical in the T–T case), and (3) near the maximum  $Q_s$  point, the flatness of the  $Q_s$  vs.  $Q_f$  curve (inverted parabola and asymmetric in the T–P case and flatter, nearly circumferential in the T–T case).

#### 4.5. Comparison with field data

The results obtained in TURBOCODE simulations are compared with field data observations from three brackish springs: S'Almadrava spring (Mallorca, Spain, Figure 1), Pantan spring (Croatia) and Almyros of Heraklion spring (Crete, Greece). Figure 11 shows the relationship of the spring

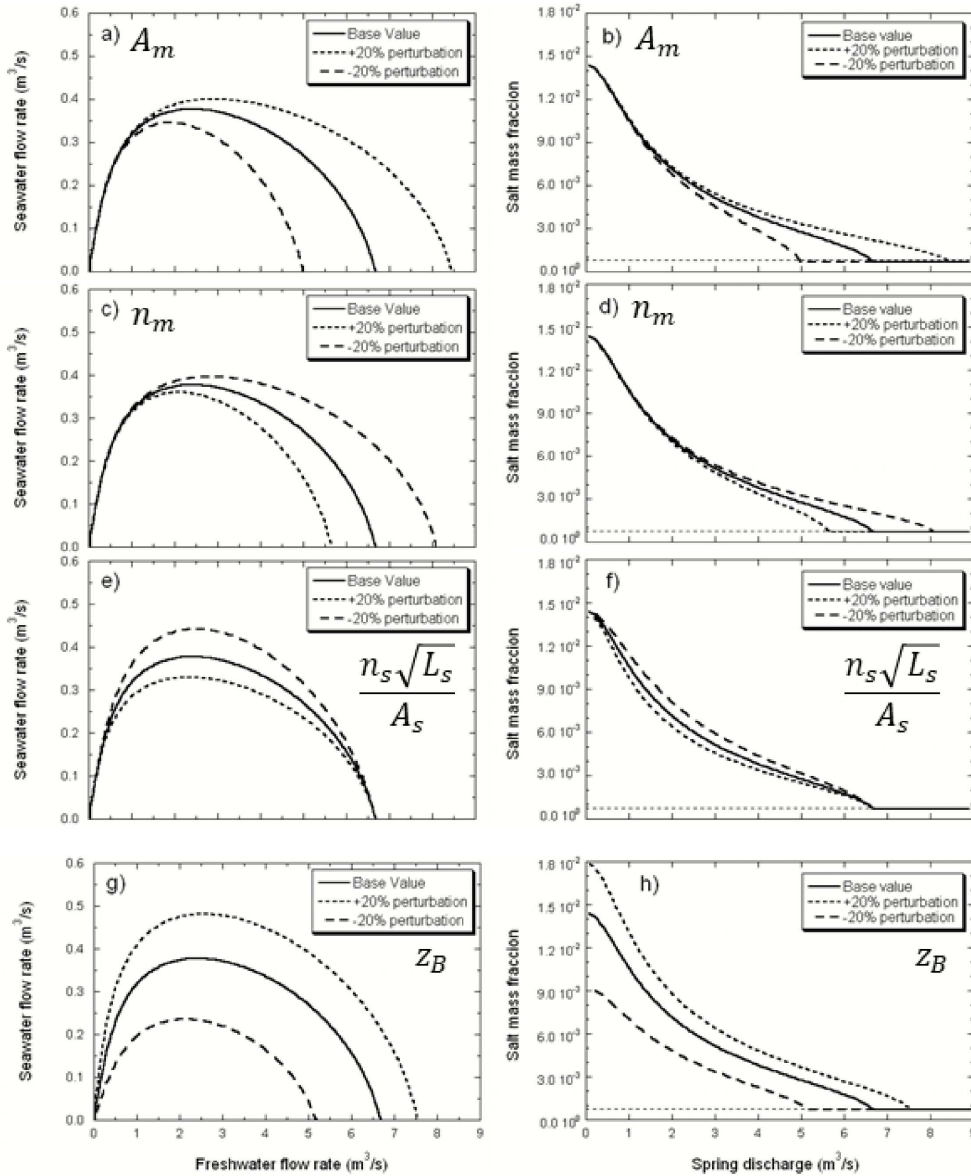
concentration vs. discharge for S'Almadrava [Sanz *et al.*, 2002], Pantan [Breznik, 1973] and Almyros [Monopolis *et al.*, 1995]. Beyond differences in number of observations and in salinity or discharge ranges, all springs display similar patterns with a decreasing salinity for increasing spring discharges (Figure 11). Based on the data available, the authors proposed that the salinization of these springs is due to the greater density of seawater in the T–T scheme, as discussed below. Field data for S'Almadrava covers a wide range on spring discharges (Figure 11a). From these observations, the conceptual model controlling the functioning of S'Almadrava spring would be a T–T case because: (1) the relation of salinity and flow rate is far from linearity for medium spring discharges, (2) for very low discharges, the spring concentration seems to reach a constant value (Figures 5b and 10, right hand side plots), and (3)  $Q_s$  remains quite constant near its maximum.

The available observations for Pantan spring also cover a wide range of spring discharges and the connection to the sea seems to be controlled by a karst conduit (T–T case) (Figure 11b). However, the limited number of data available prevents further discussion. In contrast, many observations are available for Almyros spring (Figure 11c). Data concentrates for medium to high spring discharges, where the spring clearly discharges freshwater. Field measurements seems to form a curve of high linearity. This observation seems more consistent with a T–P case (Figures 5b and 9, right hand side figures). However, the distinction of the conceptual model of Almyros turns out to be difficult due to the lack of data for very low flow discharges.

The fact that field measurements display some spreading with respect to the simulated curves for steady-flows presented in this study could be attributed to the memory effects resulting from the dynamic nature of the systems or be a consequence of the higher complexity of the karst network in the aquifers (e.g., multiple conduit branchings), compared with the simplifications considered in this study.

The freshwater seawater flow rates plots from field measurements in S'Almadrava spring and Almyros spring [Maramathas *et al.*, 2006] are shown in Figure 12. The data from S'Almadrava spring shows a zero- $Q_s$  point for  $Q_f$  approaching zero. The freshwater–seawater ratio then increases lin-

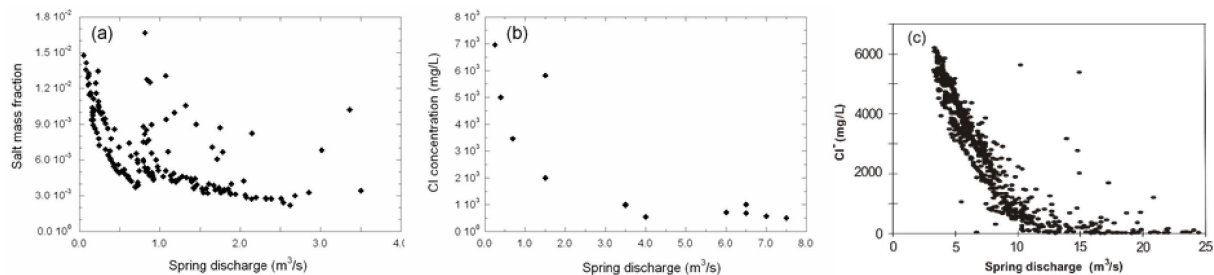




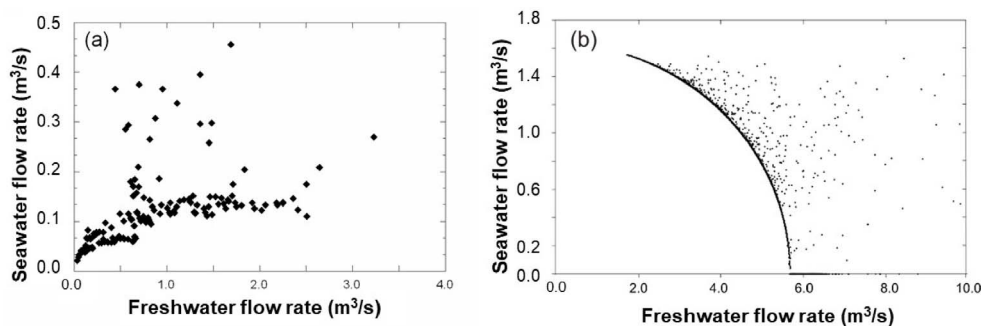
**Figure 10.** Sensitivity analysis  $Q_s$  vs.  $Q_f$  curve (left column) and  $c_m$  vs.  $Q_m$  (right column), with respect to every controlling parameters of the Turbulent–Turbulent conceptual model (T–T):  $A_m$  (cross section area of the mixed water conduit),  $n_m$  (Manning’s coefficient of the mixed water conduit),  $n_s \sqrt{L_s} / A_s$  (ratio of Manning’s coefficient, length and cross section area, of the seawater conduit) and  $Z_B$  (depth of the branching point). Horizontal dashed line in the plots on the right marks the pure freshwater salt mass fraction in the simulations.

early with  $Q_f$  before gradually stabilizing in about 0.15 m<sup>3</sup>/s of  $Q_s$  for a wide range of  $Q_f$ . This pattern agrees with simulations of this study for T–T conceptual model (Figures 5a and 10, left hand side figures).

The field data for this spring does not show any significant decrease in  $Q_s$  for higher  $Q_f$ , and therefore  $Q_f$  of at least 4 m<sup>3</sup>/s are expected to be necessary for the formation of a submarine spring.



**Figure 11.** Relationship of spring discharge and water quality for (a) S'Almadrava spring (Mallorca) [after Sanz *et al.*, 2002]; (b) Pantan spring (Croatia) [after Breznik, 1973]; and (c) Almyros spring of Heraklion (Crete) [after Monopolis *et al.*, 1995].



**Figure 12.** Seawater flow rate vs. freshwater flow rate for (a) S'Almadrava spring (Mallorca, Spain); and (b) Almyros of Heraklion spring (Crete, Greece) [after Maramathas *et al.*, 2006]. Note that the flow rate at Almyros never becomes small enough for the seawater flow rate to drop.

Field measurements for Almyros spring form a curve with an abrupt decrease in  $Q_s$  with  $Q_f$  until a zero- $Q_s$  point is reached for about  $5.6 \text{ m}^3/\text{s}$   $Q_f$ . Null seawater contribution (that is, pure freshwater discharging at the spring) is maintained for higher  $Q_f$  values, in line with the formation of a submarine spring at the sea bottom predicted in our simulations (Figure 8). The formation of submarine springs may be reconcilable with either T-T or T-P conceptual models and we would need more geology information to confirm any conceptual model. However, the abruptness of the freshwater–seawater curve when approaching the zero- $Q_s$  point indicates that a T-T scheme is more likely at Almyros spring. The different interpretations that may arise from the available observations for Almyros spring point out the difficulty of selecting the conceptual model governing brackish springs. This is especially so when the available observations do not cover a wide range of spring discharges including the two zero- $Q_s$

points. In fact, Arfib and de Marsily [2004] and Maramathas *et al.* [2003], built numerical simulations describing the T-P and T-T conceptual models, respectively, and obtained reasonable results in both cases.

## 5. Conclusions

In this study we use hydraulic models to reproduce the physics controlling the salinization of brackish springs. We derived the equations governing variable-density turbulent flow in conduits and solved them numerically. The solution has been found for systems consisting of a deep mixing place in which a freshwater conduit and another conduit connected to the sea join a third conduit with mixed water leading to the spring mouth. Two conceptual models have been tested: the Turbulent–Turbulent case in which the salinization of the conduit branching comes from a karst conduit connected to the sea, and the Turbulent–Porous case in which the

seawater intrusion occurs through a porous/fissured matrix (Figures 1 and 2, respectively). In both cases flow is governed by mass and either momentum or energy conservation. The response of the spring concentration to the variation of freshwater flow rate from the aquifer is evaluated in steady-flow conditions in terms of energy balance at the mixing point and freshwater–seawater ratios. For low freshwater flow rates, the spring response is controlled by the weight of the water column flowing from the conduit branching to the spring mouth. As the freshwater flow rate increases, the energy loss increases accordingly and the resistance of the conduit walls to flow becomes the controlling factor. Results are similar for both conceptual models for medium spring discharges, although the relation of spring salinity and discharge is more linear for the Turbulent–Porous case. Additionally, for the Turbulent–Turbulent case, the spring concentration becomes constant when the spring discharge tends to zero. This points out that in this situation the solution depends only on the water column weight in the vertical conduit. In fact, the spring dries up when the excess in elevation of the spring mouth is balanced by the excess of density of seawater with respect to that of the mixed water in the conduit connected to the spring mouth. Further, the depth of the conduit branching can be then approximated from the elevation of the spring mouth and the spring salinity at very low discharges.

The sensitivity of the spring response with respect to the parameters controlling the system is also evaluated in terms of freshwater–seawater ratios and the relation of spring discharge and salt mass fraction. For the Turbulent–Porous case, the simulation is sensitive to  $A_m$  and  $n_m$  for high freshwater flow rates, but their effect is opposite: salinity increases for larger  $A_m$  and smaller  $n_m$ . This is attributed to the quadratic increase of energy loss in the vertical conduit with the spring discharge. The solution for medium freshwater flow rates depends on  $L_s/k_s A_s$ . A higher parameter value increases the energy loss of the seawater intruding into the aquifer and thus reduces the salinization at the conduit branching. Finally, for low freshwater flow rates the solution only depends on the depth of the conduit branching since it controls the limiting water column weight that the system can support before the spring dries up. The sensitivity of the solution in the Turbulent–Turbulent case is very similar to that for the Turbulent–Porous

one for the parameters  $A_m$ ,  $n_m$  and  $Z_B$ . The resistance to flow in the conduit connected to the sea can be expressed now as  $n_s \sqrt{L_s}/A_s$ , where an increase of the resistance not only reduces the proportion of the seawater mixing at the branching point but also the salinity discharging from the spring. This effect is only important for medium freshwater flow rates since the solution for high flow rates is dominated by the energy loss in the vertical conduit rather than in the conduit connected to the sea. If the freshwater flow rate overpasses a limiting value, the high energy loss in the vertical conduit may promote a freshwater intrusion into the conduit connected to the sea. This will generate the formation of a submarine spring in the sea floor of variable density depending on the extent of the intrusion and the depth of the conduit outlet in the sea. This situation may be identified by measuring freshwater concentration during stages of high spring discharges in the spring.

The prediction curves of freshwater–seawater and the relationship of spring discharge and salt mass fraction are compared with field data from S'Almadrava spring, Pantan spring and Almyros of Heraklion spring. The simulation results show a good agreement with the field data available, and provide insights to identify the conceptual model governing every particular spring. The analysis highlights the importance of acquiring high frequency field data encompassing the whole range of spring discharges for a sound understanding of the spring functioning, specially at the extreme points of low seawater flow.

The freshwater–seawater curve is specific to every brackish spring and it is representative of the dimensions of the karst system and the salinization mechanism. In the light of our findings, an analysis based on the freshwater–seawater ratios rather than on the relationship of spring concentration and discharge proves to be more suitable for identifying the salinization mechanisms of brackish springs.

## Nomenclature

$A_l$	Cross sectional area of a conduit
$A'$	Total area of a conduit
$\varphi$	Porosity
$l$	Length along the conduit axis
$t$	Time

$\rho$	Water density, defined in Equation (11)
$\rho_0$	Density of pure water
$v$	Velocity
$f_P$	Shear component of the forces exerted by the conduit walls over the fluid
$Q$	Water flow rate at conduit
$z$	Depth
$z_{\text{spring}}$	Depth at spring mouth
$z_s$	Depth at the seawater entry point
$L$	Conduit length
$n$	Manning's coefficient
$R_H$	Hydraulic radius of the conduit (ratio of $A$ to wet perimeter)
$\mu$	Viscosity
$k$	Intrinsic permeability
$c$	Salt mass fraction
$\alpha$	Coefficient
$H$	Energy per unit volume
$H_{Bf}$	Energy at the conduit branching measured from the freshwater side
$H_{Bs}$	Energy at the conduit branching measured from the seawater side
$H_{Bm}$	Energy at the conduit branching measured from the mixed water side
$f$	Subscript for freshwater and freshwater conduit
$s$	Subscript for seawater and seawater conduit
$m$	Subscript for mixed water and mixed water conduit
$B$	Subscript for conduit branching
$\bar{\rho}$	Overbar stands for spatial average

## Conflicts of interest

Authors have no conflict of interest to declare.

## Acknowledgments

Comments by two anonymous reviewers and especially John M. (Jack) Sharp are gratefully acknowledged. Most of this work was funded by the European Union (SALTRANS project, contract no. EVK1-CT-2000-00062) and the Spanish Government (project Medistraes3. grant no. PID2019-110311RB-C21), and IDAEA-CSIC Center of Excellence Severo

Ochoa (Grant CEX2018-000794-S). ES is also grateful to a FI doctoral scholarship from the Catalan Government. The authors wish to acknowledge Dr. Guangjing Chen for his help and useful comments on the formulation and code design. Data and documentation related to S'Almadrava spring was gently provided by the Junta d'Aigües de Balears.

## References

- Acikel, S. and Ekmekci, M. (2018). Assessment of groundwater quality using multivariate statistical techniques in the Azmak Spring Zone, Mugla, Turkey. *Environ. Earth Sci.*, 77(22), 1–14.
- Ackerer, P., Carrera, J., and Delay, F. (2023). Identification of aquifer heterogeneity through inverse methods. *C. R. Géosci.*, 355(S1). Online first.
- Arfib, B. and de Marsily, G. (2004). Modeling the salinity of an inland coastal brackish karstic spring with a conduit-matrix model. *Water Resour. Res.*, 40, 1–10.
- Arfib, B., de Marsily, G., and Ganoulis, J. (2007). Locating the zone of saline intrusion in a coastal karst aquifer using springflow data. *Ground Water*, 45, 28–35.
- Arfib, B. and Ganoulis, J. (2004). Modélisation physique de l'intrusion d'eau de mer dans un aquifère karstique: cas de l'Almyros d'Héraklion (Crète). *C. R. Geosci.*, 336(11), 999–1016.
- Bakalowicz, M. (2014). Karst at depth below the sea level around the Mediterranean due to the Messinian crisis of salinity. Hydrogeological consequences and issues. *Geol. Bel.*, 17(1), 96–101.
- Bonacci, O. and Roje-Bonacci, T. (1997). Sea water intrusion in coastal karst springs: example of the Blaz Spring (Croatia). *Hydrol. Sci. J.*, 42, 89–100.
- Breznik, M. (1973). The origin of brackish karstic springs and their development. *Geologija*, 16, 83–186.
- Breznik, M. (1988). Development of the karstic spring Almyros in Greece. *Geologija*, 31, 555–576.
- Cardoso, G. (1997). *Comportamiento de los manantiales del karst nororiental de la Serra de Tramuntana, Mallorca*. PhD thesis, Dep. of Civil Eng., UPC, Barcelona.
- Chadwick, A. and Morfett, J. (1998). *Hydraulics in Civil and Environmental Engineering*. E and FN Spon, New York.

- Crosby, F. W. and Crosby, W. O. (1896). The sea mills of Cephalonia. *Technol. Q.*, 9, 6–23.
- de Marsily, G. (1986). *Quantitative Hydrogeology*. Paris School of Mines, Fontainebleau.
- de Marsily, G. (2009). *L'eau, un trésor en partage*. Dunod, Paris.
- Fleury, P., Bakalowicz, M., and de Marsily, G. (2007a). Submarine springs and coastal karst aquifers: A review. *J. Hydrol.*, 339, 79–92.
- Fleury, P., Bakalowicz, M., de Marsily, G., and Cortes, J. M. (2007b). Functioning of a coastal karstic system with a submarine outlet, in southern Spain. *Hydrogeol. J.*, 16(1), 75–85.
- Fouqué, F. (1867). Rapport sur le tremblement de terre de Céphalonie et de Mételin en 1867. In *Archives des missions scientifiques et littéraires*, volume 4, pages 445–482.
- Fuller, M. (1907). Conditions of circulation at the sea mills of Cephalonia. *Geol. Soc. Am. Bull.*, 18(1), 221–232.
- Gjurasin, K. (1943). Prilog hidrografiji primorskog krsa. *Tech. Vjesn.*, 60, 1–17.
- Guarracino, L., Carrera, J., and Vázquez-Suñé, E. (2012). Analytical study of hydraulic and mechanical effects on tide-induced head fluctuation in a coastal aquifer system that extends under the sea. *J. Hydrol.*, 450, 150–158.
- Hughes, J. D., Vacher, H. L., and Sanford, W. E. (2007). Three-dimensional flow in the Florida platform: Theoretical analysis of Kohout convection at its type locality. *Geology*, 35(7), 663–666.
- Kohout, F. A. (1965). A hypothesis concerning cyclic flow of salt water related to geothermal heating in the Floridan aquifer. *Trans. N. Y. Acad. Sci.*, 28(2), 249–271.
- Kohout, F. A., Henry, H. R., Banks, J. E., Smith, K. L., and Griffin, G. M. (1977). Hydrogeology related to geothermal conditions of the Floridan Plateau. In Smith, D. L. and Griffin, G. M., editors, *The Geothermal Nature of the Floridan Plateau*, volume 21, pages 1–41. Bureau of Geology, Div. of Resour. Manag., Tallahassee, FL.
- Kuscer, I. (1950). *Kraski izviri ob morski obali*. Razprave SAZU, Ljubljana.
- Lambrakis, N., Andreou, A. S., Polydoropoulos, P., Georgopoulos, E., and Bountis, T. (2000). Nonlinear analysis and forecasting of a brackish karstic spring. *Water Resour. Res.*, 36, 875–884.
- Leontiadis, I. L., Payne, B. R., and Christodoulou, T. (1988). Isotope hydrology of the Aghios Nikolaos area of Crete, Greece. *J. Hydrol.*, 98, 121–132.
- Lofi, J., Berne, S., Tesson, M., Seranne, M., and Pezard, P. (2012). Giant solution-subsidence structure in the Western Mediterranean related to deep substratum dissolution. *Terra Nova*, 24(3), 181–188.
- Maramathas, A., Maroulis, Z., and Marinos-Kouris, D. (2003). Brackish karstic springs model: application to almiros spring in crete. *Ground Water*, 41, 608–619.
- Maramathas, A., Pergialiotis, P., and Gialamas, I. (2006). Contribution to the identification of the sea intrusion mechanism of brackish karst springs. *Hydrogeol. J.*, 14, 657–662.
- Maurin, V. and Zoetl, J. (1967). Salt water encroachment in the low altitude karst water horizons of the island of Kephallinia (Ionian Islands). In *Hydrologie des roches fissurées*, volume 2, pages 423–438. AIHS, Dubrovnik, Yugoslavia.
- McCullough, M. L. and Land, L. S. (1992). Dynamic hydrology and diagenesis of a submerged Pleistocene fringing reef, Discovery Bay, Jamaica. *Mar. Geol.*, 104(1–4), 139–151.
- Monopolis, D., Sofiou, P., Papamastorakis, D., Steiakakis, E., Kleidopoulou, M., and Vavadakis, D. (1995). Study of hydrological parameters of almiros spring (Crete – Greece) under technical flow conditions. In *Proceedings of 3rd Congress of the Greek National Committee of Hydrogeologists*, pages 132–149.
- Payne, B. R., Leontiadis, J., Dimitroulas, C., Dounas, A., Kallergis, G., and Morfis, A. (1978). Study of the Kalamos springs in Greece with environmental isotopes. *Water Resour. Res.*, 14, 653–658.
- Pinault, J. L., Doerfliger, N., Ladouche, B., and Bakalowicz, M. (2004). Characterizing a coastal karst aquifer using an inverse modeling approach: The saline springs of Thau, southern France. *Water Resour. Res.*, 40, 1–17.
- Rezaei, M., Sanz, E., Raeisi, E., Ayora, C., Vázquez-Suñé, E., and Carrera, J. (2005). Reactive transport modeling of calcite dissolution in the fresh-salt water mixing zone. *J. Hydrol.*, 311, 282–298.
- Ryan, W. B. (1976). Quantitative evaluation of the depth of the western Mediterranean before, during and after the Late Miocene salinity crisis. *Sedimentology*, 23(6), 791–813.
- Sanford, W. E. and Konikov, L. F. (1989). Simulation of

- calcite dissolution and porosity changes in saltwater mixing zones in coastal aquifers. *Water Resour. Res.*, 25, 655–667.
- Sanz, E., Ayora, C., and Carrera, J. (2011). Calcite dissolution by mixing waters: geochemical modeling and flow-through experiments. *Geol. Acta*, 9(1), 67–77.
- Sanz, E., Custodio, E., Carrera, J., Ayora, C., Baron, A., and Gonzalez, C. (2002). Modelling coastal salty springs: first approach in carbonate media (S'Almadrava, Mallorca, Spain). In *Proceedings of 17th Salt Water Intrusion Meeting, Delft, The Netherlands*.
- Sanz, E., Custodio, E., Carrera, J., Ayora, C., Barón, A., and González, C. (2003). Modeling karst salty springs in coastal carbonate aquifers (S'Almadrava, Mallorca, Spain). In *Proceedings of 2nd Salt Water Intrusion and Coastal Aquifers, Merida, Mexico*.
- Stringfield, V. T. and LeGrand, H. E. (1969). Relation of sea water to fresh water in carbonate rocks in coastal areas, with special reference to Florida, U.S.A., and Cephalonia (Kephallinia), Greece. *J. Hydrol.*, 9, 387–404.
- Wiebel, K. W. M. (1873). *Die insel Kephallonia und die Meermühlen von Argostoli: Versuch einer Lösung dieses geophysikalischen Räthsels (The island of Kephallonia and the sea-mills of Argostoli: an attempt at solving this geophysical riddle)*. TG Meissner, Hamburg, Germany.



Research article

## Geo-hydrological Data & Models

# New insights from an enhancement of hydrogeological archives of a French karst watershed over the last 50 years

David Labat<sup>Ⓣ,\*<sup>a</sup></sup>, Thomas Guilloteau<sup>a</sup>, Sébastien Lamagnère<sup>a</sup>, Thaïs Dumareau<sup>a</sup>,  
Gabin Haryouli<sup>a</sup>, Sandra Béranger<sup>b</sup>, Franck Granouillac<sup>c</sup> and Vianney Sivelle<sup>Ⓣ,d</sup>

<sup>a</sup> Geosciences Environment Toulouse UMR CNRS IRD Université Paul Sabatier CNES,  
14 Avenue Edouard Belin 31400 Toulouse, France

<sup>b</sup> Bureau de Recherches Géologiques et Minières-Délégation Régionale Occitanie-Site  
de Toulouse, 31520 Ramonville-Saint-Agne, France

<sup>c</sup> Laboratoire Écologie Fonctionnelle et Environnement, University of Toulouse,  
CNRS, Auzeville Tolosane, 31326 Castanet Tolosan, France

<sup>d</sup> HSM, University of Montpellier, CNRS, IRD, Montpellier, France

*E-mails:* david.labat@univ-tlse3.fr (D. Labat), thomas.guilloteau@univ-tlse3.fr  
(T. Guilloteau), sebastien.lamagnere@univ-tlse3.fr (S. Lamagnère),  
thais.dumareau26@gmail.com (T. Dumareau), gabinhry85@gmail.com (G. Haryouli),  
s.beranger@brgm.fr (S. Béranger), franck.granouillac@univ-tlse3.fr (F. Granouillac),  
vianney.sivelle@umontpellier.fr (V. Sivelle)

**Abstract.** Karst aquifers, like most hydrological systems, are sensitive to the influence of changes in land use as well as current and future climate change. To be able to identify the effects of these changes, in particular on the variability of groundwater flows, we have endeavored to produce one of the longest hourly spring discharge time series available for a French karst watershed. This reconstitution is based on a digitisation of the limnigraphic archives available for the Baget karst watershed. Correlation, spectral and multifractal analyses were then performed based on the spring discharge time series to confirm the multiscale nature of the hydrological response of this basin and to highlight the non-stationarity of this response.

**Keywords.** Digitalisation, Data analysis, Karst system, Hydrogeology, Spectral analysis, Multifractal analysis.

*Manuscript received 27 June 2022, revised 26 September 2022 and 27 September 2022, accepted 28 September 2022.*

## 1. Importance and originality of karsts

From a general point of view, responsibility, negotiation, participation, solidarity are only made possi-

ble by an awareness of the “common good” character of the groundwater, a “groundwater consciousness” [Marsily and Besbes, 2017]. Groundwater resources belong to a category of natural freshwater resources that are difficult to understand and study. Since the second half of the 20th century, demographic pres-

\* Corresponding author.

sure and the development of intensive agriculture have put this essential resource at risk. This strong pressure is increasing today with the visible and future effects of climate change combining a decrease in rainfall recharge and an increase in evapotranspiration. Stevanović [2019] estimates that less than 5% of the yearly renewable water in karst aquifers is nowadays exploited. Thus, karst aquifers constitute an important source of freshwater that may allow equilibrating the increasing imbalance between freshwater availability and the need for sustaining human activities (drinking water supply, agriculture, industry, and others).

Worldwide, carbonate rock outcrops cover 12–15% of the continental surface. In France, they cover 35% of the metropolitan territory, but their extension under other geological formations is much higher. According to the most recent estimate, almost 50% of the French territory is affected by karst formations [Chen *et al.*, 2017]. Karst aquifers contribute to 40% of drinking water supply of France. [Chen *et al.*, 2017]. About 25% of the world's population is supplied with domestic water from groundwater taken from karst [Ford and Williams, 2013].

The particular character of karst aquifers is linked to the heterogeneity of the reservoir resulting from the karstification processes. The presence of karst morphology is related to the dissolution of carbonate rocks by water enriched in CO<sub>2</sub> by root respiration and microfauna combined with the oxidation of organic matter. The setting up of karst systems is therefore the result of complex processes such as fracturing, mechanical erosion and dissolution of carbonate rocks [Quinif, 1999]. The occurrence of a large-scale multi-porosity promotes the storage of a large amount of water, but also the occurrence of complex flow behaviour leads to significant difficulties in the utilisation of this freshwater resource. A large part of the flow is thus concentrated in groundwater preferential pathways, mainly conduits or unclogged fractures, with relatively rapid flow velocities compared to those observed in porous or fissured media. Therefore, darcian flows can be observed in the matrix, while velocities around 1 mm/s can be observed in the fissures or small fractures and velocities superior to 1 m/s are often measured in the conduit network.

Another part of the flow takes place in the matrix and may have some exchanges with the preferential flow path. The hydrological behaviour of

karst systems is therefore highly non-linear and non-stationary, making it difficult to characterize and predict their functioning. Karst aquifers are not only sensitive quantitatively, due to the potential evolution of recharge processes in the context of climate change [Tramblay *et al.*, 2020] but also qualitatively, due a rapid infiltration of potential contaminants within the aquifer [Hartmann *et al.*, 2021]. It is therefore important to be able to respond to the legitimate societal demand for access to an integrated groundwater management model.

## 2. Objectives

For many years, karst aquifers have been the site of numerous experiments aimed at understanding, analysing and modelling the occurrence of non-linear flow behavior. Today, this concern is stressed by the need to have access to a history of this hydrological response to identify possible non-stationarities or to corroborate the analyses already carried out for a short observation period. Several scientific operations are striving to document precisely the functioning of different hydrological compartments within OZCAR project [Gaillardet *et al.*, 2018] and now at European level with the ELTER programme. Karst aquifers are recognised as an original compartment from both a hydrological and hydrobiogeochemical point of view and the creation of the National Karst Observation Service (SNO Karst) is the logical continuation of this awareness [Jourde *et al.*, 2018]. Among the French karst watersheds, which are continuously monitored, the present study focuses on the Baget basin (Pyrénées, France), where the first investigations began in 1969 [Mangin, 1975].

Beginning with an intense period of research activities between 1969 and 1985, monitoring of the basin was then limited to the estimation of daily flows at the main outlet until 1996, when the discharge sampling step was fixed at 30 min, first sporadically, and then continuously from 2005. Former studies [Labat, 2000, Labat *et al.*, 2011] put in evidence that a 30-min sampling rate is optimal for acquiring relevant information on the hydraulic behaviour of the basin. We describe in this contribution an enhancement of limnigraphic archives that allows to propose an estimate of hourly discharge over the last fifty years. In a second step, we compare the results already obtained using time series analysis on



discharge at more restricted time intervals with the results obtained at much longer time interval in this study.

### 3. Study site: the Baget karst watershed

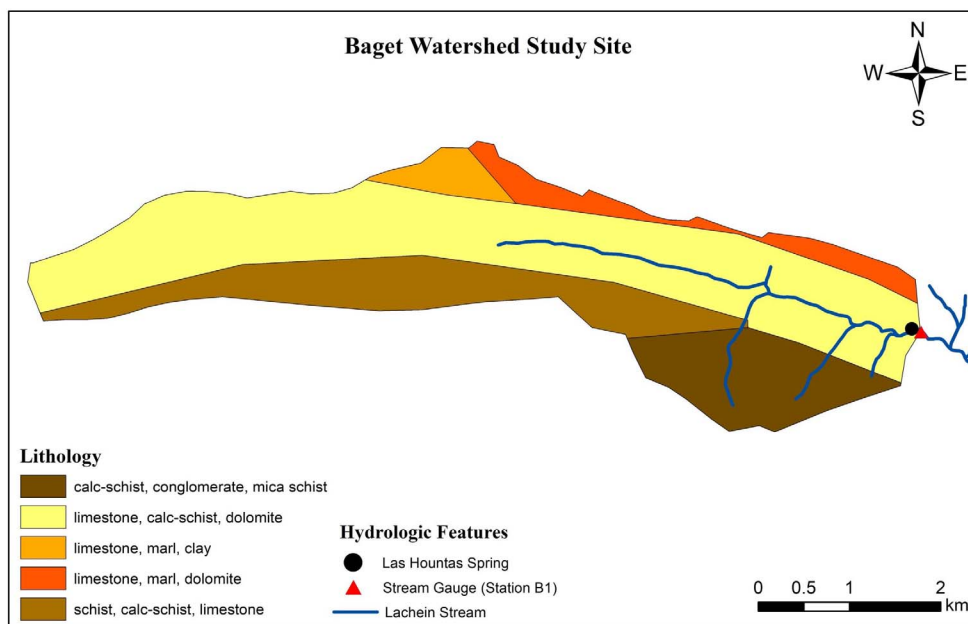
The Baget catchment area is located in the Ariège in the commune of Agert-Balaguères. The Baget valley naturally leads to the Lez valley about 10 km west of Saint-Girons. The catchment area of the Baget karst system extends from West to East over 10 km in length and 1 km width for an area of approximately 13 km<sup>2</sup>. This system is located in the metamorphic zone bounded by two faults: to the north the North Pyrenean fault, and to the south the Alas fault. The Mesozoic limestones can be considered globally as a monoclinical structure. These limestones then come into contact with the Paleozoic terrains to the north and disappear to the south under the pelitic clay-sandstone formations of the Albo-Cenomanian [Debroas, 2009]. The limestones present within the basin are very homogeneous and crystalline, with a porosity of around 2%. The Lachein valley is characterized by a fairly deep furrow of 300 m which follows the longitudinal axis of the basin. The most important reliefs are in the East with the Tuc de Graué (1417 m) while the outlet is located at 498 m. The average altitude of the basin is around 923 m and its median altitude around 930 m. Within the Baget basin, several overflows should be mentioned, including the overflows of Moulo de Jaur and that of Hillère in particular.

The non-calcareous land represents approximately 33% of the total surface area and is divided into two sectors: sandstone clay pelites in the South and Upper Cretaceous conglomerates in the North. They are characterized by significant temporary superficial flow during periods of flooding. The Lachein stream thus meets the limestones about 4.5 km upstream from Las Hountas (Figure 1). At the end of low water, the waters of the stream bury themselves very quickly in contact with the limestone. During periods of average water capacity, the surface water travels 2.5 km over the limestone before disappearing and during periods of high water they progress even further downstream to get lost at the level of the Peyrère. During episodes of strong floods, the loss of the Peyrère can behave like an emergence.

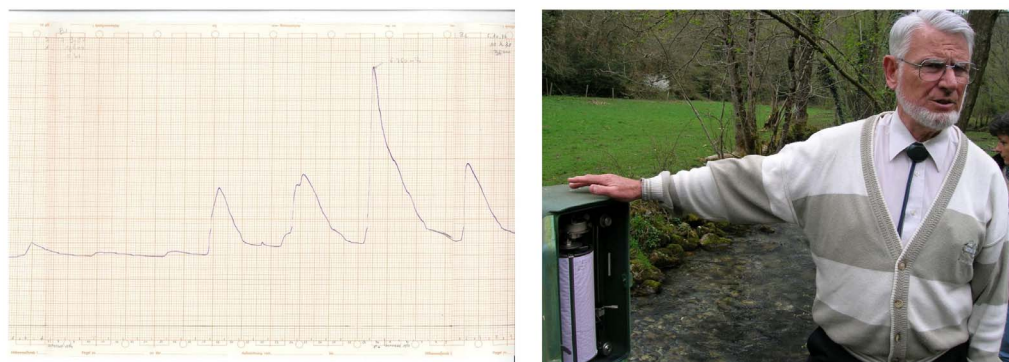
The average rainfall representative of the basin is around 1800 mm per year (with a variation rate of around 400 mm per year) with an average annual evapotranspiration estimated at around 600 mm per year. The average flow of the Baget is estimated at 0.51 m<sup>3</sup>/s for a minimum observed flow of around 0.05 m<sup>3</sup>/s. Overall for a typical hydrological cycle, it is estimated that 66% of the total flow volume comes from the perennial source of Las Hountas. In addition, several pumping tests with large flows were carried out in the karstic zone in 1991, showing a higher level of karstification in the upstream part of the basin and also the high potential of this karst in terms of water resources.

### 4. Materials

We propose here a reconstruction of the hourly flows using the limnigraphs that have been carefully kept for years. These limnigraphs have been annotated rigorously by Alain Mangin (Figure 2) through all these years. We use them to propose an estimate of the hourly flow rate at the Baget outlet over the last 50 years. The monitoring of water levels at the outlet was first carried out by the teams of the Moulis underground laboratory, with a 1/5 scale OTT 16 water level gauge until 1996. Then, a differential pressure sensor probe was installed to provide automatic acquisition. However, the discharge measurements were not completely continuous until 2005. Since 2017, the french geologic survey (BRGM Occitanie), in collaboration with the Moulis experimental ecological station, is responsible for monitoring water levels using the same equipment. Thus, for the period 2005–2012, we have at our disposal both the water level gauges and the data from the differential pressure sensor installed at the monitoring station. The long manual digitisation work, by creating the abscissa and ordinate axes from the reference data collected on-site, was carried out with the help of the software available online “PlotDigitizer”. This software allows the digitisation of the graphical documents with an average resolution of 2 to 3 h directly linked to the speed of the water level gauge clock. Thus, the periods of rising floods, in particular, are correctly estimated. In a second step, a linear interpolation was carried out to produce a continuous series of water levels over the period 1972–2008.



**Figure 1.** Geological map of the Baget karst system and location of the gauging station (thanks to the courtesy of Ibrahim El-Khoury).



**Figure 2.** Left: Example of limnigraph digitised including the annotated information. Right: A. Mangin in front of the limnigraphic station (photo by Seifeddine Jomaa).

The 2005–2008 period has been considered as a validation period for comparison of digitised data versus the differential pressure sensor data. For the period 2008–2022, we used the data available from the digitisation and the differential pressure sensor measurements were compared to validate the digitisation work. In general, the correspondence between the two sets of data was less than 3 cm. Some corrections related to the drift of the probe were also made. During the last fifty years, the hydraulic control of the rat-

ing curve at the outlet of the system has been very stable. Effectively, we estimate the discharge rate using only three successive rating curves of the station, including the construction of the artificial weir in 1976, illustrated by the step visible on the water level record. Therefore, the rating curve is still valid from 1976.

In conclusion, we already have at our disposal the daily discharge estimates from 1969 for the Baget system. Thanks to the digitization of the limni-

graphs, we propose to estimate discharge at a mean sampling rate of about 3 h (depending on the hydrologic period) over the 1970–2008 time interval. The 2005–2008 time interval allows validation of the method based on a systematic comparison with in situ differential pressure sensor probe. Therefore, a linear interpolation of the digitized data permits continuous hourly estimation of the discharge data over the last fifty years.

## 5. Methods

Correlation and spectral analysis allow us to characterise the temporal structure of a hydrological signal [Mangin, 1984, Padilla and Pulido-Bosch, 1995, Larocque *et al.*, 1998, Labat, 2000]. The decay of the correlogram characterises the hydrodynamic response of the karst system to rainfall events. The shorter the decay, the more the system is characterised by high karstic conductivity. On the contrary, the longer the decay, the more the system is characterised by low karstic conductivity [Box *et al.*, 2008]. Generally, the length of influence of an event corresponds to the memory effect with respect to the first time lag, where the correlogram reaches the value of 0.37 [Labat, 2000]. The classical threshold value of 0.2 can only be applied to daily data where the decay of the autocorrelation function is sufficiently fast.

Spectral analysis constitutes another powerful statistical technique for analyzing the distribution (over frequency) of the power contained in a signal. This technique is widely used to identify patterns in data series [Box *et al.*, 2008] and has been applied specifically in determining the presence and statistical significance of natural climate cycles in long-term series [Mathevet *et al.*, 2004]. Thus, classical spectral analyses consist of studying the variance spectral density distribution  $S(f)$  and determining at which frequency  $f$  intervals the Fourier spectral density follows power-law behavior:

$$S(f) = f^\beta \quad (1)$$

where  $\beta$  is the slope of the power law.

It is also possible to focus more specifically on the high-frequency behavior of the system to identify distinct power-law behaviors and characteristic cut-off scales.

Multifractal analysis is a method, which investigates the scale-invariance properties of a signal [Labat *et al.*, 2002, 2013]. Multifractal analyses consist of looking more carefully and systematically at the distribution of the various empirical moments distribution function of the scale of observation. The spectral analysis only focuses on the distribution of the variance (i.e., the squared signal), whereas here we provide an insight of the time series on all  $q$  power. The dyadic discrete time series under study (rainfall or runoffs here) noted  $\varphi(i)$  (with  $i = 1, \dots, 2^N$ ) is differentiated and mean-normalized to get a new signal  $\varepsilon_1(i)$ . This new signal  $\varepsilon_1(i)$  is then aggregated at intervals of increasing length  $\lambda = 2^m$ , with  $m = 2, \dots, N$ , to get  $N$  discrete signals  $\varepsilon_\lambda(i)$ , each one corresponding to a scale of observation. The next step consists in looking at the behavior of the generalized moment function  $M(\lambda, q)$  defined as the moment of the time series aggregated at a given scale  $\lambda$  and for a given power law  $q$ :

$$M(\lambda, q) = \langle (\varepsilon_\lambda)^q \rangle. \quad (2)$$

Log–log representations of  $M(\lambda, q)$  can then highlight power-law behaviors:

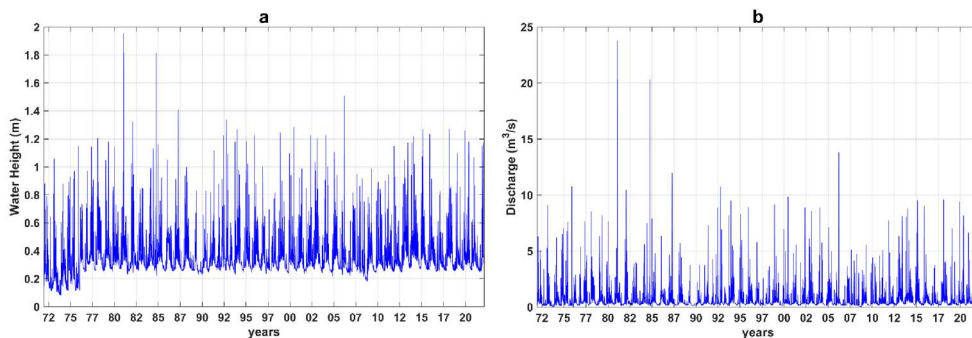
$$M(\lambda, q) \propto \left( \frac{\lambda}{L} \right)^{-K(q)} \quad (3)$$

where  $L$  is the largest scale available (here,  $L = 2^N$ ) and  $K(q)$  corresponds to the multifractal spectra of the time series.

## 6. Results

The hourly discharge time series (Figure 3) represents a continuous series of more than 415,000 data points. It is possible to carry out two types of analyses: a global analysis of the series on the one hand, and an analysis of the flow series taken over disjointed two-year intervals on the other. The second analysis highlights potential non-stationarity of statistical characteristics over the period studied.

The correlation analysis of the hourly discharge estimations over the 1970–2022 period (Figure 4) shows a global memory effect of around 110 h. Previous results based on daily data indicate a large value of around 8 days [Labat, 2000]. We also calculate the envelope curve of the auto correlation function corresponding for each delay time to the lower and higher value obtained from the auto correlation calculated over the disjointed two-year intervals. The



**Figure 3.** (a) Estimation of hourly water levels at the Baget outlet station and (b) estimation of discharge time series at the Baget station.

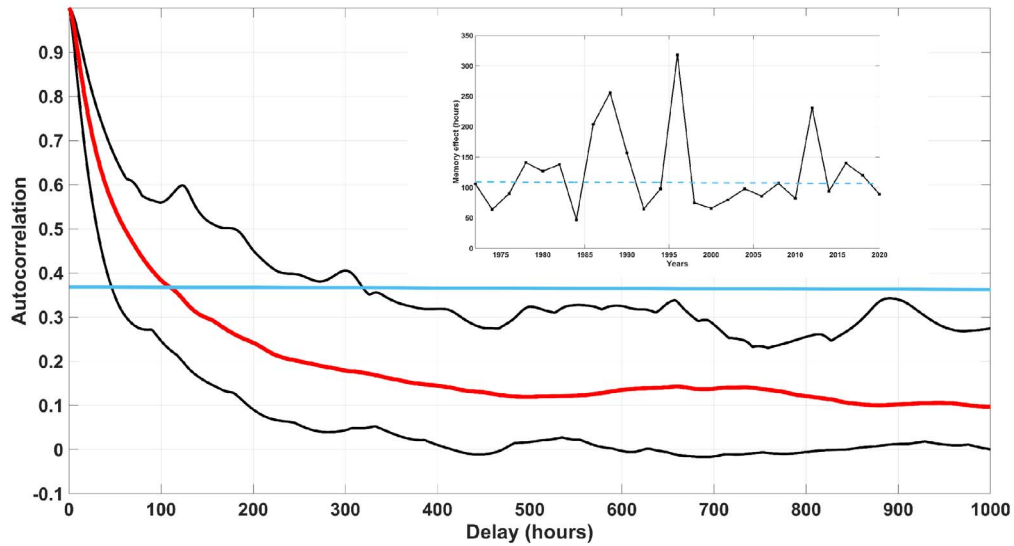
examination of the envelope curve of the autocorrelation functions puts in evidence a significant non-stationarity of the function when it is estimated over successive two-year periods. A significant variation of the memory effect is then highlighted with a global variation between 50 and 320 h.

The spectral analysis (Figure 5) of the hourly discharge estimations over the 1970–2022 period shows several power-law behaviours of the variance density spectrum. Due to the size of the time series, two spectral estimates are available: an estimation of the spectrum obtained from the whole available series with a maximum truncation equal to the length of the series divided by 3 [Mangin, 1984], and an estimation of the spectrum obtained by considering two-years series and thus a truncation of the order of 8 months. The first estimate correctly accounts for the annual hydrological cycle and the second estimate allows for identifying two power-law behaviours. Three intervals are highlighted on the variance density spectrum. For time scales of less than 2.5 h (i.e., a normalised frequency of  $0.4 \text{ h}^{-1}$ ), the spectrum is globally flat. This cut-off frequency is a little higher than the 30 min frequency previously highlighted, which is explained by the linear interpolation operations carried out on the first part of the chronicle where only the limnigraph was available. Thus, a power-law behaviour is visible between 4 h and 60 h (i.e. a normalised frequency of  $0.015 \text{ h}^{-1}$ ) followed by a second power-law behaviour over the interval from 60 h to 15 days (i.e., a normalised frequency of  $0.0025 \text{ h}^{-1}$ ). The high-frequency behaviour is characterised by a slope of 3.2, whereas the medium-frequency behaviour is characterised by a weaker

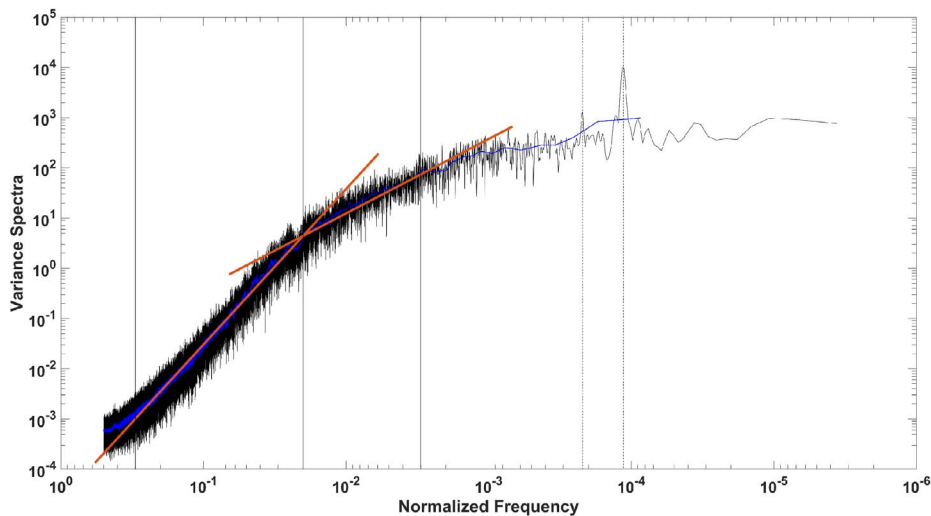
slope of around 1.5. Finally, two peaks corresponding to scales of the order of 6 months and one year are visible, classically corresponding to the influence of snow and the annual hydrological cycle. The multifractal analysis also shows a power-law behaviour (Figure 6) over all the scales from 4 h to 2 years. It is interesting to note that in previous studies [Labat *et al.*, 2002], a cut-off scale of around 8 days was posited. This difference is probably related to the size of the time interval. Thus, we are now able to propose a unique  $K(q)$  function for the Baget watershed and the corresponding multifractal parameters [Labat *et al.*, 2013] are now  $C_1 = 0.23$  and  $\alpha = 1.6$ . These two parameters are different from previous estimations. All spectral or multifractal parameters appear to be highly dependent on the length of the time series under consideration.

## 7. Conclusions and perspectives

The long-term analysis of time series is an essential step towards a better understanding of the changes occurring in hydrological systems under the simultaneous influence of climate change and various anthropogenic changes, particularly on land use. For highly reactive systems such as karst systems, the data usually available on a daily scale do not allow a high-resolution study of flood phenomena in particular. Within the framework of this study, we have proposed a method for digitising the limnigraphs, which often lie dormant in our archives. At the Baget site, it was possible to digitise these archives and thus propose an estimate of the hourly discharge time series fluctuations in this watershed



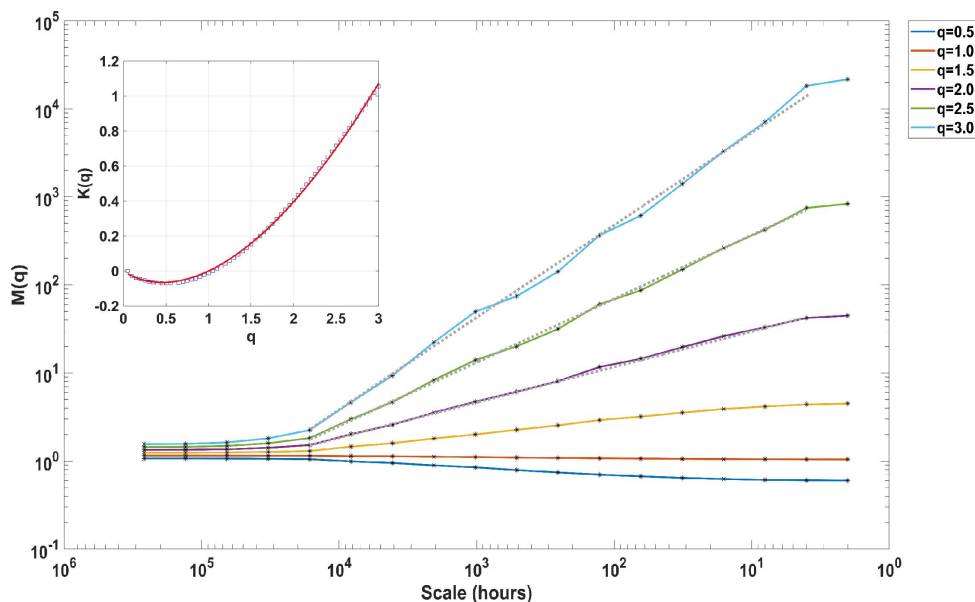
**Figure 4.** Correlation analysis of the hourly discharge time series. The red line corresponds to the correlation function obtained over the entire time interval. The two black lines corresponds to each time delay for the minimum and maximum value of the autocorrelation obtained over 25 intervals of 2 years. The memory effect obtained in these intervals is also depicted in order to highlight its variability over the last 50 years.



**Figure 5.** Spectral analysis applied to the entire interval (black) and at 2-years intervals (blue). The different power-law behaviors are put in evidence (red) together with the different scales of cut-off.

over the last 50 years. The correlation, spectral and multifractal analyses applied to this series confirm and specify the multi-scale and non-stationary nature of the hydrometric series. Other water level

gauges at strategic points of the basin (overflow, intermediate stations) are available and will soon be digitised to better understand the internal behaviour of the Baget basin, thanks to these archives of the



**Figure 6.** Multifractal analyses of the hourly discharge time series with the different power-law behavior of the empirical moments function and the corresponding  $K(q)$  function with the theoretical calibrated model.

past.

### Conflicts of interest

Authors have no conflict of interest to declare.

### Acknowledgements

The authors would like to thank the French Karst National Observatory Service (SNO KARST) initiative at the INSU/CNRS that aims to strengthen dissemination of knowledge and promotes cross-disciplinarily research on karst systems at the national scale.

### References

- Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2008). *Time Series Analysis: Forecasting and Control*. John Wiley & Sons Inc., Hoboken.
- Chen, Z., Auler, A. S., Bakalowicz, M., Drew, D., Griger, F., Hartmann, J., Jiang, G., Moosdorf, N., Richts, A., Stevanovic, Z., Veni, G., and Goldscheider, N. (2017). The world karst aquifer mapping project: concept, mapping procedure and map of Europe. *Hydrogeol. J.*, 25, 771–785.
- Debroas, E.-J. (2009). *Géologie du bassin versant du Baget*, volume 46. Association Strata, Toulouse.
- Ford, D. and Williams, P. (2013). *Karst Hydrogeology and Geomorphology*. John Wiley & Sons, Hoboken, NJ, USA.
- Gaillardet, J., Braud, I., Hankard, F., et al. (2018). OZ-CAR: The French network of critical zone observatories. *Vadose Zone J.*, 17, article no. 180067.
- Hartmann, A., Jasechko, S., Gleeson, T., Wasa, Y., Andreo, B., Barberá, J. A., Brielmann, H., Bouchaou, L., Charlier, J.-P., Darling, W. G., Filippini, M., Garvelmann, J., Goldscheider, N., Kralik, M., Kunstmann, H., Ladouche, B., Lange, J., Lucianetti, G., Martín, J. F., Mudarra, M., Sánchez, D., Stumpp, C., Zagana, E., and Wagener, T. (2021). Risk of groundwater contamination widely underestimated because of fast flow into aquifers. *Proc. Natl. Acad. Sci. USA*, 118, article no. e2024492118.
- Jourde, H., Massei, N., Mazzilli, N., et al. (2018). SNO KARST: A French network of observatories for the multidisciplinary study of critical zone processes in karst watersheds and aquifers. *Vadose Zone J.*, 17, 1–18.
- Labat, D. (2000). *Non linéarité et non stationnarité en hydrologie karstique*. PhD thesis, INP Toulouse.

- Labat, D., Ababou, R., and Mangin, A. (2002). Rainfall–runoff relations for karstic springs: multifractal analysis. *J. Hydrol.*, 256, 176–195.
- Labat, D., Hoang, C. D., Masbou, M., Mangin, A., Tchiguirinskaia, I., Lovejoy, S., and Schertzer, D. (2013). Multifractal behaviour of long-term karstic discharge fluctuations. *Hydrol. Process.*, 27(25), 3708–3717.
- Labat, D., Masbou, J., Beaulieu, E., and Mangin, A. (2011). Scaling behavior of the fluctuations in stream flow at the outlet of karstic watersheds, France. *J. Hydrol.*, 410, 162–168.
- Larocque, M., Mangin, A., Razack, M., and Banton, O. (1998). Contribution of correlation and spectral analyses to the regional study of a large karst aquifer (Charente, France). *J. Hydrol.*, 205, 217–231.
- Mangin, A. (1975). *Contribution à l'étude hydrodynamique des aquifères karstiques*. PhD thesis, Université de Dijon.
- Mangin, A. (1984). Pour une meilleure connaissance des systèmes hydrologiques à partir des analyses corrélatoire et spectrale. *J. Hydrol.*, 67, 25–43.
- Marsily, G. and Besbes, M. (2017). Les eaux souterraines. *Res. Environ.*, 86, 25–30.
- Mathevet, T., Lepiller, M., and Mangin, A. (2004). Application of time-series analyses to the hydrological functioning of an Alpine karstic system: The case of Bange-L'Eau-Morte. *Hydrol. Earth Syst. Sci.*, 8, 1051–1064.
- Padilla, A. and Pulido-Bosch, A. (1995). Study of hydrographs of karstic aquifers by means of correlation and cross-spectral analysis. *J. Hydrol.*, 168, 73–89.
- Quinif, Y. (1999). Karst et évolution des rivières: le cas de l'Ardenne. *Geodin. Acta*, 12, 267–277.
- Stevanović, Z. (2019). Karst waters in potable water supply: a global scale overview. *Environ. Earth Sci.*, 78, article no. 662.
- Tramblay, Y., Llasat, M. C., Randin, C., and Copola, E. (2020). Climate change impacts on water resources in the Mediterranean. *Reg. Environ. Change*, 20, article no. 83.







Research article

## Geo-hydrological Data & Models

# Geothermal heat advected by the recharge of underground conduit. Case study of the karstic spring of Lez (Hérault, France)

Guy Vasseur <sup>a</sup>

<sup>a</sup> Fondation CERGA: Centre pour l'Enseignement et la Recherche en Géosciences et Applications, Résidence Astérie A2, 740 rue Paul Rimbaud, 34080 Montpellier, France

E-mail: [guy.vasseur@upmc.fr](mailto:guy.vasseur@upmc.fr)

**Abstract.** The karstic flow system of the Lez spring is capturing a large proportion of the deep geothermal heat flow as observed by temperature measurements both above and below the underground karst network. For interpreting these data, an analytical model of heat/fluid interaction based on the conservation of mass and energy in a dual conduit/porous medium system is developed. In the model of energy transport, the energy of the fluid includes its enthalpy plus its gravity potential; the hydraulic head is shown to be the correct potential accounting for both pressure and gravity. The thermal features of the model are expressed as a function of a few parameters: from comparison with the actual data, the depth of the conduit appears to lie around 400 m and the number  $Pe$  characterizing the recharge is about 6. When estimated with available thermal data, the amount of geothermal energy captured by the flow system in steady state conditions is significantly lower than the actual energy output of the spring. The possible origin of this offset is discussed: effect of gravity potential, 3D convergence of geothermal heat flux lines, transient effect. Moreover, the mapping of the vertical temperature gradient at low depth indicates the general pattern of the recharge zone energy.

**Keywords.** Temperature, Analytical solutions, Groundwater recharge, Energy budget, Karst, Steady state, Gravity potential.

*Manuscript received 9 March 2022, revised 5 September 2022, accepted 12 September 2022.*

## 1. Introduction

The temperature of groundwater is a potential tracer of circulation that can provide insights into the dynamics of hydro-geological processes. When the pattern of the groundwater flow is diffuse, either through small pores of the rock matrix or through multiple fractures, the concept of representative elementary volume (REV) in porous media offers a conceptual framework for describing the heat and mass flow [de Marsily, 1995, p. 12, p. 42]. In such a REV, the water fluxes are described by an average quantity (the specific discharge or Darcy velocity) and the thermal state of the rock-water system is characterized by a single temperature. In sedimentary basins, heat

transfer in porous media explains satisfactorily the temperature field associated with large scale groundwater flow [Toth, 1963, Smith and Chapman, 1983]. Inversely, very deep temperature data, obtained in oil exploration boreholes for instance [Deming et al., 1992, Vasseur and Demongodin, 1995, Reiter, 2001, Pimentel and Hamza, 2012, Tavares et al., 2014], are currently used to constrain the large scale underground flow in deep aquifers or to derive the field of permeability [Saar, 2011]. In area of high geothermal potential, such as the Snake River Plain (USA), temperature profiles in boreholes and temperature gradients are useful to decipher the underground flow pattern and to infer its geothermal potential [McLing

et al., 2016, Lachmar et al., 2017]. Alternatively, they can be used to estimate the area of the recharge [e.g. Ingebritsen et al., 1989].

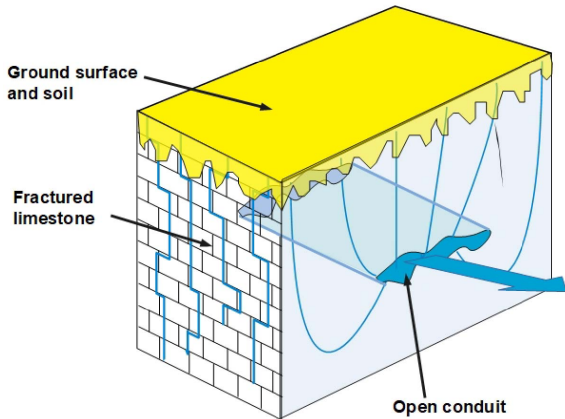
In contrast with diffuse type water flow, in karstic aquifers, heat (and fluid) transfer cannot be described simply by transfer in porous media because a major component of the porosity is composed of very large pores and open channels. These conduits are focusing the initial slow water flow of the surrounding fractured medium and present very active flow following the laws of hydraulics. In fact, the surrounding rocks which provide the water recharge toward the localized conduit flow [Atkinson, 1977], can generally be modelled as a porous medium; but inside the conduits, the water fluxes are best modelled by a network of pipes embedded in a porous rock matrix. Hybrid models have been used to simulate the transport of heat and reactive solute transport in coupled porous media and conduits [Liedl and Sauter, 1998, Liedl et al., 2003, Birk et al., 2006, Covington et al., 2011] using mathematical and numerical models.

Many observations show that, when observed elsewhere than directly in karst intersections, the vertical gradient of temperature in karstic areas is generally much lower than expected [Mathey, 1974, Luetscher and Jeannin, 2004]. This suggests that a large proportion of the regional geothermal heat flow may be captured by the water flow and eventually warms the out-flowing waters. Quantitative studies of heat flow processes within the saturated zone of karst systems are available [Benderitter et al., 1993, Lismonde, 2010, Badino, 2005]; these studies are mainly based on the temperature of the karstic conduit whereas the quantitative understanding of the physical interaction between the water flow and the heat flow requires to account not only for the temperature within the conduit, but also for the temperature field inside the surrounding rocks.

This study develops an analytic model describing the temperature field in a system based on the conceptual model of fluvio-karst described by White [2002, 2006]. This model is applied to the actual karstic zone of the Lez spring where a series of thermal data are available. The underground system of the Lez spring, located about 15 km North of the city of Montpellier, is a classical example of “karst barré” [or “damed karst”, Bakalowicz, 2005]; it gives birth to the Lez river which ends southward in the Mediter-

anean sea (28 km length). The watershed area feeding this spring has had relatively low structural deformation: the geometry of the corresponding underground karstic system is not directly known but it is likely to be relatively horizontal as is the general stratification. The Lez karstic network may thus be considered as a convenient example of underground sub-horizontal conduit.

The generic hydrological model used is illustrated on Figure 1 and is inspired from Ewers and Quinlan [1985]: rain infiltrates in the fractured medium, then merges toward a horizontal open conduit thereby contributing to the horizontal flux toward the natural outlet. Several simplifying assumptions are used: an idealized geometry is assumed and composed of a porous/fractured matrix connected to a horizontal conduit flow bounded by two horizontal planes as illustrated farther by Figure 8. The transverse dimensions of the conduit are small enough so that their internal temperature is nearly homogeneous. Following Bovardsson [1969] and Ziagos and Blackwell [1986] for open fractures—but also Vasseur et al. [2002] and Burns et al. [2016] for a thin horizontal aquifer instead of open conduit—the presence of such a horizontal concentrated flow can be accounted for through a specific boundary condition. This boundary condition, based on the energy and mass budget of the fluid system, is imposed on the limits of the porous medium. Only the steady state temperature is considered which implies that the actual recharge is simplified to an average constant rain. Moreover the porous medium is saturated and the fluid inside the conduit is assumed to be well mixed—or even turbulent—so that it is characterized by a single temperature. Using several other assumptions (mainly a very simple geometry of the flow and a negligible horizontal component of thermal conduction), it is possible to obtain simple analytical functions relating available hydrothermal data to a few hydrogeological parameters. This model is then applied to available observations of the Lez spring to obtain a better understanding of the interaction between the geothermal heat flow and the hydrological system.



**Figure 1.** Block-diagram inspired by Ewers and Quinlan [1985] showing the underground flow of water in karstic districts. The two types of water flow are illustrated in blue: slow infiltration in fractured rocks and relatively rapid flow in concentrated conduits.

## 2. Presentation of the Lez spring system and of its geothermal context

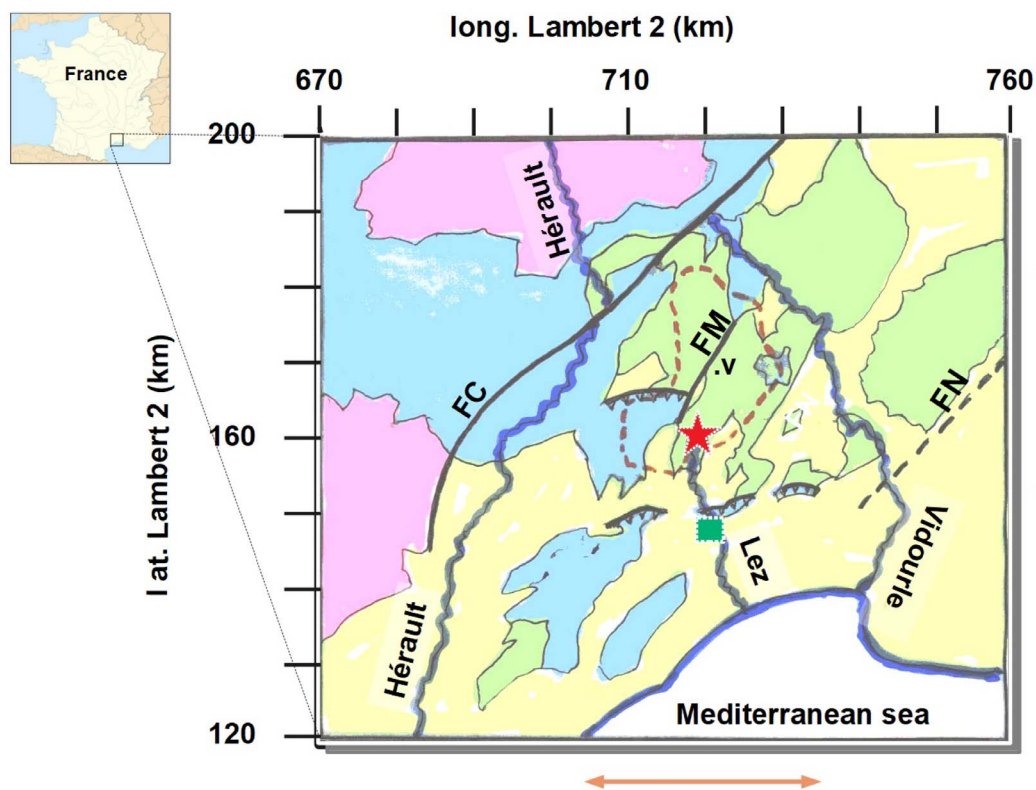
The general context of the Lez spring system has been intensively investigated because it has been for a long time and is still the main source of drinking water for the urban agglomeration of Montpellier [Drogue, 1969, 1985, Lacas, 1976, Avias, 1992, 1995, Guilbot, 1975, Uil, 1978, Bérard, 1983, Botton, 1984, Touet, 1985, Karam, 1989, Malard and Chapuis, 1995, Dausse, 2015].

### 2.1. General presentation of the system

The Lez karstic system belongs to the Mesozoic carbonate series of Languedoc which were deposited from Lias to Cretaceous on the north-western margin of the Tethyan Ocean, controlled by N-E trending faults such as the Cevennes fault (CF on Figure 2) and the Nimes fault (NF on Figure 2) [Husson, 2013]. During Cenozoic times, several geological events have imprinted the area: the Eocene Pyrenean orogeny followed by the Oligocene rifting of the gulf of Lion imposed a drainage system flowing southward and a reactivation of fault systems oriented North-East.

The result is a system of large faults (mainly normal to strike slip) in the North to North-East direction as well as thrusts trending in the East–West direction located North of the Montpellier city. The last transgression occurred during early Miocene followed by an uplift during which the base level falls of 1500 m. These episodes have sculpted a large karstic system in such a way as to adapt the drainage hydro-geological flow from mountains North of the Cevennes fault toward the Mediterranean sea to the topographic evolution [Avias, 1992, Bakalowicz, 2005, Husson, 2013].

The watershed of the Lez spring is part of the karstic hydrogeological system called the North Montpellier Garrigue unit, delimited by the Hérault river (to the West) and the Vidourle river (to the North and East). The geological setting and the major tectonic structures characterizing this unit are illustrated in Figure 2 from CERGH [1978]. The karstic aquifer layer has an estimated thickness of 650 to 800 m within the Upper Jurassic layers on both sides of the Matelles Fault (MF on Figure 2) with a thick succession of Lower Valanginian (Lower Cretaceous) marls and marly limestones; the aquifer itself straddles in the Upper Jurassic and the Berriasian (Lower Cretaceous). The karstic conduit has been recognized by diving with a cross sectional width of a few meters at a depth of about 100 m, and was explored only in the vicinity of the spring; the general geometry of the network remains speculative as largely deduced from geological observations and hydrological measurements such as dye, piezometer, chemistry... [Marechal et al., 2014]. Following a proposition by Dubois [1964] and using hydro-dynamical observations, Karam [1989] speculates on the existence of a single major karstic conduit running from North to South, along the major fault of “les Matelles” (“FM” on Figure 2). The major hydrogeological characteristics of the system have been summarized in Avias [1995]: with a natural overflow lying at 65 m (asl), the Lez spring is the main outlet of a large hydrogeological catchment; its area—illustrated on Figure 2—is estimated around 250 km<sup>2</sup>, based on geology, dye tracings, and groundwater level dynamics in the network of observation boreholes during the draw down [Ladouche et al., 2014, Thiery et al., 1983, Dausse, 2015, Marechal et al., 2014]. Its recharge zones include (i) the Jurassic limestone outcrop (between 80 and 100 km<sup>2</sup>), (ii) the Cretaceous marly



**Figure 2.** Simplified hydrogeologic map of the Languedoc karstic area with coordinates in km (Lambert 2 projection). The background illustrates the simplified geology: pre-Jurassic (basement to Lias) in pink, Jurassic in blue, Cretaceous in green and Cenozoic (Oligocene to Quaternary) in yellow. The thick lines are the main tectonic accidents: strike slip faults (*FC* for Cevennes fault, *FM* for Matelles fault and *FN* for Nimes fault) and overthrust faults are toothed black lines. The heavy blue lines show the main rivers (Hérault, Lez, Vidourle). The Lez spring is the red star and its recharge basin is shown by a dotted red line. The green square is for the Montpellier town, the symbol “v” for the Valflaunes meteorological station. The horizontal and the vertical orange scale bars present the limits in the EW and NS directions of the studied zone (restricted zone of Figures 3 and 5).

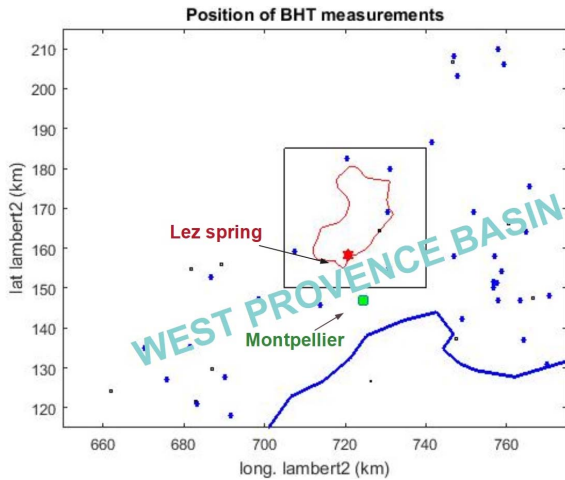
limestone formations through leakage toward the underlying Jurassic. Tertiary formations are generally impervious and contribute insignificantly to the karst aquifer recharge.

For more than 200 years, the Lez spring has been the main supply of drinking water for the city of Montpellier but the equipment for collecting the spring has evolved. Prior to 1968, the resource was exploited through spring overflow collection between 25 and 600 l/s [Paloc, 1979], then from 1968 to 1982, through pumping in the pan hole at a rate of some 800 l/s. Following this period, a so-called “active management” of the Lez karst system was undertaken with pumping rates in the summer periods

exceeding the natural discharge; from 1982, deep wells were drilled into the main karst conduit upstream of the spring in order to reach a maximum yield of 2000 l/s [Avias, 1995] through pumping units.

## 2.2. Estimate of the average volumetric and energetic discharge

The natural volumetric discharge of the karst system is not directly observed, because the Lez spring has been tapped since 1854. However, its pumping rate has been measured since 1974 and the spring’s residual overflow discharges were reliably measured between 1987 and 2007. From these observations,



**Figure 3.** Location of the deep exploration boreholes in the western part of the South Provence basin where BHT temperature measurements are available. The green square is the town of Montpellier, the red star is the Lez spring and the red curve bounds its surface recharge. A restricted  $35 \times 30 \text{ km}^2$  area around this recharge area is shown by a rectangular inset and corresponds to the range illustrated by orange arrows of Figure 2.

the yearly flow is about  $62 \times 10^6 \text{ m}^3 \cdot \text{y}^{-1}$  which corresponds to a mean of about  $2 \text{ m}^3 \cdot \text{s}^{-1}$ . In fact, during floods, the instantaneous flow rate may exceed  $10 \text{ m}^3 \cdot \text{s}^{-1}$ ; analyses of time series of water precipitation, flow rate and hydraulic head [Ladouche et al., 2014] emphasize the relatively inert behaviour of the system with respect to the infiltration of rain, since its characteristic time constant is on the order of 100 days.

A few records of the water temperature at the end of the spring (i.e. in the conduit itself) are available before the development of the underground pumping factory in 1982 [Lacas, 1976]: the average temperature was around  $15 \text{ }^\circ\text{C}$  with a slight seasonal variation: the temperature was nearing  $17 \text{ }^\circ\text{C}$  during low-water summer and tending toward lower values of about  $14.5 \text{ }^\circ\text{C}$  following strong rain episodes. With the new technical development of the spring tapping and pumping system, hourly records of the pumped water are now available from 2006 on; it appears that the temperature of the pumped water does still vary between  $14.5 \text{ }^\circ\text{C}$  and  $17 \text{ }^\circ\text{C}$  and is clearly anti-correlated

with the pumped flow rate. For the present study, possible artificial effects associated with the recent technical developments are ignored and the period around 1980 is chosen as reference one, when most of the present data were acquired. For this time, the mean annual temperature of the discharging spring is estimated to  $15.75 \text{ }^\circ\text{C}$ .

In order to obtain an energy budget of the karstic system, it is also necessary to assess the input temperature of the rain water. Meteorological observations are available at several locations of the watershed area [Thiery et al., 1983] providing monthly averages of precipitation and temperature. The average annual surface temperature is somewhat dependent on the altitude of the station varying from  $14 \text{ }^\circ\text{C}$  in its lower part (altitude 65 m) to  $12.5 \text{ }^\circ\text{C}$  in its upper one (altitude 200 m). The Valflaunes meteorologic station (altitude 135 m, station labelled “v” on Figure 2) located in the centre of the watershed has an average annual surface temperature of  $13.5 \text{ }^\circ\text{C}$ . Taking into account the fact that most rainy events occur during equinoxes, the average temperature of the rain is assumed to be equal to that of this later representative station (i.e.  $13.5 \text{ }^\circ\text{C}$ ). Another argument in favour of this choice is obtained from vertical profiles of temperature obtained in near surface boreholes of the area: as shown by Touet [1985], when vertical temperature profiles in the 30–100 m depth range can be extrapolated toward the ground surface, the extrapolated ground surface temperatures lies in the range ( $12.75 \text{ }^\circ\text{C}$ – $13.8 \text{ }^\circ\text{C}$ ). A value of  $13.5 \text{ }^\circ\text{C}$  is therefore assumed to be the average temperature of the ground surface (soil + water).

In conclusion, the average discharge temperature of the spring  $T_{\text{disch}} = 15.75 \text{ }^\circ\text{C}$  is  $2.25 \text{ }^\circ\text{C}$  higher than the assumed rain temperature (average  $T_{\text{input}} = 13.5 \text{ }^\circ\text{C}$ ). Taking into account the average flow rate of the spring ( $q = 2 \text{ m}^3 \cdot \text{s}^{-1}$ ), this implies that during its underground path the water flow has been receiving a mean energetic input  $E_{\text{spring}}$  given in watts (W) by:  $E_{\text{spring}} = (\rho C) q (T_{\text{disch}} - T_{\text{input}}) = 4.18 \times 10^6 \cdot 2.25 \cdot 2 \text{ W} = 18.8 \times 10^6 \text{ W} = 18.8 \text{ MW}$ , where ( $\rho C$ ) is the volumetric heat capacity of water. It is reasonable to assume that this heat exchange process occurs at steady state. Obviously this energetic contribution is mainly provided by the geothermal heat flux.

### 2.3. Deep regional geothermal heat flow

Over continents, heat flow data are generally obtained from temperature profiles measured in boreholes with a few hundreds of m depth; in particular, heat flow maps of France [Lucazeau and Vasseur, 1982, Bonte and Guillou-Frottier, 2006] show that the area of southern Languedoc is characterized by a regional heat flow of 80 to 90 mW·m<sup>-2</sup>. An essential condition for obtaining significant heat flow data is to avoid areas where heat convection occurs through active water flow; this precludes the use of such temperature profiles in karstic areas. However in most sedimentary basins, the oil industry has been drilling very deep boreholes for hydrocarbon exploration, reaching depths (several km) where active hydraulic fluxes are certainly much smaller; bottom hole temperatures (BHT) are recorded during well-log acquisition, being imposed more by technical purposes than by scientific ones. These rough BHT data suffer of biases reaching 5 to 15 °C due to thermal perturbations during the drilling operations but these data can be corrected leading to a set of un-biased but noisy temperatures [Vasseur et al., 1985].

A recent compilation of these data pertaining to the Provence basin (South-East of France) has been published showing a relatively consistent pattern of the deep temperature field [Garibaldi et al., 2010, Garibaldi, 2010]. For the present study, the analysis is focused on the western part of the Provence basin, i.e. west of 5 °E longitude and illustrated in Figure 3. The location of used wells (111 BHT data of which 89 at depth larger than 1000 m) is shown on Figure 3 together with a scatter plot of obtained temperatures versus depth on Figure 4. There is a clear linear increase of the temperature as a function of depth. The standard deviation (SD) of the departures between the linear fit and the data reach about 14 °C. The coefficients of the best square linear fit, with their SD writes as:

$$T(z) = 13.5 \text{ (s.d. 0.9) } ^\circ\text{C} + 0.0318 \text{ (s.d. 0.001)} z \text{ (m)}. \quad (1)$$

The value at the earth surface (13.5 ± 0.9 °C) agrees with the one previously estimated.

Figure 4 also presents the temperature versus depth plot for a more localized area covering the immediate vicinity of the catchment zone and implying a smaller number of data (14 BHT). When compared

to the previous linear trend, most data plot at a significantly lower temperature; indeed the regression of these data gives the following linear trend:

$$T(z) = 2 \text{ (s.d. 3.1) } ^\circ\text{C} + 0.0304 \text{ (s.d. 0.004)} z \text{ (m)} \quad (2)$$

pointing out a consistent slope but a smaller ground surface temperature than the whole data set. This offset of about 11 °C between the extrapolated ground temperatures of these two sets of data appears significant.

It is clear that a mean regional gradient ( $\gamma = dT/dz$ ) of 0.031 °C·m<sup>-1</sup> (or in practical units 3.1 °C/100 m) applies to deep temperatures in the zone of interest. With such a gradient, the regional basal heat flow—based on the assumption that conduction is dominating at great depth—can be evaluated as  $KdT/dz$  where  $K$  is the average thermal conductivity. The formations encountered in the various boreholes are mainly composed of carbonates and dolomites of Jurassic and Cretaceous age. The thermal conductivity  $K$  of such rocks can be evaluated from its composition and porosity; following Brigaud and Vasseur [1989] a value  $K = 2.75 \pm 0.5 \text{ W}\cdot^\circ\text{C}^{-1}\cdot\text{m}^{-1}$  is assessed. Assuming that conduction dominates at depth at this regional scale, these values imply a regional heat flow  $KdT/dz$  of about  $85 \times 10^{-3} \text{ W}\cdot\text{m}^{-2}$  in agreement with available heat flow maps of France [Lucazeau and Vasseur, 1982, Bonte and Guillou-Frottier, 2006].

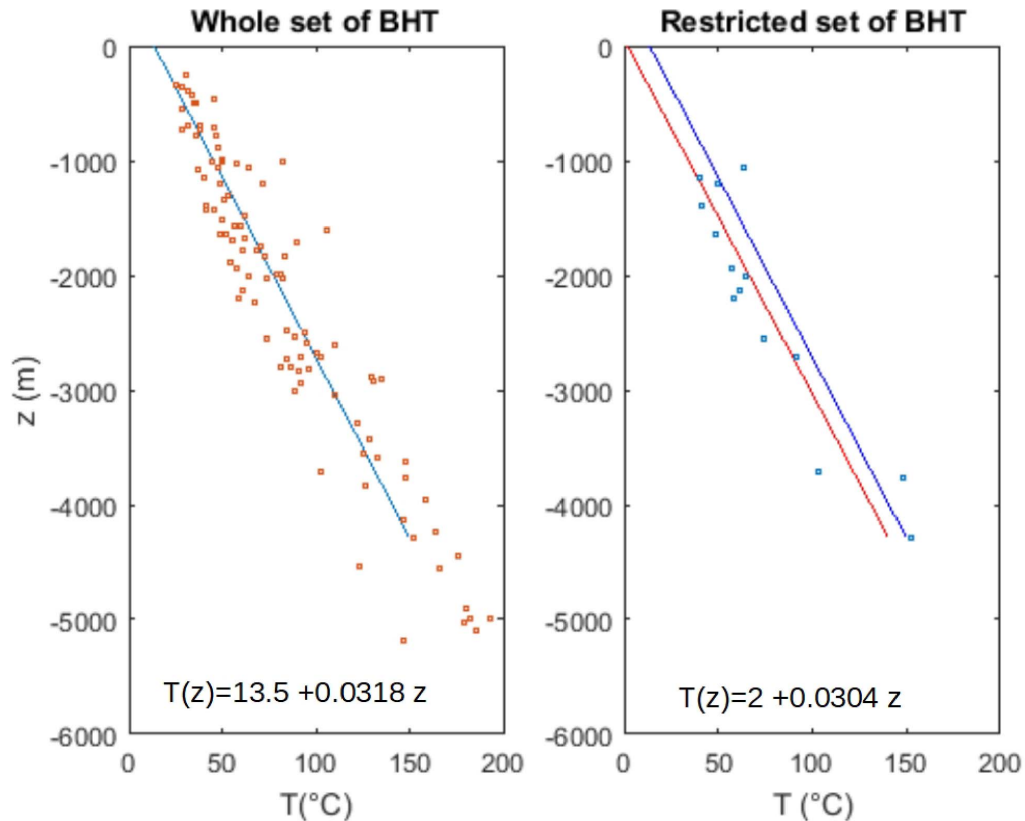
Over the catchment area (~250 km<sup>2</sup>), the integrated energy from this geothermal flux is:

$$\begin{aligned} E_{\text{geoth}} &= 0.085 \text{ W}\cdot\text{m}^{-2} * 250 \text{ km}^2 = 21.25 \times 10^6 \text{ W} \\ &= 21 \text{ MW}. \end{aligned} \quad (3)$$

To within experimental errors, this is very close to the mean energy  $E_{\text{spring}}$  needed for heating the spring which was estimated above from independent data. Assuming steady state conditions, everything happens as if the geothermal flux was completely absorbed for heating the water flowing in the upper crust of the Earth and emerging at the Lez spring. This absorption can indeed be verified experimentally through available temperature profiles observed in shallow wells.

### 2.4. Shallow temperature gradients at basin scale

Many temperature profiles in relatively shallow boreholes of the studied area have been obtained in the

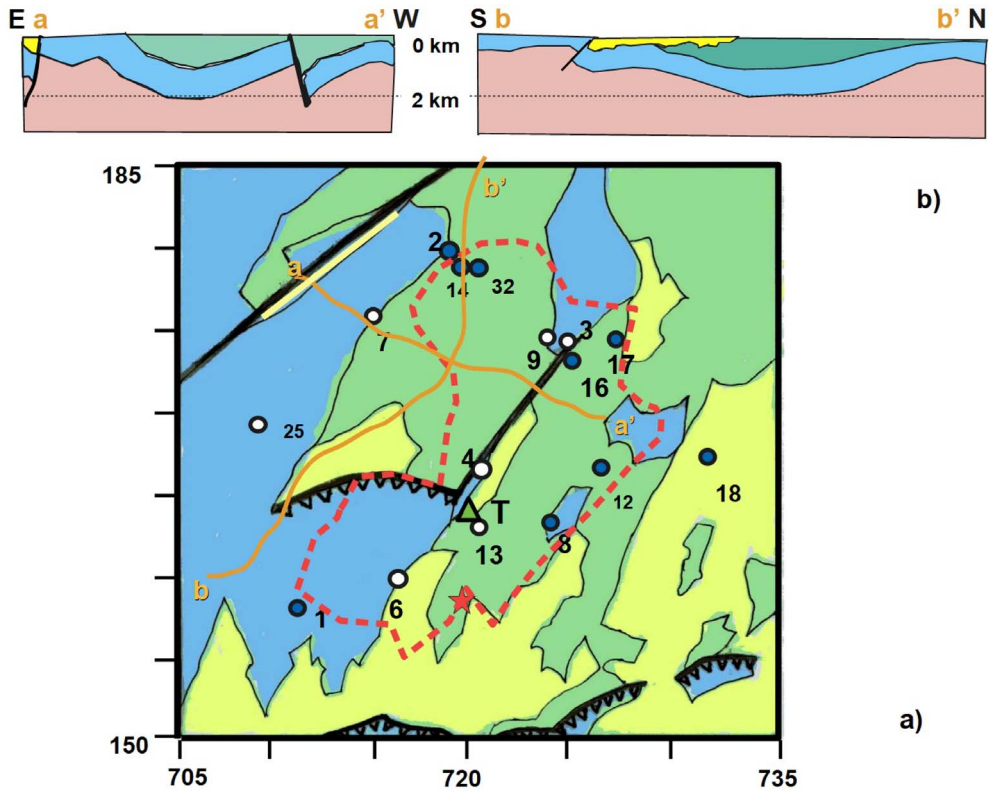


**Figure 4.** Plot of the measured BHT temperature versus their measurement depth ( $z$ ) and best linear fitting lines. The left plot is for the whole area of Figure 3 and the blue line for best fit has equation  $T(z) = 13.5 \text{ °C} + 0.0318z$ . The right plot deals with the restricted area of the inset of Figure 3 and the fitting line  $T(z) = 2 \text{ °C} + 0.0304z$ , in red, is compared with the previous one of the whole area, in blue.

80s: Uil [1978], then Botton [1984] and Drogue [1985] obtained measurements in several boreholes 60 m deep located in the immediate vicinity of the spring. A wide survey at large scale was provided by Touet [1985] from high precision temperature profiles in several superficial boreholes 50 to 200 m deep located in various parts of the catchment. Most of these profiles have reached the homo-thermic zone (where seasonal effects are very small) and many of them present a clear linear trend as a function of depth. However some of these temperature profiles reflect local disturbances due to local fluid flow such as spikes due to occasional fast fluid flow across boreholes that are easily recognizable, or to flow along boreholes between two levels that is more difficult to detect.

A careful examination of these profiles allows sort-

ing out those temperature profiles which present a regular linear trend versus depth—excluding profiles with spikes indicating lateral water arrival [Ge, 1998, Vasseur et al., 2002]. Moreover these profiles should exhibit an extrapolated ground surface consistent with meteorological data. About 17 thermal profiles have been retained, the location of which is illustrated on Figure 5 together with their geological background and two interpreted seismic cross sections of the area. Figure 6 displays seven of these profiles obtained in 1983 by Touet [1985]. The temperature gradients are all lower than  $3 \text{ °C}/100 \text{ m}$  (from 0 to  $2.9 \text{ °C}/100 \text{ m}$ ) with a mean value on the order of  $1.2 \text{ °C}/100 \text{ m}$ ; from their geographic distribution, it seems that smaller values (less than  $0.9 \text{ °C}/100 \text{ m}$ ) are obtained toward the centre and western part of the catchment whereas higher values (near  $2 \text{ °C}/100 \text{ m}$



**Figure 5.** (a) Location of near surface thermal gradient data [Touet, 1985] in their geological context [CERGH, 1978]. The colour code is the same as for Figure 2. The Lez spring and the limit of the basin are shown in red and the dots represent the location of individual thermal gradients with their label. Open symbols are for thermal gradient  $< 1 \text{ }^\circ\text{C}/100 \text{ m}$  and filled ones for  $> 1 \text{ }^\circ\text{C}/100 \text{ m}$ . The triangle  $T$  is the location of Terrieu. Orange lines  $a$ – $a'$  and  $b$ – $b'$  correspond to available seismic profiles. (b) Two cross sections along approximate NS and EW directions obtained from migrated seismic data as interpreted by Husson [2013].

or above) are observed toward its northern, eastern and southern edges. A trend analysis of these data, fitted with a polynomial of second degree in  $x$  and  $y$  as shown on Figure 7 confirms this trend: a zone of depleted gradient occurs toward the centre of the recharge zone.

### 3. A simplified model of karst circulation

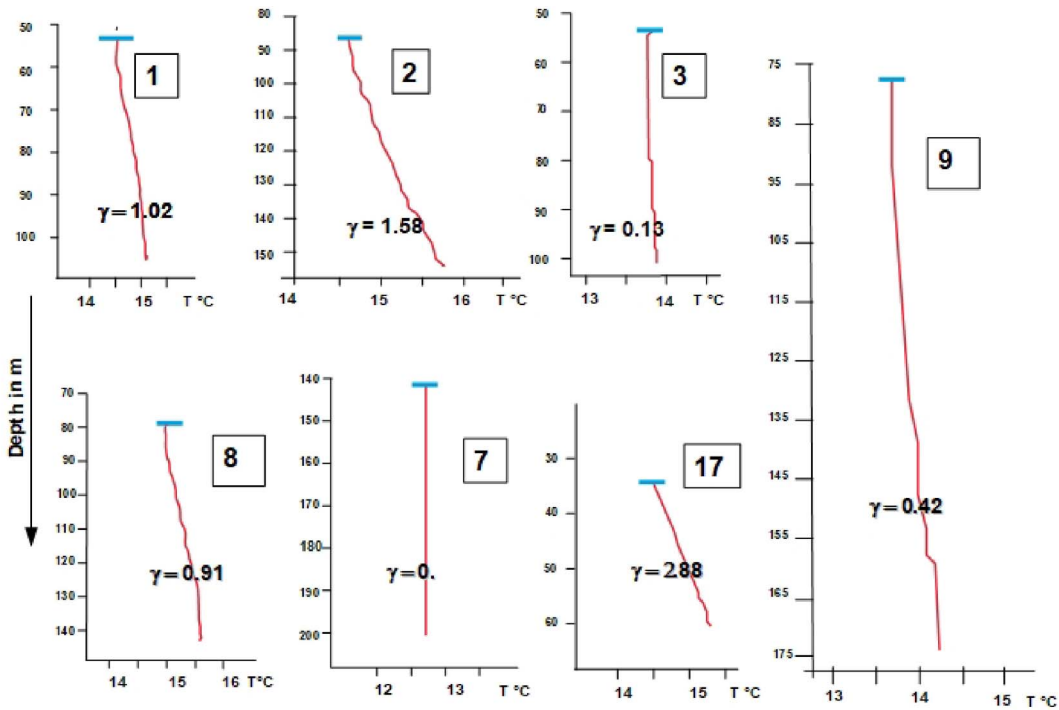
Based on the scheme described on Figure 1, a simplified model is designed where the karstic conduit, simulated by a horizontal fracture, is embedded in a porous medium simulating the fractured medium. This model, illustrated by Figure 8, reduces the problem to a pseudo 2-dimensional one and allows a sim-

ple analytical solution. Notations are summarized in Table 1.

#### 3.1. Basic assumptions

The reference frame illustrated on Figure 8 is such that  $Oxy$  is the ground surface assimilated to the water table,  $Oz$  the downward vertical, and  $Ox$  the horizontal direction parallel to that of the water flow inside the conduit. In the porous rock matrix ( $0 < z < h$ ), the porous water flow is characterized by the downward component of the specific discharge  $\vec{V}$  and its temperature  $T$  is in equilibrium with that of the rock.





**Figure 6.** Illustration of seven typical near surface thermal profiles with corresponding thermal gradient  $\gamma = dT/dz$  in  $^{\circ}\text{C}/100\text{ m}$ . The profiles measured in 1983 are redrawn in red from Touet [1985]. The ordinate is the depth from the ground surface; the level of the water table is shown as a blue line and the number in square inset correspond to the label located on Figure 5.

The thermal energy is transported by conduction in the porous medium with an energy flux given by— $K \text{ grad } T$  ( $K$  being the thermal conductivity of the porous medium) and also by convection of the moving fluid. The total energy per unit mass transported by this moving fluid is composed of its thermal energy written as  $CT$  ( $C$  the heat capacity per unit mass) plus other types of energy: internal fluid pressure, kinetic energy and potential energy. The necessary introduction of gravity potential in the total energy budget of porous flow has been pointed out in several theoretical studies [Sposito and Chu, 1981, Hasanizadeh and Gray, 1980]. More recently Stauffer et al. [2014]—see also Burns et al. [2016]—have enlightened the subject in practical applications; following Stauffer et al. [2014], the methalpy  $\theta$  sums up the specific enthalpy  $e = CT + p/\rho$  ( $p$  being the fluid pressure,  $\rho$  its density) plus its kinetic energy  $V^2/2$  (in fact negligible in porous media) and finally its po-

tential energy due to its evolution in the gravity potential i.e.  $-gz$  ( $z$  being directed downward). The methalpy is defined by:

$$\theta = CT + \frac{p}{\rho} + \frac{V^2}{2} - gz. \tag{4}$$

Assuming that the solid phase of the porous medium is rigid and in thermodynamic equilibrium with the fluid, the principle of energy conservation at steady state [Stauffer et al., 2014, Burns et al., 2016] implies that the divergence of the conductive flux  $-K\text{grad}(T) = -K\vec{\nabla}T$  in the porous plus the divergence of the convective flux of  $\theta$  is 0:

$$0 = \text{div}(-K\vec{\nabla}T + \rho\theta\vec{V}) = -K\Delta T + (\rho C)\vec{V} \cdot \vec{\nabla}T + \vec{V} \cdot \vec{\nabla}(p - \rho gz) \tag{5}$$

where  $\Delta T$  is the Laplacian of  $T$  (i.e.  $\text{div}(\text{grad}(T))$ ). For obtaining the last right hand side (rhs) of (5), use was made of the fact that  $\text{div}(\vec{V}) = 0$  (from fluid mass

**Table 1.** List of used symbols

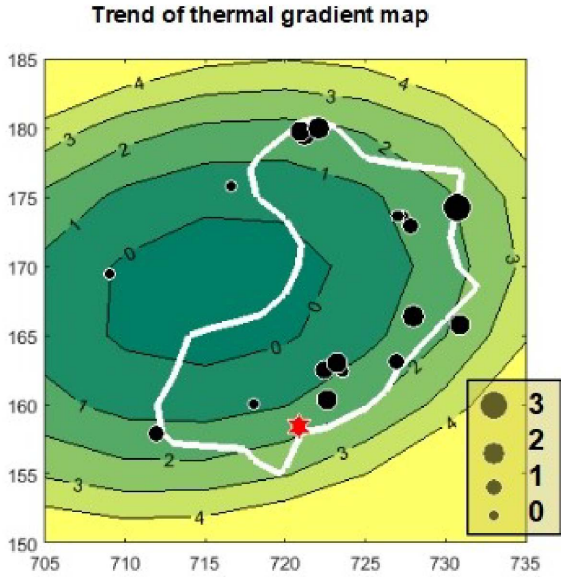
Symbol	Significance	Unit
$C$	Heat capacity of the water per unit mass	$\text{J}\cdot\text{kg}^{-1}\cdot\text{°C}^{-1}$
$E$	Energy output	W
$EE$	Function defined by $EE(u) = \text{erfc}(u) \exp(u^2)$	
$F_S$	Conductive heat flow at the ground surface	$\text{W}\cdot\text{m}^{-2}$
$G = g/C$	Thermal gradient associated with gravitational potential dissipation	$\text{°C}\cdot\text{m}^{-1}$
$H$	Hydraulic head = $p/(\rho g) - z$	m
$K$	Thermal conductivity of porous medium	$\text{W}\cdot\text{m}^{-1}\cdot\text{°C}^{-1}$
$L$	Length of the conduit ( $x$ direction)	m
$Pe$	Peclet number	
$R$	Radius of a horizontal disk	m
$S_f$	Shape factor	m
$T(x, z), T^+, T^-$	Temperature field in the porous medium, above $h$ , below $h$	$\text{°C}$
$T_0, T_{\text{ground}}, T_{\text{extrap}}$	Temperature in the conduit of the ground and extrapolated	$\text{°C}$
$V$	Specific discharge in $z$ direction. Eventually vector (Darcy vel.).	$\text{m}\cdot\text{s}^{-1}$
$e$	Specific enthalpy $e = CT + p/\rho$	$\text{J}\cdot\text{kg}^{-1}$
$g$	Acceleration of gravity ( $g = 9.81$ )	$\text{m}\cdot\text{s}^{-2}$
$h$	Depth of the conduit	m
$q$	Horizontal flux in the $x$ direction per unit $y$	$\text{m}^2\cdot\text{s}^{-1}$
$x$	Horizontal coordinate along the flow	m
$y$	Horizontal coordinate perpendicular to $x$	m
$t$	Time	s
$z$	Vertical coordinate orientated downward	m
$\alpha$	Vertical gradient of $H$	m/m
$\beta$	Horizontal gradient of $H$	m/m
$\varepsilon$	Thickness of the conduit	m
$\gamma$	Geothermal gradient	$\text{°C}\cdot\text{m}^{-1}$
$\rho$	Density of water	$\text{kg}\cdot\text{m}^{-3}$
$\kappa$	Diffusivity of the porous medium $(\rho C)_r / K$	$\text{m}^2\cdot\text{s}^{-1}$
$\theta$	Specific methalpy $\theta = e - gz$	$\text{J}\cdot\text{kg}^{-1}$
$(\rho C)$	Volumetric specific heat of water	$\text{J}\cdot\text{kg}^{-1}\cdot\text{°C}^{-1}$
$(\rho C)_r$	Volumetric specific heat of medium, of water	$\text{J}\cdot\text{kg}^{-1}\cdot\text{°C}^{-1}$
$\tau_{\text{diff}}, \tau_{\text{conv}}$	Time constant for diffusion, convection	s

conservation) and the kinetic term of  $\theta$  was omitted as negligible. The first term of the rhs arises from conduction, the second from water convection transporting internal heat and the last term acts as a source term which accounts for the transport of pressure and gravity potential variations. This last term is usefully expressed as a function of the hydraulic head arbitrarily defined with respect to the ground level

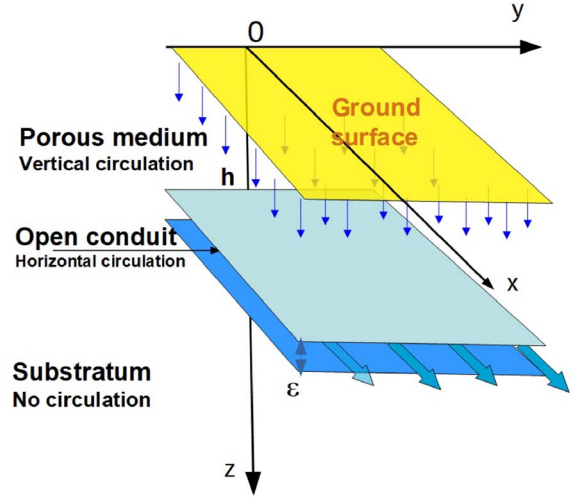
$z = 0$ :

$$H = \frac{p}{\rho g} - z. \quad (6)$$

The head  $H$ , i.e. the potential (in m) associated with the dynamics of porous flow through the Darcy law, is thus the relevant quantity for including the effect of gravity potential and pressure in the energy budget. The coupling of these two items is not astonish-



**Figure 7.** The large scale trend of the surface thermal gradient fitted by a polynomial surface of second degree in latitude–longitude on the same area as Figure 6. The values are in °C/100 m and the position of individual data is illustrated by dots. The size of dots is proportional to their value according to the bottom right legend (in °C/100 m).



**Figure 8.** Simplified geometry of the model used; the reference frame is  $Oxyz$ ,  $Oz$  being the downward vertical,  $Oxy$  the ground plane (yellow) and  $Ox$  the direction of the conduit flow. The conduit is a thin horizontal fault of thickness  $\epsilon$  at  $z = h$ . The water fluxes, illustrated by blue arrows, are vertical in the surrounding porous medium; they converge into the conduit to follow the  $Ox$  direction.

ing since variations of  $p/\rho$  can be seen as the internal response to variations of the altitude  $z$ . (This is opposite to a previous proposition of the author, who neglected the pressure term in (5); he was obviously wrong.)

Therefore, in the whole porous domain, the energy conservation in steady state conditions is expressed as:

$$K\Delta T - (\rho C)\vec{V} \cdot \vec{\nabla}T - (\rho g)\vec{V} \cdot \vec{\nabla}H = K\Delta T - (\rho C)\vec{V} \cdot \vec{\nabla}T - (\rho C)G\vec{V} \cdot \vec{\nabla}H = 0 \quad (7)$$

where  $G$  was introduced as the useful parameter:

$$G = g/C. \quad (8)$$

As seen from (7) this parameter  $G$  has a dimension of a temperature gradient and its value is  $9.8 \text{ m}\cdot\text{s}^{-2}/4180 \text{ J}\cdot\text{kg}^{-1}\cdot\text{C}^{-1} = 0.0023 \text{ }^\circ\text{C}\cdot\text{m}^{-1}$ .  $G$  is much smaller than the classical geothermal gradient  $\gamma$  which is in the order of  $3 \text{ }^\circ\text{C}$  per 100 m.

At depth  $z = h$ , as shown on Figure 8, lies a conduit assimilated to a “horizontal feature” with vertical thickness  $\epsilon$  ( $\epsilon \ll h$ ). This feature is in fact associated with a hydraulically conductive structure which is stretched in the  $x, y$  directions. It could even be a thin aquifer much more conductive than the overlying medium; in the following it is referred to as “plane conduit”. Inside this conduit, the horizontal water flow  $q(x)$  is the cumulative result from the downward flow occurring in the overlying porous medium (vertical component  $V$  along coordinates  $z$ ). At steady state, mass conservation imposes that the vertical specific discharge (Darcy velocity) is related to the increase along  $x$  of the water flow rate  $q$  (in  $\text{m}^2\cdot\text{s}^{-1}$  i.e. per unit length in the  $y$  direction) so that  $q$  and  $V$  are related by:

$$V = \frac{dq}{dx}. \quad (9)$$

A further approximation is that the horizontal conduction along the direction can be neglected since horizontal gradients are much smaller than vertical ones; nevertheless, the problem remains a 2D one

because thermal coupling along Ox is realized by the cumulative flow of heat and mass into the conduit. By hypothesis  $T$  does not depend on  $y$  and is written as  $T(x, z)$ .

In the porous medium located above the conduit ( $z < h$ ), the temperature field indexed as  $T^+(x, z)$  is submitted to the downward water flow and from (7) satisfies the following equation:

$$\text{for } 0 < z < h \quad K \frac{\partial^2 T^+}{\partial z^2} - (\rho C) V \frac{\partial T^+}{\partial z} - (\rho C) V G \frac{dH}{dz} = 0 \quad (10)$$

whereas, deeper than the conduit ( $z > h$ ), conduction is the only mechanism. There, the temperature field, indexed as  $T^-(x, z)$  satisfies:

$$\text{for } z > h \quad \frac{\partial^2 T^-}{\partial z^2} = 0. \quad (11)$$

Classical boundary conditions include imposed temperature at the ground surface (arbitrarily set to 0) and imposed vertical gradient  $\gamma$  (or given geothermal heat flux  $\varphi_{\text{geoth}} = K\gamma$ ) at great depth, so that:

- at  $z = 0$

$$T^+(x, 0) = 0 \quad (12)$$

- at  $z = +\infty$

$$\frac{\partial T^-}{\partial z}(x, +\infty) = \gamma. \quad (13)$$

At the level  $z = h$ , two specific boundary conditions are required: one accounts for the continuity of the temperature and the other for a possible discontinuity of its derivatives accounting for the energy conservation in the conduit. Since the thickness  $\varepsilon$  is very small, the temperature above and below are continuous (but not their vertical derivatives) so that:

- at  $z = h$

$$T^+(x, h) = T^-(x, h) = T_0(x) \quad (14)$$

where  $T_0$  is the temperature of the conduit. The fourth condition stems from the energy budget at the conduit level and involves not only the convective and conductive heat flow but also the pressure and gravity potential of the fluid phase. This is obtained using the surface integral equivalent of the energy conservation (7). For a volumetric element ( $dx dy \varepsilon$ ) of the plane conduit (Figure 9), the input of conductive plus convective energy including methalpy is 0 for steady state conditions:

- at  $z = h$

$$\begin{aligned} 0 &= -K \frac{\partial T}{\partial z} \Big|_{-}^{+} - (\rho C) \frac{d(qT_0)}{dx} + (\rho C) V T_0 \\ &\quad - (\rho C) G \frac{\partial(qH)}{\partial x} \Big|_{z=h} + (\rho C) G V H|_h \\ &= -K \frac{\partial T}{\partial z} \Big|_{-}^{+} - (\rho C) q \frac{dT_0}{dx} - (\rho C) G q \frac{\partial H}{\partial x}. \quad (15) \end{aligned}$$

Immediately to the right of the sign “0 =”, the first term containing the discontinuity of temperature gradient at  $z = h$  is the net conductive heat flow from above and below; the second and third terms are the convective heat input respectively in the horizontal direction and in the vertical one and the fourth and fifth ones the corresponding contributions of fluid pressure and gravity potential (both included in the head  $H$ ). The horizontal conductive term is omitted because it is negligible with respect to the horizontal convection. The simplification leading to the final expression arises from the relation  $V = dq/dx$ .

Since steady state is assumed, the energy budget of the conduit element equal to 0 provides the fourth boundary condition as (15). This boundary condition seems quite similar to the one stated by Bovardsson [1969], Lowell [1975], Ziagos and Blackwell [1986] for the case of a horizontal aquifer. In fact it differs because the energy contains more than internal heat as already stated by Burns et al. [2016]. But the main difference is that horizontal water flux  $q(x)$  is not constant with  $x$ : it is accumulating the water flux supplied vertically by the porous flow.

In the next paragraphs useful analytical solutions are obtained for cases where the parameters  $\alpha = -dH/dz$  in the porous medium, and  $\beta = -dH/dx$  in the horizontal conduit are assumed to be constant on the whole domain. Both parameters  $\alpha, \beta$  are related the dynamics of the underground water flow along the path. Since the flow is orientated downward and toward the spring,  $\alpha$  and  $\beta$  are strictly positive; This implies that  $-dH/ds$  (i.e.  $\alpha, \beta$ ) acts as a positive energy contribution in (15). Moreover they belong to the range  $[0, 1]$ ; whenever  $\alpha$  or  $\beta$  are very small, the variations of fluid pressure  $p$  are nearly compensating those of the gravity potential  $\rho gz$ ;  $H$  is then nearly constant along the water flow lines i.e. hydraulic equilibrium is nearly satisfied and the head contribution to the energy budget vanishes. On the contrary, when  $\alpha$  increases toward 1, the gradient of fluid pressure decreases with respect to that of the



Similarly, that of (18) with boundary condition (14) is:

$$T^-(x, z) = T_0(x) + \gamma(z - h). \quad (20)$$

These two solutions are continuous at the level of the conduit  $z = h$ : i.e.  $T^+(x, z = h) = T^-(x, z = h) = T_0(z)$ . However their derivatives are discontinuous so that the jump of conductive flux at  $z = h$ :

$$K \frac{\partial T}{\partial z} \Big|_{-z=h}^+ = K \left[ \frac{Pe}{h} \frac{T_0 - G\alpha h}{\exp(-Pe) - 1} - G\alpha + \gamma \right]. \quad (21)$$

Applying the fourth boundary condition, i.e. the energy budget described by (15) leads to the following equation for  $T_0$ :

$$\frac{(\rho C)V}{K} \frac{(T_0 - G\alpha h)}{1 - \exp(-Pe)} - \gamma + G\alpha + \frac{(\rho C)q}{K} \frac{d(T_0 - G\beta x)}{dx} = 0 \quad (22)$$

which can also be written as:

$$\begin{aligned} & \frac{(\rho C)V}{K} \frac{(T_0 - G\alpha h - G\beta x)}{1 - \exp(-Pe)} \\ & + \frac{(\rho C)q}{K} \frac{d(T_0 - G\alpha h - G\beta x)}{dx} \\ & = \gamma - G\alpha - \frac{(\rho C)V}{K} \frac{G\beta x}{1 - \exp(-Pe)}. \end{aligned} \quad (23)$$

This is a first order differential equation for  $T_0(x)$  which can be solved analytically in two specific cases of interest, depending on the assumption on  $V(x)$  (therefore  $q(x)$  and  $Pe(x)$ ).

### 3.2. $Pe(x)$ is $\gg 1$ in the entire domain of infiltration

Now, consider that the domain where vertical infiltration occurs is limited from  $x = 0$  to  $x = L$  so that  $q(0) = 0$  and  $q(L)$  is the final outflow rate. Let assume that in this range of  $x$  value,  $Pe(x)$  is large enough so that  $e^{-Pe(x)} \ll 1$  (in practice  $Pe > 3$ ), then the lhs of (23) can be modified to highlight a primitive:

$$\begin{aligned} & \frac{(\rho C)V(x)}{K} (T_0 - G\alpha h - G\beta x) \\ & + \frac{(\rho C)q(x)}{K} \frac{d(T_0 - G\alpha h - G\beta x)}{dx} \\ & = \gamma - G\alpha - G\beta \frac{(\rho C)V(x)x}{K}. \end{aligned} \quad (24)$$

It can be integrated from  $x = 0$  (where  $q(0) = 0$ ) to  $L$  leading to:

$$\begin{aligned} & \frac{(\rho C)q(L)}{K} (T_0 - G\alpha h - G\beta L) \\ & = (\gamma - G\alpha)L - \frac{(\rho C)}{K} G\beta \int xV(x)dx \end{aligned} \quad (25)$$

or:

$$T_0(L) = G\alpha h + G\beta(L - L_b) + \frac{(\gamma - G\alpha)h}{\langle Pe \rangle_L} \quad (26)$$

where  $L_b$  denotes the barycentre of the infiltration velocity ( $=L/2$  when  $V$  is homogeneous) and  $\langle Pe \rangle_L$  the average Peclet number over the range  $(0, L)$  both defined by:

$$\langle Pe \rangle_L = \frac{(\rho C)h}{K} \frac{q(L)}{L} \quad L_b = \frac{\int xVdx}{q(L)}. \quad (27)$$

These results can be applied for any  $x < L$  when replacing in (26)–(27)  $L$  by  $x$ .

### 3.3. $Pe(x)$ remains constant in the entire domain of infiltration

Assuming that  $V$  (and so  $Pe$ ) remains constant with  $x$  (without restriction on its value), the first order differential equation has a particular solution which writes:

$$\begin{aligned} T_0(x) = & G\alpha h + \frac{(\gamma - G\alpha)h}{Pe} (1 - \exp(-Pe)) \\ & + G\beta \left( \frac{1 - \exp(-Pe)}{2 - \exp(-Pe)} \right) x. \end{aligned} \quad (28)$$

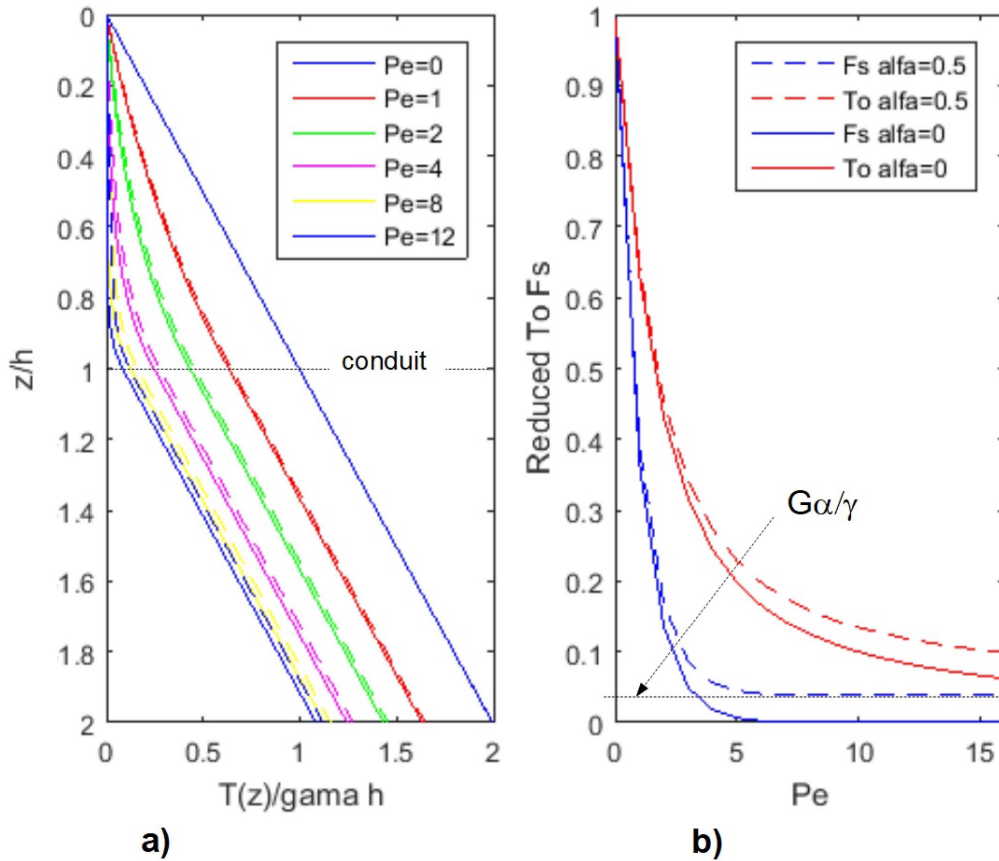
From (16)–(17) it corresponds to the vertical profile of temperature:

$$\begin{aligned} T^+(x, z) = & \left( \frac{(\gamma - G\alpha)h}{Pe} + \frac{G\beta x}{2 - \exp(-Pe)} \right) \\ & \times \frac{\exp(Pez/h) - 1}{\exp(Pe)} + G\alpha z \quad \text{for } z < h \end{aligned} \quad (29)$$

$$\begin{aligned} T^-(x, z) = & G\alpha h + \frac{(\gamma - G\alpha)h}{Pe} (1 - \exp(-Pe)) \\ & + G\beta \left( \frac{1 - \exp(-Pe)}{2 - \exp(-Pe)} \right) x + \gamma(z - h) \quad \text{for } z > h. \end{aligned} \quad (30)$$

The resulting solution  $T(x, z)$  is illustrated in non-dimension units ( $T/\gamma h$  and  $z/h$ ) on Figure 10a omitting its linear variation in  $x$  (i.e.  $\beta = 0$ ): vertical profiles of temperature are presented for various values of  $Pe$ . A comparison of  $T(\cdot, z)$  profiles obtained with  $\alpha = 0$  with those obtained with  $\alpha = 0.5$  illustrates the relatively small but substantial effect of this last parameter. The profile bending at  $z = h$  can be characterized by the intercept of the linear trend of the deep temperature  $T^-(x, z)$  when extrapolated at  $z = 0$ :

$$\begin{aligned} T_{\text{extrap}}(x) = & T_0(x, 0) - \gamma h = G\alpha h \\ & + \frac{(\gamma - G\alpha)h}{Pe} (1 - \exp(-Pe)) \\ & + G\beta \left( \frac{1 - \exp(-Pe)}{2 - \exp(-Pe)} \right) x - \gamma h. \end{aligned} \quad (31)$$



**Figure 10.** (a) Normalized temperature ( $T/\gamma h$ ) versus normalized depth ( $z/h$ ) for the plane conduit model in the established regime for various values of  $Pe$  and 2 values of  $\alpha$ . For emphasizing the effect of gravity-head potential, a strong value of  $\alpha$  (0.5) is chosen. Continuous lines are for  $\alpha = 0$  and dotted ones for  $\alpha = 0.5$  (i.e.  $G\alpha/\gamma = 0.039$ ). (b) Corresponding value of the normalized conduit temperature  $T_0$  ( $T_0/\gamma h$ ) and of the normalized surface heat flow  $F_S$  ( $dT/dz/\gamma$ ). Continuous lines are for  $\alpha = 0$  and dotted ones for  $\alpha = 0.5$  (i.e.  $G\alpha/\gamma = 0.039$ ).

At the ground surface, the “residual” conductive heat flow  $F_S$  is:

$$F_S = K \frac{\partial T}{\partial z} \Big|_{z=0} = K\gamma \exp(-Pe) + KG\alpha(1 - \exp(-Pe)) + KG\beta \frac{x}{h} \frac{Pe \cdot \exp(-Pe)}{(2 - \exp(-Pe))}. \quad (32)$$

The first term of the rhs is the deep geothermal heat flow, once attenuated by an exponential factor, the second and third ones accounts for the gravitational effect during the downward infiltration the final horizontal flow. When  $Pe \rightarrow \infty$   $F_S$  (and therefore the vertical thermal gradient) does not tend toward 0 but toward a threshold given by  $KG\alpha$ . It can be verified from (32) that the input of thermal energy for heating water from its initial temperature  $T = 0$  at the sur-

face to its value  $T_0(x)$  along the conduit is equal to the bottom heat flow minus the residual surface heat flow plus two components associated with gravity effects: the first is associated with the vertical path and the second one to the horizontal path as shown below:

$$(\rho C)VT_0 = \frac{(\rho C)Vh}{Pe}(\gamma - G\alpha)[1 - e^{-Pe}] + (\rho C)VG\alpha h + (\rho C)VG\beta \frac{(1 - e^{-Pe})}{(2 - e^{-Pe})}x = K\gamma - F_S + (\rho C)VG \left[ \alpha h - \frac{\beta x}{2 - e^{-Pe}} \right]. \quad (33)$$

Figure 10b depicts how  $T_0$  and  $F_S$  depends on  $Pe$ ; the  $x$  dependence is omitted (i.e.  $\beta = 0$ ) and both quantities are normalized to their undisturbed values  $\gamma h$

(for temperature) and  $K\gamma$  (for flux); the role of the gravity potential component is illustrated by comparison of two values either  $\alpha = 0$  and  $\alpha = 0.5$ . Both  $T_0$  and  $F_S$  decrease with  $Pe$ ; for high  $Pe$ ,  $T_0$  and  $F_S$  tend exponentially toward a limit ( $Gah$  and  $KG\alpha$  respectively) which becomes 0 when  $\alpha = 0$ .

3.4. Case of a recharge window ( $Pe = Cst$  for  $0 < x < L$  and  $Pe = 0$  elsewhere)

General solutions of (23) are now investigated for the case of a recharge window ( $V \neq 0$ ) limited to the range  $x = 0$  to  $L$  as illustrated by Figure 11a. Only the case  $\beta = 0$  is considered for sake of simplicity.

- for  $x < 0$ ,  $V = 0$  and the conduit is not recharged so that  $q(x) = 0$ ; thus the standard conductive equilibrium  $T(x, z) = \gamma z$  applies;
- for  $0 < x < L$ , water infiltration with constant  $V$  occurs; it is supplying the conduit flow resulting in a flow rate increasing linearly according to  $q(x) = Vx$ ;
- for  $x > L$ ,  $V = 0$ ; for  $x > L$ ,  $q(x)$  is now constant and becomes  $q(x) = VL$ .

For  $x < 0$ , the assumed initial conduit temperature is just  $T_{0ini} = \gamma h$ .

For  $0 < x < L$ ,  $T_0(x)$  is solution of:

$$x \frac{d(T_0 - Gah)}{dx} + \frac{(T_0 - Gah)}{1 - \exp(-Pe)} = \frac{(\gamma - G\alpha)h}{Pe} \quad (34)$$

The general solution of this first order differential equation is:

$$T_0(x) = Gah + \frac{(\gamma - G\alpha)h}{Pe} [1 - \exp(-Pe)] + Cst \left(\frac{1}{x}\right)^{\frac{1}{1 - \exp(-Pe)}} = T_0^* + Cst \left(\frac{1}{x}\right)^{\frac{1}{1 - \exp(-Pe)}} \quad (35)$$

where  $T_0^*$  is the particular temperature (28) previously described, also called established temperature. The constant  $Cst$  is defined by the boundary condition ( $T_0(0) = T_{0ini} = \gamma h$ ) at  $x = 0$ . In fact, due to the singular behavior of the solution at  $x = 0$ , this boundary condition cannot be imposed at all. This difficulty is associated with the oversimplification of zero horizontal conduction in the whole porous medium which results in cancelling any heat sink at  $x = 0$ . Nevertheless it can be qualitatively overcome according to the following approximate scheme.

It is possible to avoid the singularity of the solution (35) at  $x = 0$  by adding into (34) a dissipating term in the energy budget; this dissipating term accounts for

horizontal conduction along  $x$  limited to a horizontal strip of half thickness  $h/2$  around the conduit. The boundary condition now applies: after a few development detailed in Appendix A, for large  $Pe$ , the solution of the modified equation is now:

$$T_0(x) = T_0^* + (T_{0ini} - T_0^*)EE[\sqrt{Pe}x/h] \quad (36)$$

where  $EE(u)$  is the function given by  $EE(u) = \exp(u^2)\text{erfc}(u)$ .  $EE(u)$  is a classical function in heat transfer problems [Carslaw and Jaeger, 1959, p. 482]: it is equal to 1 for  $u = 0$ , behaves as  $1/u$  for large  $u$  where it can be roughly approximated  $1/(1 + 1.5u)$ . Since  $EE(u) = 1/e$  for  $u = 1.23$ , the horizontal distance  $z$  for which  $T_0(z) - T_0^* = (T_{0ini} - T_0^*)/e$  is roughly  $L_u \sim 2.18h/\sqrt{Pe}$ . This emphasizes the rapid convergence of  $T_0(z)$  toward the established regime  $T_0^*$  and especially for high  $Pe$ .

When for  $z > L$ , the recharge is stopped;  $q(x)$  remains constant and equal to  $VL$  but the model has to be adapted. This is because, in this case, Equation (10) becomes simply  $\partial^2 T^+ / \partial z^2 = 0$ : the vertical temperature profile is simply linear above and below the conduit. But, at the level  $z = h$ , the slopes of these two linear segments are different in order to compensate for the horizontal decay of the convective heat flow in the conduit. The heat budget of the conduit writes:

$$-(\rho C)q \frac{dT_0}{dx} - K \frac{\partial T}{\partial z} \Big|_{-}^{+} = -(\rho C)q \frac{dT_0}{dx} - K \left(\frac{T_0}{h} - \gamma\right) = 0 \quad (37)$$

or:

$$PeL \frac{dT_0}{dx} + T_0 = \gamma h \quad (38)$$

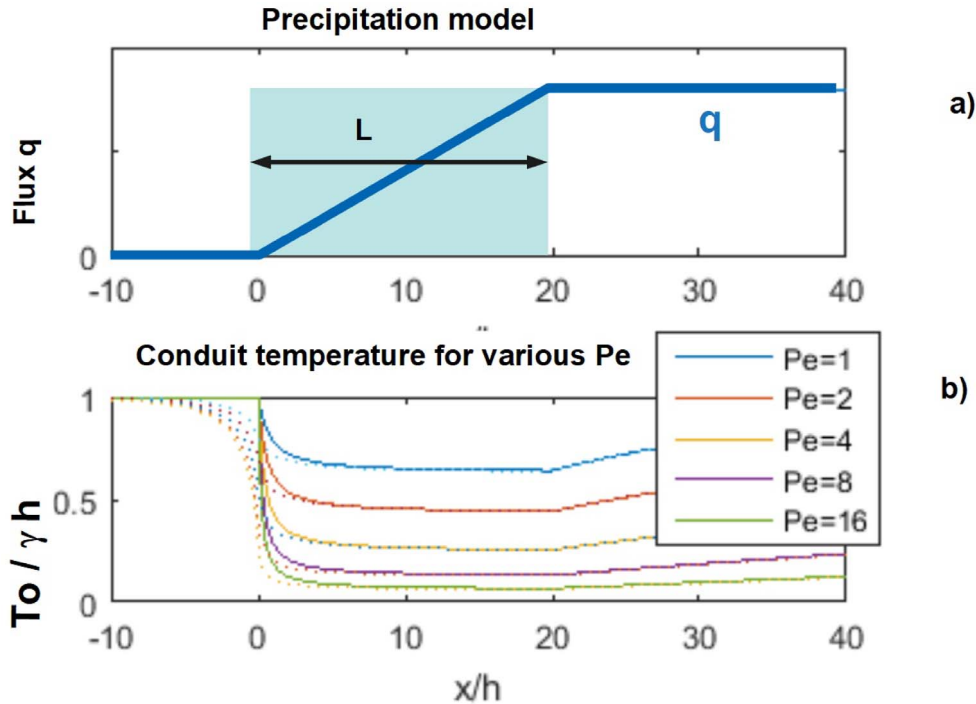
the solution of which is:

$$T_0(x) = \gamma h + (T_{0L} - \gamma h) \exp[-(x - L)/PeL] \quad (39)$$

with the boundary condition at  $z = L$  that  $T_0(L) = T_{0L}$  at  $z = L$  (if  $L \gg h$  for example,  $T_{0L} \sim T_0^*$ ). Therefore, when the recharge of the conduit flow stops, its temperature returns exponentially toward the undisturbed temperature  $\gamma h$  with a characteristic length  $PeL$ . These results, although obtained with an approximate account of lateral conduction, have been ascertained numerically with the help of numerical computations using finite differences and accounting for 2-D conduction as illustrated by Figure 11b.

The final result of this study is that, when infiltration supplying the conduit flow occurs abruptly in  $x$ , the established regime is rapidly dominant in





**Figure 11.** Cross section along  $x$  direction showing the effect of a limited infiltration ( $0 < x < L = 20$  h) recharging the conduit. (a) Represents the evolution of the water flux  $q(z)$  in the conduit and (b) plots the normalized conduit temperature  $T_0(x)$  for several  $Pe$  values. The dotted lines show the numerical value and the full lines the approximation of (36) of Section 3.4.

$x$  except in a narrow range of  $x$  values (on the order of  $h/\sqrt{Pe}$ ). Also when the supply is suppressed for  $x > L$ ,  $T_0$  decreases exponentially with a characteristic length  $PeL$  so that this decrease in  $x$  is very slow. Furthermore the integral solution described above allows to generalize these results for varying  $Pe$ ,  $Pe$  being replaced by its arithmetic average. Since in the actual case  $h \ll L$  and  $Pe > 1$ , the practical conclusion is that the actual thermal data of  $T_0$  can be directly compared to the prevision of the model through (26)–(28). These analytical results can now be compared to actual thermal data.

#### 4. Application of the model to Lez data; results and discussion

The available thermal data are: (1) the temperature of the Lez spring which is identified to  $T_0$  at the output of the system, (2) the near surface temperature gradients in shallow boreholes ( $Fs/K$ ) and (3) the deep

temperature trend extrapolated onto the ground surface ( $T_{\text{extrap}}$ ) from oil exploration data extrapolation. Some parameters are well known constants such as ( $\rho C$ ) =  $4.18 \times 10^6 \text{ W}\cdot\text{°C}^{-1}\cdot\text{m}^{-3}$ ,  $G = 0.23 \text{ °C}/100 \text{ m}$ ; others apply to the site (cf Section 2) such as rock conductivity  $K = 2.75 \text{ W}\cdot\text{m}^{-1}\cdot\text{°C}^{-1}$ , deep geothermal gradient  $\gamma = 3.1 \text{ °C}/100$ , length of the system  $L = 15,000 \text{ m}$ .

However the head gradients along the vertical path  $\alpha = -dH/dz$  and along the horizontal paths  $\beta = -dH/dx$  also need to be defined; in the following, their value is speculated on the basis of available piezometric data. Water level has been continuously monitored in about 20 piezometers beginning in the 1980s [Avias, 1995]. The observed difference of water levels between the upper part of the watershed and the spring is on the order of some tens of m and rarely exceeds 100 m [Karam, 1989, Avias, 1995]. This decrease of  $H$  along the water path is associated partly with the downward infiltration path (coefficient  $\alpha$ ) and partly with the horizontal path in

the conduit (coefficient  $\beta$ ); none of them is available from actual data. In order to appreciate the importance of the gravity-head potential term on the result, two extreme assumptions will be made:

- either  $\alpha$ ,  $\beta$  are both 0 or at least very small i.e. a quasi hydraulic equilibrium occurs along the path; this implies no contribution to the energy transport since the gravity potential effect is fully compensated by that of the water pressure;
- or a head decrease on the order of  $\delta H = 100$  m is occurring between the ground level and the spring; half of this decrease is assigned to the vertical path ( $\alpha = 0.1$  m/m for a vertical depth around 500 m as obtained below) and the other half to the horizontal one ( $\beta = 0.0067$  m/m for a horizontal length around 15 km). The fact that  $\alpha > \beta$  reflects the fact that the impedance of the conduit is obviously much smaller than that of the porous medium.

#### 4.1. Global parameters: use of $T_0$ and $T_{\text{extrapol}}$

The observations of  $T_0$  and  $T_{\text{extrapol}}$  are now used in the above 1D model to derive from (26)–(31) unknown parameters. These unknown quantities are  $h$  and  $Pe$ , but since  $Pe = (\rho C)Vh/K$ , it is clear the actual unknowns are  $h$  and  $V$ : in practice  $Pe$  may be expressed as  $Pe = 0.049Vh$  with  $h$  in m and  $V$  in m/year.

From the data  $T_0$  and  $T_{\text{extrapol}}$ ,—eventually expressed as deviations, with respect to  $T_{\text{ground}}$ —the value of  $h$  is easily computed using (31): the difference  $\gamma h = T_{\text{extrapol}} - T_0$  gives  $h = 440$  m. The average value of  $Pe$  is then obtained from (26). In the case  $\alpha = \beta = 0$  (26) reduces to  $\langle Pe \rangle = \gamma h / T_0 = 6$ . In the case  $\alpha = 0.1$ ,  $\beta = 0.01$ , Equation (26) leads to  $\langle Pe \rangle = 6.8$  which is significantly higher.

From these values of  $Pe$ , one can deduce the vertical specific discharge  $V = 0.28$  m·y<sup>-1</sup> in the first case ( $\alpha = \beta = 0$ ) and  $V = 0.31$  m·y<sup>-1</sup> in the second one. These values of  $V$  are quite consistent with the estimate deduced from the water budget at the emergence: in order to supply the Lez spring, ( $2$  m<sup>3</sup>·s<sup>-1</sup> i.e.  $63 \times 10^6$  m<sup>3</sup>·y<sup>-1</sup>), a precipitation area of  $63/0.28 = 225$  km<sup>2</sup> is required in the first case and  $63/0.315 = 200$  km<sup>2</sup> in the second one. This is smaller than the catchment area previously quoted ( $\sim 250$  km<sup>2</sup>) but not in contradiction with various estimates of the

area feeding the spring [Dausse, 2015, Marechal et al., 2014].

#### 4.2. Detailed interpretation of $F_S/K$ data

According to (33) the apparent surface heat flow  $F_S$ —the residual geothermal heat flow—is equal to the deep heat flow attenuated by a factor  $e^{-Pe}$  plus a contribution of the gravity-head potential which tends toward  $KG\alpha$  for  $Pe \rightarrow +\infty$ . The last contribution proportional to  $\beta$  and related to the conduit flow is probably very small and therefore neglected; therefore  $F_S/K$  should be larger than  $G\alpha$  (i.e.  $0.23\alpha$  in °C/100 m). Actually some of the measured thermal gradients are very small and even reach 0. It cannot be excluded that some of these measurements are biased by vertical water circulation inside the well but most measurements seem sound. Since the value of  $\alpha$  may be vanishing for these wells, experimental  $T(z)$  profiles don't provide any experimental proof for the occurrence of a positive lower bound for the temperature gradient which would indicate some significant positive value of  $G\alpha$  for large  $Pe$ .

Previous studies have already proposed that gravity potential dissipation does affect the thermal gradient in underground water system [Manga and Kirchner, 2004, Lismonde, 2010]. In particular, Luetscher and Jeannin [2004] compiled observations of vertical temperature gradients in several thick karst systems; these gradients tend toward a small but positive value on the order of  $0.3$  °C/100 m (i.e. close to  $G$ ) which could be associated with the dissipation of gravity potential. This result is clearly in contradiction with the present observations of thermal gradients but this also suggests that the temperature gradient was not estimated in the same conditions. According to the present model of saturated porous flow, a lower bound of thermal gradient near to  $G$  could occur only for very small fluid pressure (i.e.  $\alpha \sim 1$ ); therefore the occurrence of lower bound close to  $G$  would imply that, for such data, water is flowing at pressure conditions which remain close to the atmospheric pressure i.e. in unsaturated conditions at least during part of the flow circuit. Therefore differences in saturation conditions could explain the apparent paradox.

Anyway, the horizontal variations of present observed thermal gradients may be interpreted qualitatively in term of  $Pe$  according to  $Pe = -\log(F_S/K)$  so

that  $F_S/K$  reflects at least qualitatively the local value of the product  $Vh$ . According to measurements, a zone of very low gradient values and therefore of large  $Pe$  is observed on Figure 6. The wells with labels 3 Mas de Vedel (0.13 °C/100 m), 4 Bois de St Mathieu (0.5 °C/100 m), 6 les Matelles (0.13 °C/100 m), 7 Ferrieres (0 °C/100 m), 25 Frouzet (0.15 °C/100 m), 9 Claret (0.42 °C/100 m) have been drilled in Jurassic layers; all of them are located in recognized zones of water loss physically connected to the Lez spring as shown by dye tracing [Marechal et al., 2014]. This confirms that low thermal gradients areas correspond to high infiltration rate (i.e. high  $V$ ). Inversely, relatively high values of gradient ( $> 2$  °C/100 m) occur in the eastern and northern border of the catchment; this is consistent with the lack of connection with the spring as observed in these areas by dye tracing [Marechal et al., 2014]. Besides, the polynomial trend of Figure 7 suggests that the central zone of very low thermal gradient (high infiltration?) extends toward the West. This rapid comparison indicates the interest of high precision temperature measurements in shallow boreholes for mapping recharge zones.

#### 4.3. Global evaluation of integrated energy fluxes

Due to the small number and values of measured surface gradients, it is impossible to define a single value characterizing the near surface heat flow. Nevertheless an integrated value is available through surface integration of the observed values of  $F_S$  over the watershed area would reflect the integrated value of the residual geothermal flux—once part of it is absorbed for heating the water. Evaluation of  $E_{\text{resid}} = \int F_S$  requires the knowledge of  $F_S$  on the whole area. As the distribution of the available data is uneven and quite sparse, the trend displayed on Figure 7 is used and the surface integral over the whole domain is computed. The resulting residual energy  $E_{\text{resid}}$  is 5.8 MW. In the interpolation of Figure 7 some negative values of  $F_S$  have been obtained in the centre of the basin; when these values are forced to 0, the change in  $E_{\text{resid}}$  is very small (6.02 MW instead of 5.8). On Table 2 this integral is compared with other values of the energy associated with the system.

The difference between the geothermal flux of deep origin  $E_{\text{geoth}}$  and that observed at the surface  $E_{\text{resid}}$  is of the same order as the observed  $E_{\text{spring}}$  but

somewhat smaller by about 25%. It is therefore interesting to study whether the gravity head potential term could contribute significantly to the output temperature: according to the heat budget of the model (33), the component of gravity-head potential should be added as an extra energy to the difference  $E_{\text{geoth}} - E_{\text{resid}}$ . In the case of  $\alpha = 0.1$  which is an upper bound, for  $Pe = 6.8$  as suggested in Section 4.1, this component  $E_{\text{gravity}}$  amounts to 1.08 MW which is significant but clearly too small to explain the gap between  $E_{\text{spring}}$  and the captured geothermal heat flow  $E_{\text{geoth}} - E_{\text{resid}}$ . Many explanations for this gap may be suggested: they could proceed from errors in the energy budget such as the underestimate of  $E_{\text{geoth}}$  or the overestimate of  $E_{\text{spring}}$ . However this gap could also derive from some inherent simplification of the model; two of them are briefly discussed below.

A first explanation could be that the geothermal heat flow is in fact captured over an area larger than the assumed 250 km<sup>2</sup>, as a result of 3D concentration: under this effect, the deep flow lines of the geothermal heat flow would naturally deviate from vertical and tend to converge toward the underground exchanger associated with deep water circulation. As suggested by Badino [2005], this three dimensional effect can be estimated using standard conductive models developed for the design of heat exchangers. Let assimilate the characteristics of the exchanger (the deep conduit) to a horizontal disk with radius  $R = 8$  km (area  $\sim 200$  km<sup>2</sup>) located at depth  $h = 440$  m (simplified geometry required to evaluate its “shape factor”  $S_f = 4\pi R/\text{acotg}(R/2h)$ ) and submitted to a heat flow  $K\gamma$  of deep origin. Assume for simplicity that the imposed temperature is 0 at the ground surface and 0 at the disk surface. The linearity of the heat equation allows to superimpose two fields of temperature: (1) the field associated with a disk with imposed temperature  $-\gamma h$  and no heat flow at  $z = \infty$  plus (2) the field due to a single gradient  $\gamma z$ . The absorbed conductive heat flux transferred from the medium to the disk amounts to:

$$E = -KS_f\gamma h + K\pi R^2\gamma \quad \text{with } S_f = 4\pi R/\text{acotg}(R/2h). \quad (40)$$

The result is that the concentration factor increases the absorbed heat flow by only 1%; therefore this natural 3D effect contributes only slightly to the energy gap, except if the deep geothermal flux itself is heterogeneous and exhibits locally a positive anomaly

**Table 2.** Evaluation of the various integrated energy fluxes

Name	Explanation	Integrated energy (MW)	Corresponding area (km <sup>2</sup> )	Average energy flux (W·m <sup>-2</sup> )
$E_{\text{geoth}}$	Available geothermal flux over the catchment: $\int K\gamma dS$	21	250	0.085
$E_{\text{resid}}$	Residual thermal flux over the catchment; $\int F_S dS$	6	250	0.021
$E_{\text{geoth}} - E_{\text{resid}}$	$\int K\gamma dS - \int F_S dS$	15	250	0.064
$E_{\text{gravity}}$	Gravity head potential component $PeKG(\alpha h + \beta L/2)$	0 to 1.1	250	
$E_{\text{spring}}$	Output of the spring ( $\rho C$ ) $Q\Delta T$	19	-	-

(not detected by BHT measurements).

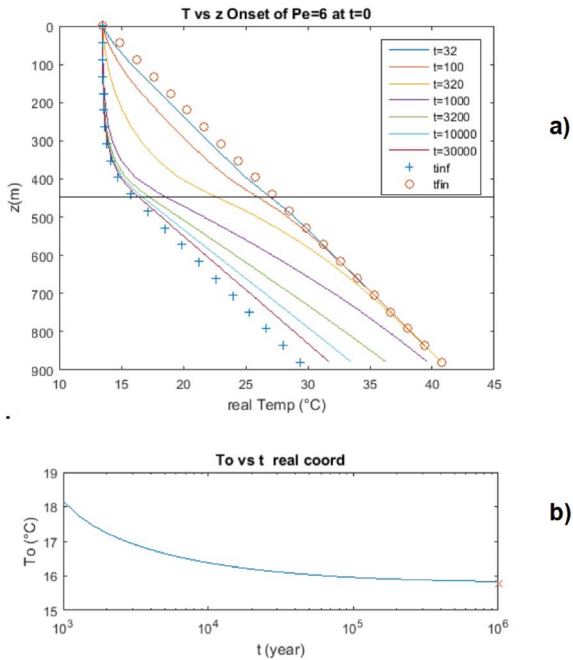
A second explanation implies possible transient effects due to external factors forcing such secular variations in water circulation. So far, the budget evaluation use the steady state assumption which rules out such possible environmental changes. The impact of sudden change such as the onset of water circulation can be studied using time dependent solutions of (10)–(15); the time dependent equations are written by replacing “= 0” in (10)–(11) by “=  $(\rho C)_r \partial T / \partial t$ ” and their transient solutions are obtained analytically by Laplace transform followed by numerical inversion. Assume, for example that, starting from an initial conductive equilibrium ( $Pe = 0$ ), the water flow is suddenly switched on with  $Pe = 6$  at time  $t = 0$ . From this epoch up to now, the temperature field evolves toward a new equilibrium and the transient response of the temperature profile propagates downward according to two characteristic times: the diffusive time  $\tau_{\text{diff}} = z^2 / \kappa = 3200$  y (where  $\kappa$  is the rock diffusivity  $\sim 2 \times 10^{-6}$  m<sup>2</sup>·s<sup>-1</sup>) and the convective time  $\tau_{\text{conv}} = z / V \sim 200$  y (for  $V = 0.5$  m·y<sup>-1</sup>). Instantaneous temperature profiles are shown in Figure 12a for various values of time following the onset of circulation; they illustrate how the upper layer is quite rapidly cooling whereas the temperature disturbance propagates downward slowly, in the conductive zone  $z > h$ . The evolution of the conduit temperature as a function of time is also illustrated by Figure 12b, showing that time for reaching the equilibrium temperature is in the order of 10,000 y; this emphasizes the delayed effect of very ancient modifications of the flow which could explain the apparent discrepancy derived from steady state assumptions (e.g. by

overestimating the output temperature of the spring and its energy output, since steady state is not yet reached).

## 5. Conclusion

Available geothermal data around the catchment of the Lez spring offer the opportunity of developing and of applying a geothermal model based on the principle of energy conservation. The energy source is mainly provided by the geothermal heat flux; this energy is exchanged with the fluid during its diffuse flow inside the karstic porous/fractured matrix connected to a conduit until the spring outlet where it is pouring. In order to assess quantitatively the nature and amount of energy captured by this natural heat exchanger, a simplified physical model is developed based on the universal laws of mass and energy conservation in steady state conditions. It leads to second order differential equations for the fluid temperature with boundary conditions which are specific of hydro-geothermal problems encountered in karstic areas. Owing to various simplifying assumptions (simple geometry of the exchanger, negligible horizontal heat conduction) the equations can be solved analytically; simple expressions for temperatures and heat fluxes are proposed as functions of  $Pe, h$  plus parameters associated with hydraulic gradients.

This model cannot pretend to account for the actual complexity of the karstic medium but it provides the parameters of a simplified system which is energetically equivalent to the actual flow. The offset observed between the actual ground temperature and the trend of the deep thermal data gives the depth



**Figure 12.** (a) Calculated profiles of  $T(z, t)$  resulting from the onset of water circulation ( $Pe = 6$ ) starting at time 0 from a conductive profile. Initial condition (conductive equilibrium) and final one (steady state) are shown as well as several intermediate profiles at various times. The parameters have been adapted to the Lez case ( $T_{\text{ground}} = 13.5 \text{ }^\circ\text{C}$ ,  $\gamma = 3.1 \text{ }^\circ\text{C}/100 \text{ m}$ ,  $h = 440 \text{ m}$ ,  $Pe = 6$ ). (b) Resulting temperature  $T_0$  at  $z = h$  evolving as a function of actual time in  $y$ .

$h \sim 440 \text{ m}$  of the conduit. A mean Peclet number  $Pe \sim 6$  can be evaluated from the excess of the spring temperature with respect to that of the rain. An energy budget integrating the various energy fluxes—i.e. (1) deep geothermal flux, (2) residual surface flux deduced from shallow temperature profiles and (3) heat advected at the spring—is presented. The difference between the two first fluxes is insufficient to account for the third one so that a further component of energy is needed. The energy component associated with the dissipation of gravity potential is too small; other explanations are discussed such as 3D effects in the geothermal flux and the occurrence of transient effects due to environmental changes.

In this work, specific attention is directed to the various energy sources affecting fluids, the potential

gravitational energy, in particular. Following Stauffer et al. [2014], the correct energy balance transports methalpy, which includes enthalpy (i.e. thermal energy plus fluid pressure) plus the potential energy of gravity. The other terms (kinetic energy and friction energy) are omitted as negligible for porous flow. In water saturated media, the conduction–advection equation for energy involves the transport of hydraulic head  $H$  which includes both pressure and gravity potentials; since the vertical fluid pressure gradient tends to compensate the effect of gravity, the contribution of  $H$  to the energy balance is much smaller than that of gravity potential alone (which would, on its own, result in a component of vertical thermal gradient up to  $G = 0.23 \text{ }^\circ\text{C}/\text{km}$ ) but nevertheless significant. This concept is tentatively applied to the model (and data) of the Lez spring at two levels. Firstly, at the global level, the contribution of gravity-head term is evaluated and could amount to less than 5% of the total energy contribution. Secondly, at the level of single thermal profiles, the gravity and head contribution would eventually result in a lower bound for thermal gradient affected by downward flow; however, from actual thermal profiles obtained in the area, no such lower bound of thermal gradient was evidenced suggesting that the gravity head contribution is not so important.

Data presented here have been obtained in the 80s before recent active management of the Lez output; a systematic campaign of measurements would certainly permit actualization and specify the regional pattern of residual heat flow.

Further investigations and in particular, high precision thermal loggings are required to answer this question definitively, as well as the role of gravitational potential.

## Conflicts of interest

The author has no conflict of interest to declare.

## Acknowledgments

This work was initiated following discussions with Michel Bakalowicz, Francis Lucazeau, Véronique Leonardi and Hervé Jourde. The development of the mathematical model received assistance from Pierre Genthon. The possible importance of the gravitational potential was pointed out by Patrick

Lachassagne and resulted in discussions with Patrick Goblet, Emmanuel Ledoux and Ghislain de Marsily. The author apologizes for committing a basic error in a previous version of the manuscript while omitting the influence of pressure in the energy budget.

Most of the data come from the collections of CERGA (Centre pour l'Enseignement et la Recherche en Géosciences et Applications) which has been recording and storing most of the hydrogeological studies on the Lez karstic system; the help of José Gravellec, Pierre Bérard, Jean Pierre Faillat and Jean Christophe Marechal is gratefully acknowledged.

## Appendix A. Analytical solution for a window of the recharge

The singularity of the solution of (23) is associated with the assumption of null horizontal conduction which prevents heat from dissipation at  $x = 0$  in the recharge shown on Figure 11a. To avoid this, a convenient approximation consists in modifying the heat budget of the conduit by introducing a further dissipative term; this term simulates horizontal conduction in a strip of porous medium of thickness  $h'$  around the conduit. This dissipative term is then— $h'K\partial^2 T/\partial x^2$  and the heat budget is modified according to:

$$h'K \frac{d^2(T_0 - G\alpha h)}{dx^2} - (\rho C)_e q \frac{d(T_0 - G\alpha h)}{dx} = K \frac{Pe(T_0 - G\alpha h)}{h(1 - e^{-Pe})} T_0 - K(\gamma - G\alpha). \quad (A 1)$$

With  $q = Vx$ , this equation becomes:

$$\frac{d^2(T_0 - G\alpha h)}{dx^2} - \frac{Pe}{hh'} x \frac{d(T_0 - G\alpha h)}{dx} - \frac{Pe(T_0 - G\alpha h)}{hh'(1 - e^{-Pe})} = \frac{1}{h'}(G\alpha - \gamma). \quad (A 2)$$

The general solution is the sum of the solutions  $T_0^* = G\alpha h + (\gamma - G\alpha)h[1 - \exp(-Pe)]/Pe$  defined in the main text—cf (28) with  $\beta = 0$ —plus a combination of two independent solutions of the corresponding homogeneous equation. Using changes  $f = (T_0 - G\alpha h)$ ,  $\omega = Pe/(hh')$  and the change of variable  $x' = x\sqrt{\omega}$ , the homogeneous equation (without second member) becomes:

$$f'' - x'f' - \frac{1}{(1 - \exp(-Pe))} f = 0. \quad (A 3)$$

With the transformation  $ff(x') = \exp(x'^2/4)f(x')$ , this equation takes the canonical form:

$$ff'' - ff' \left[ \frac{x'^2}{4} + \frac{1}{(1 - \exp(-Pe))} - \frac{1}{2} \right] = 0. \quad (A 4)$$

Two independent solutions of this equation are the parabolic cylinder functions  $u(\mu, x')$  and  $v(\mu, x')$  [also called Weber functions by Temme, 2010] which depends on the parameter  $\mu = 1/[1 - \exp(-Pe)] - 1/2$ . Of these two functions, only  $u$  has a correct (finite) behaviour for  $x' \rightarrow \infty$ . For large  $Pe$ ,  $\mu \sim 1/2$  and  $u(1/2, x') = \sqrt{\pi/2} \exp(x'^2/4) \operatorname{erfc}(x'/\sqrt{2})$ . Returning to the (33), the solution which satisfies  $T_0(0) = T_{0i} = \gamma h$  and  $T_0(x) \rightarrow T_0^*$  when  $x \rightarrow \infty$ :

$$T_0(x) = T_0^* + (T_{0i} - T_0^*) \exp\left(\frac{Pe}{2hh'} x^2\right) \operatorname{erfc}\left(\sqrt{\frac{Pe}{2hh'}} x\right). \quad (A 5)$$

A convenient rational approximation is:

$$T_0(x) = T_0^* + \frac{(T_{0i} - T_0^*)}{1 + 1.5x\sqrt{Pe/2hh'}}. \quad (A 6)$$

Numeric tests solving the 2D equation with the actual lateral conduction suggest choosing  $h' = h/2$  so that  $\omega = 2Pe/h^2$ . This solution can be extended to the case where  $Q = Q_0 + Vx$  i.e. a non zero conduit flow  $Q_0$  at  $x = 0$ , by replacing  $x$  by:  $x - Q_0/V$ .

## References

- Atkinson, T. C. (1977). Diffuse flow and conduit flow in limestone terrain in the Mendip Hills, Somerset (Great Britain). *J. Hydrol.*, 35, 93–110.
- Avias, J. (1992). Contrôles géologiques des systèmes aquifères karstiques (s.a.k) de type méditerranéen : l'exemple du S.A.K de la source du Lez-France-Karstic aquifers of Mediterranean type, geological controls: Lez Spring (North-Montpellierian karsts, France) example. In *Hydrogeology of Selected Karst Regions*, volume 13, pages 89–113. IAH, Reading.
- Avias, J. (1995). Gestion active de l'exsurgence karstique de la source du Lez (Hérault, France) 1957–1994. *Hydrogéologie*, 1, 113–127.
- Badino, G. (2005). Underground drainage systems and geothermal flux. *Acta Carstologica*, 34, 277–316.
- Bakalowicz, M. (2005). Karst groundwater; a challenge for new resources. *Hydrogeol. J.*, 13, 148–160.

- Benderitter, Y., Roy, B., and Tabbagh, A. (1993). Flow characterization through heat transfer evidence in a carbonate fractured medium: first approach. *Water Resour. Res.*, 29, 3742–3747.
- Bérard, P. (1983). Alimentation en eau de la ville de Montpellier. Captage de la source du Lez, commune de St-Clément (Hérault). Etude des relations entre la source et son réservoir aquifère. Rapport no 2.
- Birk, S., Liedl, R., and Sauter, M. (2006). Karst spring responses examined by process-based modeling. *Ground Water*, 44, 832–836.
- Bonte, D. and Guillou-Frottier, L. (2006). Nouvelles cartographies du flux de chaleur et des températures profondes en France métropolitaine. BRGM report.
- Botton, R. (1984). *Etude de certaines modalités du fonctionnement de l'aquifère karstique (zone d'infiltration et zone saturée) sur deux champs de forages Nord Montpelliérains*. PhD thesis, Université Montpellier 2.
- Bovardsson, G. (1969). On the temperature of water flowing through fractures. *J. Geophys. Res.*, 74, 1987–1992.
- Brigaud, F. and Vasseur, G. (1989). Mineralogy, porosity and fluid control on thermal conductivity of sedimentary rocks. *Geophys. J.*, 98, 525–542.
- Burns, E. R., Ingebritsen, S. E., Manga, M., and Williams, C. F. (2016). Evaluating geothermal and hydrogeologic controls on regional groundwater temperature distribution. *Water Resour. Res.*, 52, 1328–1344.
- Carslaw, H. S. and Jaeger, J. C. (1959). *Conduction of Heat in Solids*. Clarendon Press, Oxford, 2nd edition.
- CERGH (1978). Carte hydrogéologique du Languedoc et du Roussillon. Available from CERGA, Résidence Asterie A2, 740 rue Paul Rimbaud, 34080 Montpellier, France.
- Covington, M. D., Luhmann, A. J., Gabrovsek, F., Saar, M. O., and Wicks, C. M. (2011). Mechanisms of heat exchange between water monitoring of ground rock in karst conduits. *Water Resour. Res.*, 47, article no. W10514.
- Dausse, A. (2015). *Facteurs d'échelle dans la hiérarchisation des écoulements au sein d'un aquifère karstique - Analyse multi-échelles des propriétés hydrodynamiques et de transport de l'aquifère du Lez*. PhD thesis, Université Montpellier 2.
- de Marsily, G. (1995). *Quantitative Hydrogeology*. Academic Press, New York.
- Deming, D., Sass, J. H., Lachenbruch, A. H., and Derrito, R. F. (1992). Heat flow and subsurface temperature as evidence for basin-scale groundwater flow, North Slope of Alaska. *Geol. Soc. Am. Bull.*, 104, 528–542.
- Drogue, C. (1969). *Contribution à l'étude quantitative des systèmes hydrogéologiques karstiques d'après l'exemple de quelques karsts périméditerranéens*. PhD thesis, Université Montpellier 2.
- Drogue, C. (1985). Geothermal gradient and ground water circulation in fissured and karstic rocks: the role played by the structure of the permeable network. *J. Geodyn.*, 4, 219–231.
- Dubois, P. (1964). Les circulations souterraines dans la région Nord de Montpellier. *Bull. B.R.G.M.*, 2, 1–31.
- Ewers, R. and Quinlan, J. F. (1985). Groundwater flow in limestone terranes: strategy national and procedure for reliable efficient monitoring of groundwater quality in karst area. In *National Symposium and Exposition on Aquifer Restoration Ground Water Monitoring, Columbus, Ohio*.
- Garibaldi, C. (2010). *Détermination des températures profondes du bassin du Sud-Est de la France et relations entre anomalies thermiques, géologie et circulations hydrothermales par modélisation 3D*. PhD thesis, Université Nice-Sophia Antipolis.
- Garibaldi, C., Guillou-Frottier, L., Lardeaux, J. M., Bonte, D., Lopez, S., and Bouchot, V. (2010). Thermal anomalies and geological structures in the Provence basin: implications for hydrothermal circulations at depth. *Bull. Soc. Géol. Fr.*, 181, 363–376.
- Ge, S. (1998). Estimation of groundwater velocity in localized fracture zones from well temperature profiles. *J. Volcanol. Geotherm. Res.*, 84, 93–101.
- Guilbot, A. (1975). *Modélisation des écoulements d'un aquifère karstique (liaison pluie débit). Application aux bassins de Saugras et du Lez*. Thèse, Université Montpellier 2.
- Hassanizadeh, M. and Gray, W. G. (1980). General conservation equations for multi-phase systems. 3 Constitutive theory for porous media flow. *Adv. Water Resour.*, 3, 25–40.
- Husson, E. (2013). *Interaction géodynamique/karstification et modélisation géologique 3D des massifs carbonatés. Implication sur la distribution prévisionnelle de la karstification*. PhD

- thesis, Université Montpellier 2.
- Ingebritsen, S. E., Sherrod, D. R., and Mariner, R. H. (1989). Heat flow and hydrothermal circulation in Cascade range, north-central Oregon. *Science*, 243, 1458–1462.
- Karam, Y. (1989). *Essais de modélisation des écoulements dans un aquifère karstique. Exemple de la source du Lez*. PhD thesis, Université Montpellier 2.
- Lacas, J. L. (1976). *Introduction à la méthodologie d'étude et d'utilisation des champs hydrothermiques des aquifères karstiques: d'après l'exemple du site de l'exurgence de la source du Lez (Hérault, France)*. PhD thesis, Université Montpellier 2.
- Lachmar, T. E., Freeman, T. G., Sant, C. J., Walker, J. R., Joseph, F., Batir, J. F., Shervais, J. W., Evans, J. P., Nielson, D. L., and Blackwell, D. D. (2017). Effect of an 860-m thick, cold, freshwater aquifer on geothermal potential along the axis of the eastern Snake River Plain, Idaho. *Geotherm. Energy*, 5, article no. 28.
- Ladouche, B., Marechal, J. C., and Dorfliger, N. (2014). Semi-distributed lumped model of a karst system under active management. *J. Hydrol.*, 509, 215–230.
- Liedl, R. and Sauter, M. (1998). Modelling of aquifer genesis and heat transport in karst systems. *Bull. Hydrogéol.*, 16, 185–200.
- Liedl, R., Sauter, M., Huckinghaus, D., Clemens, T., and Teutsch, G. (2003). Simulation of the development of karst aquifer using a coupled continuum pipe flow model. *Water Resour. Res.*, 39, article no. 1057.
- Lismonde, B. (2010). Puissance géothermique advectée par l'eau d'une source et surface de son bassin versant. Étude Théorique. *Karstologia*, 55, 1–10.
- Lowell, R. P. (1975). Circulations in fractures, hot springs and convective transport on mid-ocean ridges crests. *Geophys. R. Astron. Soc.*, 40, 351–365.
- Lucazeau, F. and Vasseur, G. (1982). Heat flow density data from France and surrounding margins. *Tectonophysics*, 164, 251–258.
- Luetscher, M. and Jeannin, P. Y. (2004). Temperature distribution in karst systems: the role of air and water fluxes. *Terra Nova.*, 16, 344–350.
- Malard, F. and Chapuis, B. (1995). Temperature logging to describe the movement of sewage-polluted surface water infiltrating into a fractured rock aquifer. *J. Hydrol.*, 173, 191–217.
- Manga, M. and Kirchner, J. W. (2004). Interpreting the temperature of water at cold springs and the importance of gravitational potential energy. *Water Resour. Res.*, 40, article no. W05110.
- Marechal, J. C., Ladouche, B., Batiot-Guilhe, C., Borrell-Estupina, V., Caballero, Y., Cernesson, F., Dörfliker, N., Fleury, P., Jay-Allemand, M., Jourde, H., Leonardi, V., Malaterre, P. O., Seidel, J. L., and Vion, P. Y. (2014). Projet gestion multi-usages de l'hydrosystème karstique du Lez – synthèse des résultats et recommandations. Rapport BRGM/RP-61051-FR, 126 pp.
- Mathey, M. (1974). Gradient géothermique et hydraulique souterraine dans un aquifère karstique (bassin de la source de l'Areuse/NE). *Bull. Soc. Neuchatel. Sci. Nat.*, 97, 301–314.
- McLing, T. L., Smith, R. P., Smith, R. W., and Blackwell, D. D. (2016). Wellbore and groundwater temperature distribution eastern Snake River Plain, Idaho: Implications for groundwater flow and geothermal potential. *J. Volcanol. Geotherm. Res.*, 320, 144–155.
- Paloc, H. (1979). Alimentation en eau de la ville de Montpellier. Captage de la Source du Lez commune de St Clement (Hérault). Etude documentaire préalable a l'établissement des périmètres de protection. Note de Synthèse. Rapport BRGM/79SGN319LRO.
- Pimentel, E. T. and Hamza, V. M. (2012). Indications of regional scale groundwater flows in the Amazon Basins: Inferences from results of geothermal studies. *J. South Am. Earth Sci.*, 37, 214–227.
- Reiter, M. (2001). Using precision temperature logs to estimate horizontal and vertical groundwater flow components. *Water Resour. Res.*, 37, 663–674.
- Saar, M. G. (2011). Review: geothermal heat as a tracer of large scale groundwater flow and a means to determine permeability fields. *Hydrogeol. J.*, 19, 31–52.
- Smith, L. and Chapman, D. (1983). On the thermal effects of groundwater flow, I regional systems. *J. Geophys. Res.*, 88, 593–608.
- Sposito, G. and Chu, S. Y. (1981). The statistical mechanical theory of groundwater flow. *Water Resour. Res.*, 17, 885–892.
- Stauffer, P. H., Lewis, K. C., Stein, J. S., Travis, B. J., Lichner, P., and Zyvoloski, G. (2014). Joule-Thomson effects on the flow of liquid water. *Transp. Porous Media*, 105, 471–485.
- Tavares, E., Pimentel, L. C. G., and Hamza, V. M.



- (2014). Use of geothermal methods in outlining deep groundwater flow systems in Paleozoic interior basins of Brazil. *Hydrogeol. J.*, 22, 107–128.
- Temme, N. M. (2010). Parabolic cylinder function. In Olver, F. W., Lozier, D. W., Boisvert, R. F., and Clark, C. W., editors, *NIST Handbook of Mathematical Functions*. Cambridge University Press, Cambridge.
- Thiery, D., Berard, P., and Camus, A. (1983). Captage de la source du Lez: étude des relations entre la source et son réservoir aquifère. Ed. BRGM, 83 SGN 167 LRO, 95 p.
- Toth, J. (1963). A theoretical analysis of groundwater in small drainage basins. *J. Geophys. Res.*, 68, 4795–4812.
- Touet, F. (1985). *Détermination de l'origine des ressources en eau captées en bordure sud du pli de Montpellier (Hérault)*. PhD thesis, Université Paris XI Orsay.
- Uil, H. (1978). *Application du carottage thermique et de l'hydrochimie à l'étude des circulations d'eau souterraine en milieu karstique (région karstique Nord-Montpellieraine)*. PhD thesis, Université Montpellier 2.
- Vasseur, G. and Demongodin, L. (1995). Convective and conductive heat transfer in sedimentary basins. *Basin Res.*, 7, 67–79.
- Vasseur, G., Demongodin, L., and Bonneville, A. (2002). Thermal modelling of fluid flow effects in thin-dipping aquifers. *Geophys. J. Int.*, 112, 276–289.
- Vasseur, G., Lucazeau, F., and Bayer, R. (1985). The problem of heat flow density determination from inaccurate data. *Tectonophysics*, 121, 25–34.
- White, W. B. (2002). Karst hydrology: recent developments and open questions. *Eng. Geol.*, 65, 85–105.
- White, W. B. (2006). Fifty years of karst hydrology and hydrogeology. In Harmon, R. H. S. and Wicks, C. M., editors, *Perspective in Karst Geomorphology Hydrology and Geochemistry*, Special paper 404. Geological Society of America, Boulder, CO.
- Ziagos, J. P. and Blackwell, D. D. (1986). A model for the transient temperature effects of horizontal fluid flow in geothermal systems. *J. Volcanol. Geotherm. Res.*, 27, 371–397.



# **Geostatistics**





Research article

## Geo-hydrological Data & Models

# Random fields and up scaling, towards a more predictive probabilistic quantitative hydrogeology

Benoît Noetinger <sup>a</sup>

<sup>a</sup> IFP Energies Nouvelles, France

E-mail: [benoit.noetinger@ifpen.fr](mailto:benoit.noetinger@ifpen.fr)

**Abstract.** Random fields are becoming a mature tool sharing applications in many area of physics, mechanics and geosciences. In the latter, it is commonly used under the name of geostatistics. Continuous enrichment of geological/geostatistical models leads to manipulating hydrogeological models characterized by many parameters or hyperparameters corresponding to statistical aggregates that may be poorly estimated due to the scarcity of field data. Those parameters are generally support-scale-dependent and uncertain, so some inverse problem and uncertainty analysis must be carried out in practical applications that involve generally some forward calculation for example a fluid flow simulation if one is interested in transfers in the subsurface. Up scaling techniques are still required to find and to restrict in a controlled manner the more relevant parameters, allowing to lower the dimension of the parameter space. In the stochastic case, the interaction between the conductivity spatial distribution and the flow pattern can lead to non trivial behaviours that will be discussed. Fractured media will not be considered. That note does not present original results, but a selection of some potentially fruitful research avenues suggested by previous works.

**Keywords.** Applied geosciences, Porous media, Disorder, Upscaling, Geostatistics, Quenched disorder.

*Manuscript received 11 April 2022, revised 4 November 2022, accepted 22 November 2022.*

## 1. Introduction

The use of numerical models for studying subsurface flow has become common practice in hydrology and petroleum engineering over the last 50 years [Renard and De Marsily, 1997, De Marsily et al., 2005]. However, one of the major questions that still poses a problem is that the data needed to build an accurate model are incomplete in essence. Although contemporary computers are growing ever more powerful and capable of describing the relevant flows with increasing details, the required input data growth imposes a major increase in computing facilities if one wants to explore correctly the uncertain input parameter space in order to make the best decisions [Gorell et al., 2001]. In many applied geosciences issues, the basic concern is to recover, to store or to fol-

low some fluid, waste or even thermal energy in subsurface formations at good economical conditions, while minimizing overall environmental risks. Most of the associated transport processes occur in the porosity of natural rocks, perhaps fractured, so studying flow in porous media at several scales is a major issue. In order to proceed, a concise macroscopic description of flow in porous media encompassing several scales without knowing the exact structure of the subsurface at lower resolution scales must be obtained.

Basically one wants to be able to estimate with some accuracy the average value, as well as the associated variance of any simulated value of interest. Knowledge of the full probability distribution is in general not required, although knowing if the re-

sulting distribution is approximately Gaussian or not may be important. In the general case, when the driving equations relating input data to output observables are strongly non-linear, extensive Monte-Carlo sampling of the input parameter space is required to get estimations of these output statistical quantities. A typical subsurface model is generally described by some random fields of say porosity, permeability maps conditioned to several field observations, sedimentary layer shapes, aquifer positions, etc... These random maps depend in turn on some hyperparameters of statistical nature such as variance, drift, correlation scale: that is the aim of geostatistics that will be presented with more details in Section 2.1. In addition, some man-controlled parameters such as the position of wells and the associated flow rates enter as input data in the model. For applications, that influence of that set of parameters may be sampled by simulation to select the best position and management of the wells.

Using such models, we are led to consider high dimensional parameter spaces, both for the random set of natural data, and for the man-controlled parameters. Consider a parameter space of  $N = 100$  dimensions, and by hypothesis, let us assume that each parameter can take only 2 values. An exhaustive exploration of that parameter space will imply being able to perform  $2^{100}$  independent runs, that is clearly impossible even with hypothetical quantum computers. In addition, in current models, most parameters belong to some continuous interval, increasing the effective dimensions of the parameter space. So following any exhaustive approach is a dead end. Many alternative sampling techniques were developed [Zhang, 2001]. Experimental design techniques coupled with machine learning methods are being developed [Scheidt et al., 2007, Santos Oliveira et al., 2021, and references therein].

That implies that some systematic workflow is to be built in order to reduce the number of dimensions of the effective working space under consideration. As that situation presents some similarities with statistical physics, some bridges between both disciplines can be expected: a description of these connections may be found in Noetinger [2020]. To sum-up, similar issues are at the heart of one the most active area of complex systems understanding, the so-called spin glasses for which G Parisi was awarded the 2021 Nobel prize in physics [Mézard et al., 1987, Zde-

borová and Krzakała, 2007, Zdeborová and Krzakala, 2016]. In the geoscience side, some reviews illustrate similar issues [Gorell et al., 2001, Floris et al., 2001, De Marsily et al., 2005].

The goal of that short paper is to suggest some avenues of research in order to provide answers to the following questions:

- (i) What is the amount of information that is required to build a useful numerical model with a degree of sophistication consistent with the input knowledge.
- (ii) is it possible to build a model that may be continuously improved as our knowledge of the particular aquifer increases along with the volume of data acquisition.

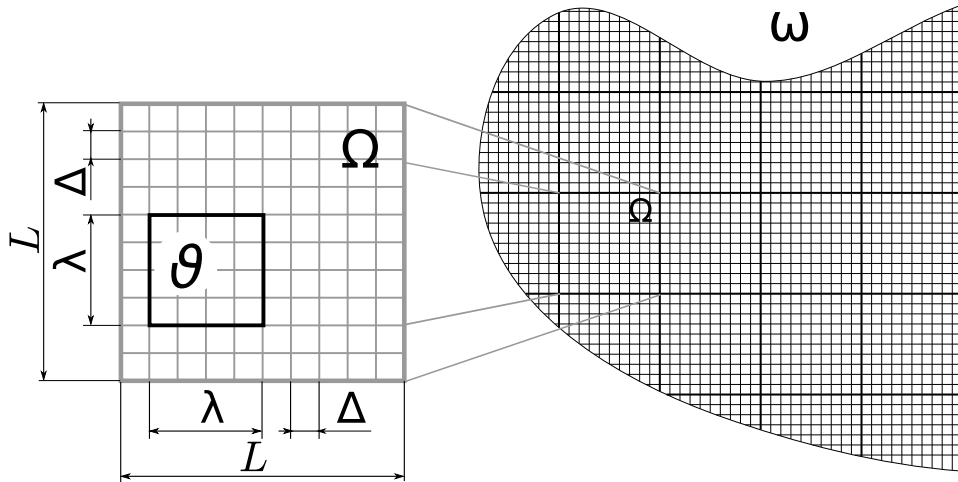
The paper is organised as follows: in Section 2, we provide some general elements. Then a focus is given in Section 2.1 about geostatistics, before giving some elements about information quantification in Section 2.2. Single phase flow problem serving as a basis for most of our discussion is introduced in Section 2.3. Then we discuss up-scaling applications in Section 3.1 by discussing the scale and support effects. Non linear issues implying a strong feed-back between the medium structure and the flow are discussed in the case of two-phase flows in Section 3.2. In Section 4, we present some ideas allowing to merge the preceding considerations between geostatistics, up-scaling and information theory for practical applications. Finally some fruitful research avenues are suggested for future investigations.

## 2. Medium description, up-scaling issues

We consider a subsurface formation of typical kilometeric scales. In Figure 1, the overall scales of the considered aquifer are represented, as well as the corresponding discretization mesh of interest, chosen cartesian for sake of simplicity. Each grid block may be populated by parameters such as porosity  $\Phi(\mathbf{r})$ , permeability  $k(\mathbf{r})$  and any other relevant parameter. That can be drawn using some deterministic or geostatistical software.

### 2.1. About geostatistics

Modelling subsurface formation using numerical models implies to account for the fact that any local



**Figure 1.** Geometry of the problem on a simplified 2D section of a subsurface model domain  $\omega$ . A coarse grid made of blocks  $\Omega$  of typical size  $L$  is super-imposed to a geological fine grid of typical size  $\Delta$  in order to solve the conservation equations of interest. The intermediate scale  $\lambda$  serves in posterior treatments for checking the overall consistency of the model. Reprinted (adapted) from Colecchio et al. [2020] Copyright 2020 with permission of Elsevier.

quantity of interest may depend on position vector  $\mathbf{r}$ . In most cases, for example the quenched<sup>1</sup> positive permeability  $k(\mathbf{r})$ , is represented as being a random function of position. It is characterized by some

mean-value and fluctuations that are measured using field data (an issue in itself!). That is the aim of the so-called geostatistical approach, originally founded by Matheron [1963], Krige [1976], De Marsily et al. [2005], Chiles and Delfiner [2009] and developed by several others, that form the conceptual basis of several popular commercial softwares. The basic idea is to interpolate known measurements on a given support on several locations by inferring some underlying statistical structure. Notice that the primary random mathematical objects are functions or fields that are characterized by an infinite number of degrees of freedom. So the most natural embedding formalism that encompasses it is the so called field-theory employed in both statistical mechanics or particle physics [Noetinger, 2020, Hristopoulos, 2020]. Another related issue that was pointed-out by Tarantola [2005] is that working with continuous variables implies that there is not a reference entropy measure allowing to quantify easily the information of a geostatistical distribution: the overall results are sensitive to the coordinate system used by the modeller. At this stage, the associated reference “non-informative” reference distribution [Tarantola, 2005, Abellan and Noetinger, 2010] is not determined. That explains some deep theoretical and practical difficulties for generalizing the approach to curved

<sup>1</sup>The term quenched is employed in statistical physics of disordered systems, it corresponds to systems having a fixed but unknown disorder without thermal fluctuations that reorganise it, a glass or an amorphous solid are typical quenched material.

spaces or to non-uniform grid, Mourlanette et al. [2020]: such a situation may occur if tectonics imply strong deformation of a sedimentary basin. In most practical cases, some implicit cartesian assumption is made. Probably it could be justified using the sedimentary nature of most aquifers. In the similar statistical physics case, it may be shown that the continuous Boltzmann–Gibbs distribution that is the basis of most applications may be derived by a suitable limit of the underlying quantum mechanics processes that poses paradoxically less difficulty owing to their discrete nature (Setting-up a natural probability measure to a 6 face dice rather than to a continuous process). In geoscience, approximate Log-normal distributions (the logarithm of the conductivity is a Gaussian distributed variable) were observed at the core scale within several geological environments [Gelhar, 1993]. Note that in most cases, the amount of data is not sufficient to provide the form of the probability distribution function (pdf) of  $k$ , and higher order correlation functions are almost impossible to be determined from field data. In the current modelling practice, they are implicitly fixed by choosing some additional assumptions based on geological expertise on similar environments. As a consequence, even the input stochastic model is questionable.

Another class of models that intend to add information on high order correlation functions are multipoint or so-called process-based models, intending to reconstruct the apparent randomness of the subsurface by simulating the long term history of the formation using some dynamical physically based model: Diggle et al. [1998], Granjeon and Joseph [1999], Hu and Chugunova [2008], Michael et al. [2010]. These models introduce more physically based parameters and are more realistic. However, they may depend on assumptions regarding the time-dependant external forcing at their boundaries, introducing stochasticity in a somewhat another manner. An essential specificity of geosciences is the so-called measurement support effect: any data is obtained using some instrumentation that averages the details of the process at hand on a characteristic length that depends on the physical process. That scale may vary from some centimeters for laboratory/logging tools, to hectometers in geophysics corresponding to the typical wavelength of geophysical waves, or to the support of transient well tests in-

terpretation. Any measurement is characterized by a support scale characterizing the measurement itself, and by another spatial scale that is the typical resolution.

Finally, due to the data scarcity, it is necessary to up-date any prior model of the subsurface by accounting for the additional data that are continuously enriching the available information. That new data can be provided by pressure and rates measurements at some producing wells, time-dependant tracer concentration showing connections between several regions of the subsurface etc... This is the so called inverse problem issue that was addressed by many authors [Jacquard, 1965, LaVenue et al., 1995]. It is well-known that such problems are ill conditioned and that some regularization is required [Tarantola, 2005, Oliver and Chen, 2011, and references therein]. That can be achieved accounting for a prior geostatistical structure [LaVenue et al., 1995, Le Ravalec et al., 2000, Hu and Chugunova, 2008] that is easier to carry-out considering gaussian priors.

## 2.2. Quantification of information

As data may be scarce (but even redundant), it is essential to be able to quantify the amount of information provided by a measurement, as well as the quantity of information that is required to make faithful predictions. That idea can be formalized using information theory, as attempted in Abellan and Noetinger [2010, and references therein]. A quantitative estimate can be defined introducing the relative entropy (Kullback–Leibler divergence) between posterior and prior distribution that serves as a non-informative reference distribution. That quantity is defined by:

$$I(p) = I(p|\mu) = \int p(\mathbf{m}) \log \frac{p(\mathbf{m})}{\mu(\mathbf{m})} d\mathbf{m}. \quad (1)$$

Here,  $p(\mathbf{m})$  is the posterior distribution of  $\mathbf{m}$  given that some measurement was carried out on the system.  $\mu(\mathbf{m})$  denotes the prior distribution  $\mathbf{m}$ . For clarity, it is easier to introduce some discretization of  $\mathbf{m}$  on a regular grid. So  $\mathbf{m}$  is represented as a realization of the  $M$  dimensional vector, so  $\mathbf{m} = (m(\mathbf{x}_1), \dots, m(\mathbf{x}_M))^t$ . The corresponding  $\mu(\mathbf{m})$  is assumed to be centered and multigaussian so the corresponding probability measure is given by  $\mu(\mathbf{m}) \sim$



$\exp -1/2(\mathbf{m} \cdot \mathbf{C}_p \cdot \mathbf{m})$  with a gaussian covariance matrix with elements that are given by:

$$\mathbf{C}_p(\mathbf{x}_i, \mathbf{x}_j) = \sigma_0^2 \exp\left(\frac{-|\mathbf{x}_i - \mathbf{x}_j|^2}{\ell_c^2}\right) \quad (2)$$

with  $\ell_c^2$  is the correlation length, and  $\sigma_0^2$  the associated variance. Then the relative entropy  $I_{MN}$  of a posterior distribution corresponding to  $N$  linear measurements with independant errors  $\sigma_d^2$  gathered as a  $N$  dimensional vector  $\mathbf{F} \cdot x$  is given by  $I_{MN}$ :

$$I_{MN} = \frac{1}{2} \log |\mathbb{I}_N + \mathbf{C}_v^{-1} \cdot \mathbf{F}^t \mathbf{C}_p \mathbf{F}| \quad (3)$$

where  $|\dots|$  is the determinant of the square matrix  $\dots$ . Considering the specific problem of local measurements of the random field  $\mathbf{m}$  at  $N$  different locations  $(\mathbf{x}_1, \dots, \mathbf{x}_N)$ ,  $N < M$  with measurement errors  $\sigma_d^2$ , one is led to evaluate:

$$I_{MN}(\mathbf{x}_1, \dots, \mathbf{x}_N) = \frac{1}{2} \log \left| \left( \delta_{ij} + \frac{\sigma_0^2}{\sigma_d^2} \exp \frac{-|\mathbf{x}_i - \mathbf{x}_j|^2}{\ell_c^2} \right)_{1 \leq i, j \leq N} \right|, \quad (4)$$

when the measurement locations are far apart (typically all distances greater than  $\ell_c$ ), one gets the estimation:

$$I_{MN}(x_1, \dots, x_N) = \frac{N}{2} \log(1 + \mu^2) \quad \text{with } \mu^2 = \frac{\sigma_0^2}{\sigma_d^2}. \quad (5)$$

Figure 2 illustrates present results that were also generalised to pumping test analysis and non-linear flow driven issues in Abellan and Noetinger [2010].

### 2.3. Fluid-flow modelling

In most cases, modelling fluid flow in a porous formation [De Marsily, 1986] leads to solve a diffusion-like equation that reads:

$$\phi c_t \frac{\partial p(\mathbf{r}, t)}{\partial t} = \nabla \cdot (k(\mathbf{r}) \nabla p(\mathbf{r}, t)) + f(\mathbf{r}). \quad (6)$$

Here, the parameters  $p(\mathbf{r}, t)$ ,  $c_t$  and  $f(\mathbf{r})$  are respectively the time-dependent potential of interest, the total compressibility, and a source term. Dirichlet or Neumann Boundary conditions are known at the boundary of the domain  $\omega$ . Once that potential is known, associated transport problems can be solved, including multiphase flow issues that are not in the scope of present paper.

The basic issue is to be able to solve such equations, accounting for both the the effect of heterogeneities, and to be able to quantify the related uncertainties. In that context, analytical methods were

developed by hydrogeologists [Gelhar, 1993, Dagan, 1989, Indelman and Abramovich, 1994, Abramovich and Indelman, 1995, Renard and De Marsily, 1997], mathematicians [Jikov et al., 2012, Armstrong et al., 2019] and physicists, among other [King, 1989, Noetinger, 1994, Hristopulos and Christakos, 1999, Noetinger, 2000, Attinger, 2003, Eberhard et al., 2004, Teodorovich, 2002, Stepanyants and Teodorovich, 2003].

On the numerical side, a possible strategy is to transfer the small scale spatial fluctuations to a large scale support that encompasses the low spatial frequency components of the fields of interest. The calculations are carried out in practice using some numerical model in which the Laplace equation is solved using a grid of resolution  $L$  generally considerably much coarser than the input fine geological grid  $\Delta$  (notations in Figure 1), because the available computing power leads also to continuous improvement of local geological 3D representations. This implies obtaining a coarse grained Laplace equation with a renormalized conductivity map accounting as best as possible to the local subgrid variations. In the case of multimodal distributions, in which connectivity aspects are dominant, specific methods can be developed within the framework of percolation theory, such as real-space renormalization techniques [Berkowitz and Balberg, 1993, Hunt et al., 2014, Hristopulos, 2020].

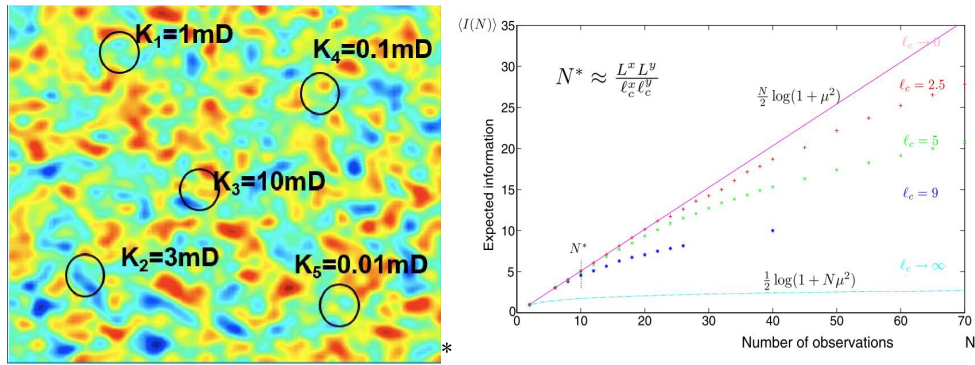
In the stochastic context, a strategy is to average the solution of Equation (6) to get the average head (or pressure)  $\langle p(\mathbf{r}, t) \rangle$  in which the ensemble-average  $\langle \dots \rangle$  is to be taken on the disorder of the conductivity field  $k(\mathbf{r})$  corresponding to a Monte-Carlo average over the disorder of the medium. Then, one can look for the equation that may drive that average head<sup>2</sup> [Noetinger and Gautier, 1998].

## 3. About flow up-scaling

### 3.1. Single phase steady state up-scaling

It can be shown that under quite general hypothesis (statistical stationarity and convergence conditions)

<sup>2</sup>At first sight, averaging the driving equation looks to be simpler, and would evidently provide the arithmetic average for the conductivity, a clearly wrong result.



**Figure 2.** Left: A typical realization of a random map 2D section of a subsurface model on a domain  $\Omega$ , representing  $\text{Log}(k)$ . The superimposed circles correspond to  $N = 5$  physical measurements of permeability on a support scale with the corresponding radius, with several realistic values. The average distance between the circles corresponds to the coverage of the area. In a perfect situation, all the area should be covered, that is seldom the case. Right: the amount of information provided by  $N$  local measurements defined by the relative entropy of posterior distribution versus prior. Note the saturation of the information once the medium is well covered by the measurements,  $\mu$  is defined in the main text.  $L^x$  and  $L^y$  are the overall size of the domain in the  $x$  and  $y$  directions,  $\ell_0^x$  and  $\ell_0^y$  denote the correlation length in corresponding directions.  $N^*$  is thus the number of correlated volumes inside the overall domain. It corresponds to the crossover the transition of information content between independent measurements and redundant ones. Copyright 2020 with permission of Springer-Nature.

that the low frequency components of the average potential  $\langle p(\mathbf{r}, t) \rangle$  is driven by an effective equation that reads:

$$\langle \phi \rangle_{c_t} \frac{\partial \langle p(\mathbf{r}, t) \rangle}{\partial t} = \nabla \cdot (K_{\text{equ}} \nabla \langle p(\mathbf{r}, t) \rangle) + f(\mathbf{r}). \quad (7)$$

Here, the average  $\langle \dots \rangle$  corresponds to an averaging over the randomness of the medium. The parameter  $K_{\text{equ}}$  is called the equivalent conductivity. It corresponds to the “natural” large scale relation between the mean flux and the large scale pressure gradient that can be provided by homogenization theories solving the so-called “auxiliary problem” to be solved numerically in  $x$ ,  $y$  and  $z$  directions [Armstrong et al., 2019, and references therein], as it is illustrated Figure 3.<sup>3</sup> Many expressions were proposed to relate  $K_{\text{eff}}$  to the underlying disorder [Renard and

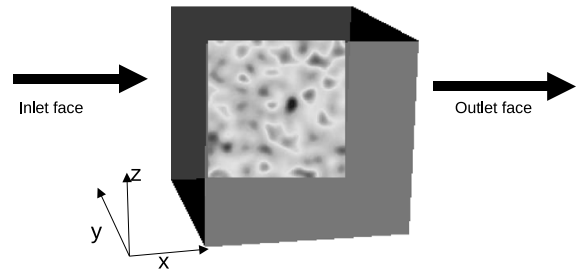
De Marsily, 1997]. For one dimension, an elementary analytical calculation provides the harmonic average  $K_{\text{eff}} = \langle k^{-1} \rangle^{-1}$ . Generalizations of such a simple formula were proposed by many authors [Landau and Lifshitz, 1960, Matheron, 1967, King, 1987, 1989, Dagan, 1993, Neuman and Orr, 1993, Noetinger, 1994, Indelman and Abramovich, 1994, De Wit, 1995, Teodorovich, 2002, Stepanyants and Teodorovich, 2003, Colechio et al., 2020, Nøttinger, 2000]. Looking at the evolution of the permeability distribution, in 2 dimensions, numerically the log normal distribution appears to be stable under the up-scaling transformation, Figure 4 from Colechio et al. [2020], analogous to the central limit theorem. That may be related to the duality argument of Matheron [1967] that justifies the geometric mean in 2D. In the 3D case, studying the emergence of a “stable” conductivity distribution invariant on the up scaling transformation would also be useful for studying strongly correlated systems having conductivity correlations decaying as a power law with the lag distance. This may be illustrated on Figure 4 below: in the practical side, in most cases, at present times,

<sup>3</sup>Notice that in the current practice, the result of the evaluation of these formula, or of the numerical solution of the auxiliary problem is used to populate the coarse grid of  $\Omega$  blocks, using the underlying local fine grid data. So the equivalent  $K_{\text{equ}}$  may still depend on position rather than the so-called  $K_{\text{eff}}$  that is intrinsic, it would correspond to working with an infinite volume  $\Omega$ .

using a numerical technique is sufficient. In case of extremely heterogeneous media (that can correspond to bimodal media) at percolation threshold, a percolation transition may occur [Charlaix et al., 1987, Berkowitz and Balberg, 1993, Hunt et al., 2014, Stauffer and Aharony, 2014, Colecchio et al., 2020]. Its main consequence, a stabilization of the effective conductivity over a very large scale can be observed on Figure 5. In both situations, it appears that at some scale (becoming infinite just at the percolation threshold for an infinite small-scale conductivity contrast), the formation can be treated as an almost continuous medium: that is the so called self-averaging property that manifests itself once the characteristic size of the flow is larger than the typical correlation scale. Long ranged media, as well as media close to a percolation threshold remain a notable exception. The results presented in Figures 4 and 5 were obtained using a suitable post treatment of a single phase flow simulation on a large domain, without adding any external expansive Monte Carlo loop: that is sufficient to capture the dominant characteristic scales of the flow. In addition, in Figure 5, middle, it is illustrated the emergence of the so-called localization phenomena that is a consequence of high permeability contrasts [Arnold et al., 2019] that localizes high velocity regions in tiny regions, even in non-fractured media. Such flow-based maps are well suited for post-treatment: helping to build a coarse mesh, estimating critical paths that may control the overall uncertainties.

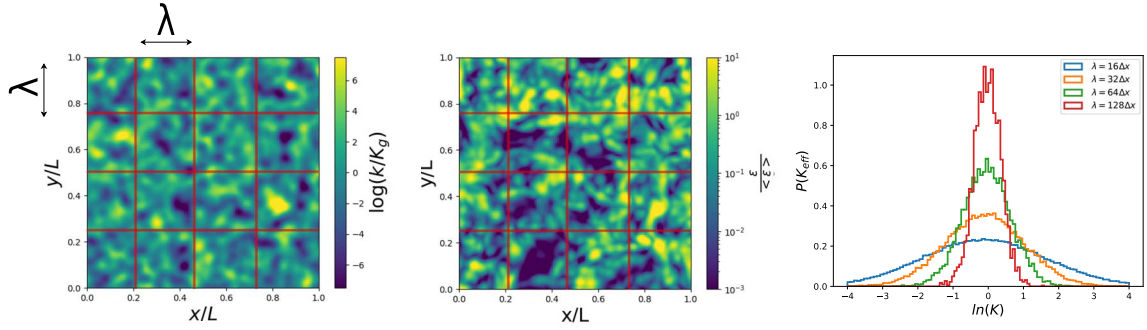
### 3.2. *About the interaction between the conductivity spatial distribution and the flow pattern*

In several situations, such as when considering multiphase flow flows in heterogeneous media, some strong coupling between the flow and the underlying medium may occur. This can be the case when considering water displacing oil, CO<sub>2</sub> injection etc... This may give rise to an instability driven by a mechanism is similar to the well-known Saffman–Taylor instability that leads to viscous fingering if the displacing fluid is more mobile than the fluid initially in place. This emergence problem was studied by many authors such as Saffman and Taylor [1958], Homay [1987], Tabeling et al. [1988], Tang [1985], Shraiman [1986]. A generalization to immiscible two phase flows

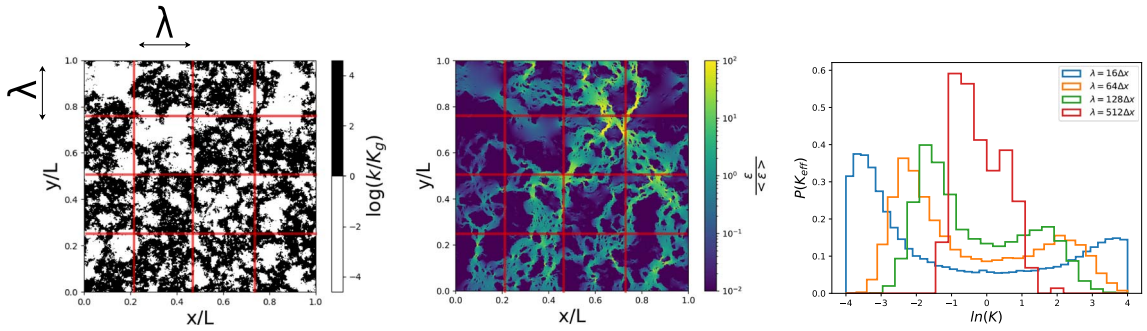


**Figure 3.** Up-scaling geometry. The coarse block of size  $L$  have a detailed conductivity map given by geology. It is up-scaled by solving a steady-state quasi Laplace equation to determine an effective conductivity in the mean flow direction. Boundary conditions are usually no-flux parallel to the imposed mean flow, or periodic.

was carried-out in the paper of King and Dunayevsky [1989, and references therein]. In that case, it appears a well defined sharp front separating both phases moves at the local velocity of the fluid, as shown in Figure 6, top. That front may be in unstable, by the same mechanism than the Saffman–Taylor instability, as shown in Figure 6, bottom. In the unstable case, an amplification of the heterogeneity effect is expected, while some smoothing arises in the stable case. The intuition suggests that a highly mobile



**Figure 4.** Monte Carlo study of the dependence of the effective conductivity pdf with coarsening scale  $\lambda$ . The overall flow is solved using several realizations of the input log conductivity map (left). The associated local dissipation map (center) allows to evaluate a distribution of coarsened effective conductivities averaged at intermediate scale  $\lambda$  [Colecchio et al., 2020]. The associated pdf's are plotted (right). The self averaging feature is highlighted by the sharply peaked distribution around the geometric average for  $\lambda = 128$  units. Reprinted (adapted) from Colecchio et al. [2020] Copyright 2020 with permission of Elsevier.



**Figure 5.** Monte Carlo study of the effective conductivity pdf with coarsening scale  $\lambda$ . The overall flow is solved using several realizations of the input bimodal map with conductivity contrast of  $10^4$ . The associated local dissipation map (center) highlights the localization phenomenon and allows to evaluate pdf's of coarsened effective conductivities averaged at scale  $\lambda$  [Colecchio et al., 2020]. The resulting scale-dependant pdf's are plotted (right). As the scale is increasing, the two peaks merge, the bimodal distribution disappears and becomes close to a log-normal distribution. The convergence to this asymptotic distribution shows critical slowing-down when the facies proportions are close to percolation threshold. In the infinite contrast case, scaling-laws are recovered [Stauffer and Aharony, 2014]. Reprinted (adapted) from Colecchio et al. [2020] Copyright 2020 with permission of Elsevier.

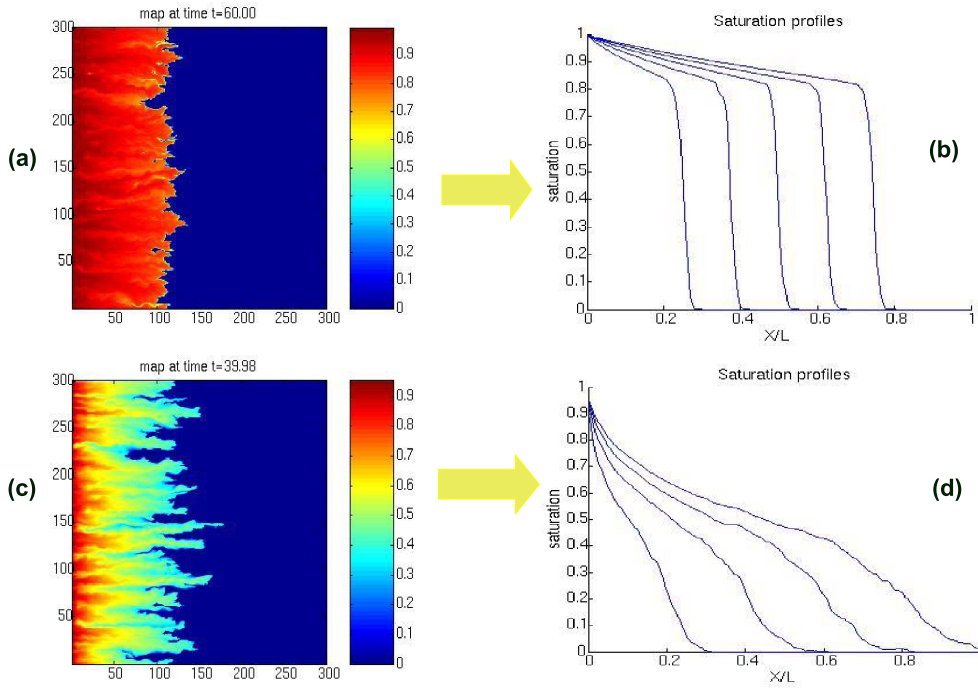
fluid will follow high velocity paths, its presence amplifying thus that advantage by a positive feedback loop. This is the so called channeling issue. This problem was addressed by De Wit and Homsy [1997a,b], and revisited in the stochastic context by Artus et al. [2004], Noetinger et al. [2004], Artus and Noetinger [2004], merging perturbation theory within the framework developed by King and

Dunayevsky [1989]. In order to be more specific, the underlying equations reads:

$$\nabla \cdot [\lambda(S(\mathbf{r}, t))k(\mathbf{r})\nabla p(\mathbf{r}, t)] = 0 \quad (8)$$

$$\phi \frac{\partial S(\mathbf{r}, t)}{\partial t} + \nabla \cdot (f(S(\mathbf{r}, t))\mathbf{U}(\mathbf{r}, t)) = 0 \quad (9)$$

$$\mathbf{U}(\mathbf{r}, t) = -\lambda(S(\mathbf{r}, t))k(\mathbf{r})\nabla p(\mathbf{r}, t). \quad (10)$$



**Figure 6.** Dynamics of a two phase flow front moving in a heterogeneous rock, imposed mean flow from left to right, underlying random log normally distributed conductivity map top (a) stable case (b)  $Y$  averaged saturation at different times, bottom (c) fingering in the unstable case, (d) associated  $Y$ -averaged saturation at different times.

Here,  $S(\mathbf{r}, t)$ ,  $\lambda(S(\mathbf{r}, t))$  and  $f(S(\mathbf{r}, t))$ ,  $S$  denote respectively the water saturation (local % of water, total mobility and the so-called fractional flow of water). This set of coupled equations may be solved numerically (it is at the heart of any multiphase flow in porous media simulator, to which additional complexities such as phase transitions and boundary condition management must be added). The saturation equation (9) is hyperbolic, leading to the formation of a shock front, whose stability is controlled by the jump of the total mobility  $\lambda(S(\mathbf{r}, t))$  at the front.

The technical difficulty for setting-up a perturbation expansion comes from the presence of the front that implies a mobility jump that renders perturbation theory a bit tricky [King and Dunayevsky, 1989, Nøetinger et al., 2004]. The difficulty may be avoided by a suitable change of variable, using a working variable  $x(S, y, t)$  rather than  $S(x, y, t)$ . It is thus possible to introduce the function  $x(S_f, t)$ , in which  $S_f$  is the saturation of water just behind the front

$M_f$  is the corresponding total mobility jump  $M_f = \lambda(S_f)/\lambda(S = 0)$ . The randomness of the underlying conductivity field propagates to the randomness of  $x(S_f, t)$ , of average value  $\langle \mathbf{U} \rangle t$ . At long times, the associated two point correlation function can be shown to converge to a well defined function in the stable case, while in unstable case it diverges, a manifestation of the spreading of the front [Tallakstad et al., 2009, Toussaint et al., 2012]. A possible approach of practical interest close to the single phase flow approach would be to look at an effective equation driving the ensemble-averaged water saturation  $\langle S(\mathbf{r}, t) \rangle$ , or the  $Y$ -averaged saturation  $S(x, t)$ . A diffusive-like regime arises in the case  $M_f = 1$ , that corresponds to a marginal stability criterion. In the general case, several proposals were reported long ago for characterizing the emerging large scale transport equation [Koval et al., 1963, Todd et al., 1972, Yortsos, 1995, Blunt and Christie, 1993, Sorbie et al., 1995]. In the unstable case, one can consider that long fingers parallel to the imposed flow may be treated as a stratified medium.

This leads to modify the fractional flow function with an ad-hoc change [Koval et al., 1963]. In the stable case, it can be shown that the competition between the disorder that distorts the front and the viscous forces that tends to sharpen the front [Noetinger et al., 2004] must lead to another form of the effective fractional flow, including some macrodispersion representing the net effect of the averaged disorder. In the infinite contrast case (e.g. immiscible gas injection) Diffusion Limited Aggregation (DLA) models were proposed, leading to a very rich literature involving percolation invasion models, fractals [Witten Jr and Sander, 1981, Wilkinson and Willemsen, 1983, Paterson, 1984, Masek and Turcotte, 1993] with many contributions of SP.

#### 4. Geostatistics, up-scaling and inverse modelling

Our original issue was to know the relevance of a complex model in regard to the input knowledge. It appears that if one is interested by predicting the average and variance of the output of some transport process, a compromise must be found between the overall accuracy of the calculation involving a high resolution grid implying large computing costs, and the statistical accuracy that requires being able to perform a huge number of Monte-Carlo simulations<sup>4</sup> required to get representative statistical averages of the process under consideration. The overall simulation budget  $B(\text{Size}_{\text{model}}, N_{\text{MC}})$  in terms of hours of CPU depends on the model size  $\text{Size}_{\text{model}}$  and the number of Monte Carlo simulations  $N_{\text{MC}}$ . The latter depends on the number of relevant input parameters that control the overall behaviour of the system. It is possible to determine the couple  $(\text{Size}_{\text{model}}, N_{\text{MC}})$  such that the variance  $\sigma^2(\text{Size}_{\text{model}}, N_{\text{MC}})$  of the prediction errors of the simulation output is minimized under the constraint of fixed  $B(\text{Size}_{\text{model}}, N_{\text{MC}})$ . As presented in the preceding section, up-scaling techniques may provide some tools allowing to get a case-dependent prior estimate of that variance without running the whole set of simulations.

Practitioners want to get estimations of the average value and of the variance of the output of interest

(water supply, heat recovery etc...). So the approach developed by Abellan and Noetinger [2010] could be carried-out in the prediction space. The advantage of such an approach will be to be closer to operational quantities of practical interest, and also to be more independent on the dimension of the input parameter space. Such an avenue needs to be developed.

#### 5. Conclusions and some perspectives

In that short note, we discussed some ideas and previous findings illustrating mainly the interaction of the quenched disorder of the natural media present at all scales to the flow. It appears that hydrologists could be greatly helped by developing an approach combining pragmatic practice, advanced averaging techniques and optimization techniques merged within a framework of physically guided artificial intelligence techniques. An approach is to draw the “phase diagram” of the problem at hand, i.e. the set of dominant parameters controlling the overall behaviour of the system, and providing descriptions of the critical behaviour of the system between the different regions of the phase diagram. Information theory concepts can be helpful to quantify the disorder and the amount of information of the system, providing useful links to statistical physics concepts and methods.

Such an approach is evidently process-dependant and must be adapted for heat or tracer transport and mixing, as well as reactive flows, issues that were not discussed in that short paper.

A promising approach would be to map the linear systems to be solved that are essentially of Laplacian nature into random matrix theory close to random graphs theory [Mohar, 1997, Bordenave and Lelarge, 2010, Potters and Bouchaud, 2019]. That area of research is a fast growing field with applications for modelling networks such as the Internet, urban and road traffic, smart electrical networks, brain modelling, finance etc... Some applications may be found in the modelling of flow in fractured rocks [Hyman et al., 2017, 2019]. A possible mathematical approach is to study the spectrum of the large matrices obtained for solving the quasi Darcy flow equation that depends randomly on the geological input parameters. That may provide information about the effective number of relevant degrees of freedom that must be retained to describe the system [Biroli et al., 2007].

<sup>4</sup>We implicitly consider that fully analytically solvable problems are very scarce!

Depending on the degree of disorder, some eigenvalues of the associated Laplacian matrix can be expected to provide information while other eigenvalues may follow some universal distribution such as a Marchenko–Pastur distribution [Marchenko and Pastur, 1967]. Such topics are deeply connected to classification methods by neural nets [Louart et al., 2018, Dall’Amico et al., 2019].

Finally, as illustrated by De Marsily [2009], water supply issues will become more and more critical, implying great applied and academic research efforts. As it was explained in present paper, expecting tremendous computing power cannot be the single option. The community will have to rely on innovative approaches at the cutting edge of modern science, as our mentors did [Matheron and De Marsily, 1980, De Marsily et al., 2005].

### Conflicts of interest

The author has no conflict of interest to declare.

### Acknowledgements

The author is indebted to MM Alexandre Abellan, Alejandro Boschan, Ivan Colecchio Pauline Mourlanette and Alejandro Otero for their useful contributions.

### References

- Abellan, A. and Noetinger, B. (2010). Optimizing subsurface field data acquisition using information theory. *Math. Geosci.*, 42(6), 603–630.
- Abramovich, B. and Indelman, P. (1995). Effective permittivity of log-normal isotropic random media. *J. Phys. A: Math. Gen.*, 28(3), 693–700.
- Armstrong, S., Kuusi, T., and Mourrat, J.-C. (2019). *Quantitative Stochastic Homogenization and Large-scale Regularity*, volume 352. Springer, Berlin.
- Arnold, D. N., David, G., Filoche, M., Jerison, D., and Mayboroda, S. (2019). Localization of eigenfunctions via an effective potential. *Commun. Partial Differ. Equ.*, 44(11), 1186–1216.
- Artus, V. and Noetinger, B. (2004). Up-scaling two-phase flow in heterogeneous reservoirs: current trends. *Oil Gas Sci. Technol.*, 59(2), 185–195.
- Artus, V., Noetinger, B., and Ricard, L. (2004). Dynamics of the water–oil front for two-phase, immiscible flow in heterogeneous porous media. 1–stratified media. *Transp. Porous Media*, 56(3), 283–303.
- Attinger, S. (2003). Generalized coarse graining procedures for flow in porous media. *Comput. Geosci.*, 7(4), 253–273.
- Berkowitz, B. and Balberg, I. (1993). Percolation theory and its application to groundwater hydrology. *Water Resour. Res.*, 29(4), 775–794.
- Biroli, G., Bouchaud, J.-P., and Potters, M. (2007). Extreme value problems in random matrix theory and other disordered systems. *J. Stat. Mech.*, 2007(07), article no. P07019.
- Blunt, M. and Christie, M. (1993). How to predict viscous fingering in three component flow. *Transp. Porous Media*, 12(3), 207–236.
- Bordenave, C. and Lelarge, M. (2010). Resolvent of large random graphs. *Random Struct. Algorithms*, 37(3), 332–352.
- Charlaix, E., Guyon, E., and Roux, S. (1987). Permeability of a random array of fractures of widely varying apertures. *Transp. Porous Media*, 2(1), 31–43.
- Chiles, J.-P. and Delfiner, P. (2009). *Geostatistics: Modeling Spatial Uncertainty*, volume 497. John Wiley & Sons, New-York.
- Colecchio, I., Boschan, A., Otero, A. D., and Noetinger, B. (2020). On the multiscale characterization of effective hydraulic conductivity in random heterogeneous media: a historical survey and some new perspectives. *Adv. Water Resour.*, 140, article no. 103594.
- Dagan, G. (1989). *Flow and Transport in Porous Formations*. Springer-Verlag GmbH & Co. KG, Berlin.
- Dagan, G. (1993). Higher-order correction of effective permeability of heterogeneous isotropic formations of lognormal conductivity distribution. *Transp. Porous Media*, 12(3), 279–290.
- Dall’Amico, L., Couillet, R., and Tremblay, N. (2019). Classification spectrale par la laplacienne déformée dans des graphes réalistes. In *GRETSI 2019 - XXVIIème Colloque francophone de traitement du signal et des images, August 2019, Lille, France*. fffhal-02153901.
- De Marsily, G. (1986). Quantitative hydrogeology. Technical report, Paris School of Mines, Fontainebleau.
- De Marsily, G. (2009). *L’eau, un trésor en partage*. Dunod, Paris.

- De Marsily, G., Delay, F., Gonçalves, J., Renard, P., Teles, V., and Violette, S. (2005). Dealing with spatial heterogeneity. *Hydrogeol. J.*, 13(1), 161–183.
- De Wit, A. (1995). Correlation structure dependence of the effective permeability of heterogeneous porous media. *Phys. Fluids*, 7(11), 2553–2562.
- De Wit, A. and Homsy, G. (1997a). Viscous fingering in periodically heterogeneous porous media. I. Formulation and linear instability. *J. Chem. Phys.*, 107(22), 9609–9618.
- De Wit, A. and Homsy, G. (1997b). Viscous fingering in periodically heterogeneous porous media. II. Numerical simulations. *J. Chem. Phys.*, 107(22), 9619–9628.
- Diggle, P. J., Tawn, J. A., and Moyeed, R. A. (1998). Model-based geostatistics. *J. R. Stat. Soc.: Ser. C Appl. Stat.*, 47(3), 299–350.
- Eberhard, J., Attinger, S., and Wittum, G. (2004). Coarse graining for upscaling of flow in heterogeneous porous media. *Multiscale Model. Simul.*, 2(2), 269–301.
- Floris, F. J., Bush, M., Cuypers, M., Roggero, F., and Syversveen, A. R. (2001). Methods for quantifying the uncertainty of production forecasts: a comparative study. *Pet. Geosci.*, 7(S), S87–S96.
- Gelhar, L. W. (1993). *Stochastic Subsurface Hydrology*. Prentice-Hall, Upper Saddle River, NJ.
- Gorell, S., Bassett, R., et al. (2001). Trends in reservoir simulation: Big models, scalable models? Will you please make up your mind? In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, Houston.
- Granjeon, D. and Joseph, P. (1999). Concepts and applications of a 3-D multiple lithology, diffusive model in stratigraphic modeling. In *Numerical Experiments in Stratigraphy: Recent Advances in Stratigraphic and Sedimentologic Computer Simulations*. SEPM Society for Sedimentary Geology, Oklahoma City.
- Homsy, G. M. (1987). Viscous fingering in porous media. *Annu. Rev. Fluid Mech.*, 19(1), 271–311.
- Hristopulos, D. and Christakos, G. (1999). Renormalization group analysis of permeability upscaling. *Stoch. Environ. Res. Risk Assess.*, 13(1–2), 131–161.
- Hristopulos, D. T. (2020). *Random Fields for Spatial Data Modeling A Primer for Scientists and Engineers*. Springer, Berlin.
- Hu, L. and Chugunova, T. (2008). Multiple-point geostatistics for modeling subsurface heterogeneity: A comprehensive review. *Water Resour. Res.*, 44(11), article no. W11413.
- Hunt, A., Ewing, R., and Ghanbarian, B. (2014). *Percolation Theory For Flow in Porous Media*, volume 880. Springer, Berlin.
- Hyman, J. D., Dentz, M., Hagberg, A., and Kang, P. K. (2019). Emergence of stable laws for first passage times in three-dimensional random fracture networks. *Phys. Rev. Lett.*, 123(24), article no. 248501.
- Hyman, J. D., Hagberg, A., Srinivasan, G., Mohd-Yusof, J., and Viswanathan, H. (2017). Predictions of first passage times in sparse discrete fracture networks using graph-based reductions. *Phys. Rev. E*, 96(1), article no. 013304.
- Indelman, P. and Abramovich, B. (1994). A higher-order approximation to effective conductivity in media of anisotropic random structure. *Water Resour. Res.*, 30(6), 1857–1864.
- Jacquard, P. (1965). Permeability distribution from field pressure data. *Soc. Pet. Eng. J.*, 5(04), 281–294.
- Jikov, V. V., Kozlov, S. M., and Oleinik, O. A. (2012). *Homogenization of Differential Operators and Integral Functionals*. Springer Science & Business Media, Berlin.
- King, M. J. and Dunayevsky, V. A. (1989). Why waterflood works: a linearized stability analysis. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, Houston.
- King, P. (1987). The use of field theoretic methods for the study of flow in a heterogeneous porous medium. *J. Phys. A: Math. Gen.*, 20(12), 3935–3947.
- King, P. (1989). The use of renormalization for calculating effective permeability. *Transp. Porous Media*, 4(1), 37–58.
- Koval, E. et al. (1963). A method for predicting the performance of unstable miscible displacement in heterogeneous media. *Soc. Pet. Eng. J.*, 3(02), 145–154.
- Krige, D. (1976). A review of the development of geostatistics in South Africa. In *Advanced Geostatistics in the Mining Industry*, pages 279–293. Springer, Berlin.
- Landau, L. and Lifshitz, E. (1960). *Electrodynamics of Continuous Media*, volume 8. Mir, Moscow.
- LaVenue, A. M., RamaRao, B. S., De Marsily, G., and Marietta, M. G. (1995). Pilot point methodology for automated calibration of an ensemble of con-



- ditionally simulated transmissivity fields: 2. Application. *Water Resour. Res.*, 31(3), 495–516.
- Le Ravalec, M., Noetinger, B., and Hu, L. Y. (2000). The FFT moving average (FFT-MA) generator: An efficient numerical method for generating and conditioning Gaussian simulations. *Math. Geol.*, 32(6), 701–723.
- Louart, C., Liao, Z., Couillet, R., et al. (2018). A random matrix approach to neural networks. *Ann. Appl. Probab.*, 28(2), 1190–1248.
- Marchenko, V. A. and Pastur, L. A. (1967). Distribution of eigenvalues for some sets of random matrices. *Mat. Sb.*, 114(4), 507–536.
- Masek, J. G. and Turcotte, D. L. (1993). A diffusion-limited aggregation model for the evolution of drainage networks. *Earth Planet. Sci. Lett.*, 119(3), 379–386.
- Matheron, G. (1963). Principles of geostatistics. *Econ. Geol.*, 58(8), 1246–1266.
- Matheron, G. (1967). *Éléments pour une théorie des milieux poreux*. Masson, Paris.
- Matheron, G. and De Marsily, G. (1980). Is transport in porous media always diffusive? a counterexample. *Water Resour. Res.*, 16(5), 901–917.
- Mézard, M., Parisi, G., and Virasoro, M. (1987). *Spin Glass Theory and Beyond: An Introduction to the Replica Method and Its Applications*, volume 9. World Scientific Publishing Company, Singapore.
- Michael, H., Li, H., Boucher, A., Sun, T., Caers, J., and Gorelick, S. (2010). Combining geologic-process models and geostatistics for conditional simulation of 3-d subsurface heterogeneity. *Water Resour. Res.*, 46(5), article no. W05527.
- Mohar, B. (1997). Some applications of laplace eigenvalues of graphs. In *Graph Symmetry*, pages 225–275. Springer, Berlin.
- Mourlanette, P., Biver, P., Renard, P., Noetinger, B., Caumon, G., and Perrier, Y. A. (2020). Direct simulation of non-additive properties on unstructured grids. *Adv. Water Resour.*, 143, article no. 103665.
- Neuman, S. P. and Orr, S. (1993). Prediction of steady state flow in nonuniform geologic media by conditional moments: Exact nonlocal formalism, effective conductivities, and weak approximation. *Water Resour. Res.*, 29(2), 341–364.
- Noetinger, B. (1994). The effective permeability of a heterogeneous porous medium. *Transp. Porous Media*, 15, 99–127.
- Noetinger, B. (2000). Computing the effective permeability of log-normal permeability fields using renormalization methods. *C. R. Acad. Sci. - Ser. IIA - Earth Planet. Sci.*, 331(5), 353–357.
- Noetinger, B. (2020). Statistical physics and applied geosciences: some results and perspectives. *C. R. Phys.*, 21(6), 539–560.
- Noetinger, B., Artus, V., and Ricard, L. (2004). Dynamics of the water–oil front for two-phase, immiscible flow in heterogeneous porous media. 2–isotropic media. *Transp. Porous Media*, 56(3), 305–328.
- Noetinger, B. and Gautier, Y. (1998). Use of the Fourier-Laplace transform and of diagrammatical methods to interpret pumping tests in heterogeneous reservoirs. *Adv. Water Resour.*, 21(7), 581–590.
- Oliver, D. S. and Chen, Y. (2011). Recent progress on reservoir history matching: a review. *Comput. Geosci.*, 15(1), 185–221.
- Paterson, L. (1984). Diffusion-limited aggregation and two-fluid displacements in porous media. *Phys. Rev. Lett.*, 52(18), 1621–1625.
- Potters, M. and Bouchaud, J.-P. (2019). *A First Course in Random Matrix Theory*. Cambridge University Press, Cambridge.
- Renard, P. and De Marsily, G. (1997). Calculating equivalent permeability: a review. *Adv. Water Resour.*, 20(5), 253–278.
- Saffman, P. G. and Taylor, G. I. (1958). The penetration of a fluid into a porous medium or Hele-Shaw cell containing a more viscous liquid. *Proc. R. Soc. Lond. A. Math. Phys. Sci.*, 245(1242), 312–329.
- Santos Oliveira, D., Horowitz, B., and Rojas Tueros, J. A. (2021). Ensemble-based method with combined fractional flow model for waterflooding optimization. *Oil Gas Sci. Technol. - Rev. IFP Energies nouvelles*, 76, article no. 7.
- Scheidt, C., Zabalza-Mezghani, I., Feraille, M., and Collombier, D. (2007). Toward a reliable quantification of uncertainty on production forecasts: adaptive experimental designs. *Oil Gas Sci. Technol. - Revue de l'IFP*, 62(2), 207–224.
- Shraiman, B. I. (1986). Velocity selection and the Saffman–Taylor problem. *Phys. Rev. Lett.*, 56(19), 2028–2032.
- Sorbie, K., Zhang, H., and Tsibuklis, N. (1995). Linear viscous fingering: new experimental results, direct simulation and the evaluation of averaged models. *Chem. Eng. Sci.*, 50(4), 601–616.
- Stauffer, D. and Aharony, A. (2014). *Introduction to*

- Percolation Theory: Revised Second Edition*. CRC Press, London.
- Stepanyants, Y. A. and Teodorovich, E. (2003). Effective hydraulic conductivity of a randomly heterogeneous porous medium. *Water Resour. Res.*, 39(3), article no. 1065.
- Tabelling, P., Zocchi, G., and Libchaber, A. (1988). An experimental study of the Saffman–Taylor instability. In *Dynamics of Curved Fronts*, pages 219–234. Elsevier, Amsterdam.
- Tallakstad, K. T., Knudsen, H. A., Ramstad, T., Løvoll, G., Måløy, K. J., Toussaint, R., and Flekkøy, E. G. (2009). Steady-state two-phase flow in porous media: statistics and transport properties. *Phys. Rev. Lett.*, 102(7), article no. 074502.
- Tang, C. (1985). Diffusion-limited aggregation and the Saffman–Taylor problem. *Phys. Rev. A*, 31(3), 1977–1987.
- Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*, volume 89. SIAM, Philadelphia.
- Teodorovich, E. (2002). Renormalization group method in the problem of the effective conductivity of a randomly heterogeneous porous medium. *J. Exp. Theor. Phys.*, 95(1), 67–76.
- Todd, M., Longstaff, W., et al. (1972). The development, testing, and application of a numerical simulator for predicting miscible flood performance. *J. Pet. Technol.*, 24(07), 874–882.
- Toussaint, R., Måløy, K. J., Méheust, Y., Løvoll, G., Jankov, M., Schäfer, G., and Schmittbuhl, J. (2012). Two-phase flow: Structure, upscaling, and consequences for macroscopic transport properties. *Vadose Zone J.*, 11(3), article no. vzj2011.0123.
- Wilkinson, D. and Willemsen, J. F. (1983). Invasion percolation: a new form of percolation theory. *J. Phys. A: Math. Gen.*, 16(14), 3365–3376.
- Witten Jr, T. and Sander, L. M. (1981). Diffusion-limited aggregation, a kinetic critical phenomenon. *Phys. Rev. Lett.*, 47(19), 1400–1404.
- Yortsos, Y. C. (1995). A theoretical analysis of vertical flow equilibrium. *Transp. Porous Media*, 18(2), 107–129.
- Zdeborová, L. and Krzakala, F. (2016). Statistical physics of inference: Thresholds and algorithms. *Adv. Phys.*, 65(5), 453–552.
- Zdeborová, L. and Krzakala, F. (2007). Phase transitions in the coloring of random graphs. *Phys. Rev. E*, 76(3), article no. 031131.
- Zhang, D. (2001). *Stochastic Methods for Flow in Porous Media: Coping With Uncertainties*. Elsevier, Amsterdam.



Research article

Geo-hydrological Data & Models

# Multiscale modeling of reservoir systems using geostatistical methods

Hayet Chihi<sup>①,\*</sup>, Mohamed Amin Hammami<sup>①,a</sup>, Imen Mezni<sup>①,a,b</sup>, Habib Belayouni<sup>①,c</sup> and Abdallah Ben Mammou<sup>①,d</sup>

<sup>a</sup> Georesources Laboratory, Center for Water Research and Technologies, University of Carthage, Tunis, Tunisia

<sup>b</sup> Tunisian National Oil Company, Tunis, Tunisia

<sup>c</sup> Department of Geology, Faculty of Sciences of Tunis, University of Tunis El Manar, Tunisia

*E-mails:* hayet\_chihi@yahoo.fr, hayet.chihi@certe.rnrt.tn (H. Chihi), mohamedamine.hammami@certe.rnrt.tn (M. A. Hammami), imen.mezni@certe.rnrt.tn (I. Mezni), hbelayouni@yahoo.com (H. Belayouni), abdallah.benmammou@fst.rnu.tn (A. Ben Mammou)

**Abstract.** Identifying factors and quantifying the processes, influencing the extent of reservoir rocks and the distribution of their heterogeneities are fundamental prerequisites to improve the characterization of reservoir connectivity and its dynamic functioning. In this paper, we propose a multiscale approach based on combined geological–geophysical investigations, geostatistics, and 3D geological modeling to build a structural model of faulted-reservoir systems, at regional and local scales within the Jeffara basin, southeast Tunisia.

The regional modeling procedure is calibrated by outcrops, a Digital Elevation Model (DEM), boreholes, and well data logs. The geological processing, itself, is used as a tool to extend the database from a variety of documents such as geological maps, geological cross-sections, outcrop descriptions, and especially the wealthy geological knowledge. The original and extended data are then coded and stored in a common georeferenced database. This provides a detailed input for the geostatistical modeling procedures that enabled to precisely capture the varying extent and shape of the hydrogeological units at regional scale.

Geophysical data, available along the Jeffara plain, were added as a complement to the Jeffara geodatabase and used to establish the fault network and to build a local architectural model of the coastal aquifer system, based on time-to-depth conversion using kriging with external drift.

The results have allowed (i) a major update of the geological configuration of the Jeffara basin, (ii) more precision on the geometries and extent of the stratigraphic sequences and (iii) an accurate prediction of the main characteristics of water reservoirs, i.e. their occurrence, thickness, facies and dynamic properties variation at the basin and reservoir scale. Such multiscale modeling provided effective tools for better understanding the hydrostratigraphic setting, the hydrodynamic functioning of neighboring aquifers, and to assist water resources management.

**Keywords.** Multiscale modeling, Kriging with external drift (KED), 3D Geological modeling, 3D potential field interpolation, Combined geological–geophysical investigations, Uncertainty, Resources management.

\* Corresponding author.

*Manuscript received 30 March 2023, accepted 3 April 2023.*

## 1. Introduction

Reservoir system architecture and heterogeneities, that are controlled by stratigraphic framework and sedimentological trends, are difficult to predict, particularly when all knowledge is derived from scattered boreholes and 2D seismic data. Obviously, modeling procedures are more intricate when it comes to drawing the geological configuration at a regional and local scale and to unraveling the stratigraphic/structural pattern of sedimentary sequences. Such a 3D structural model of the subsurface provides further imperative tools that help efficiently, to (i) gain a better understanding of the stratigraphic framework, the structural setting and the reservoir discontinuities and heterogeneities, (ii) quantify dynamic processes, (iii) assess natural georesources and (iv) make the most appropriate decisions. de Marsily et al. [1992] affirmed, “*We claim that the model helps us today to take the best engineering decisions, and the better the “verification” of the model, the better the “validation” of the model, and, therefore, the better the decision.*”

On the other hand, the challenges in regional investigations and modeling come, mainly from the size and the quality of the data to gather and visualize. Typically, such data are limited, discrete, derived from various sources, and heterogenous, associating one or several scalar values to a set of points in space. Additionally, we suppose that, within each of these regional sequences, the facies and stacking patterns of the reservoir formations, are tectonically controlled, consequently, each of them contains a distinct set of features, that is indigenous to its position within the basin compartments [Mariethoz and Lefebvre, 2014, Caumon et al., 2016]. Uncertainties are usually inherent to these local and regional trend variabilities of the geological phenomena and to the limited size and resolution of datasets [Chilès and Delfiner, 2012, Wellmann and Caumon, 2018].

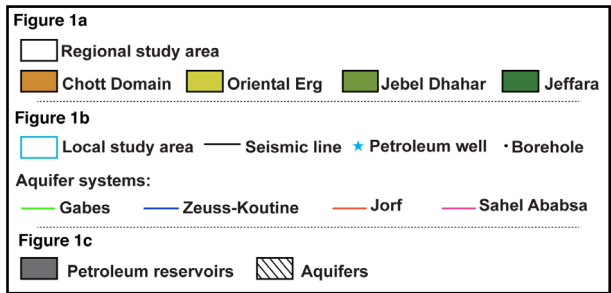
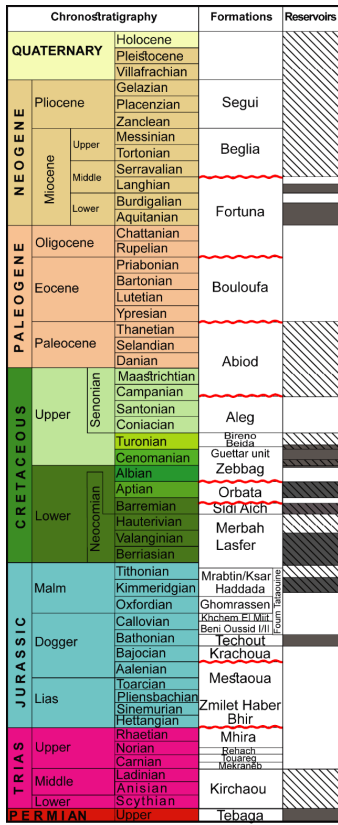
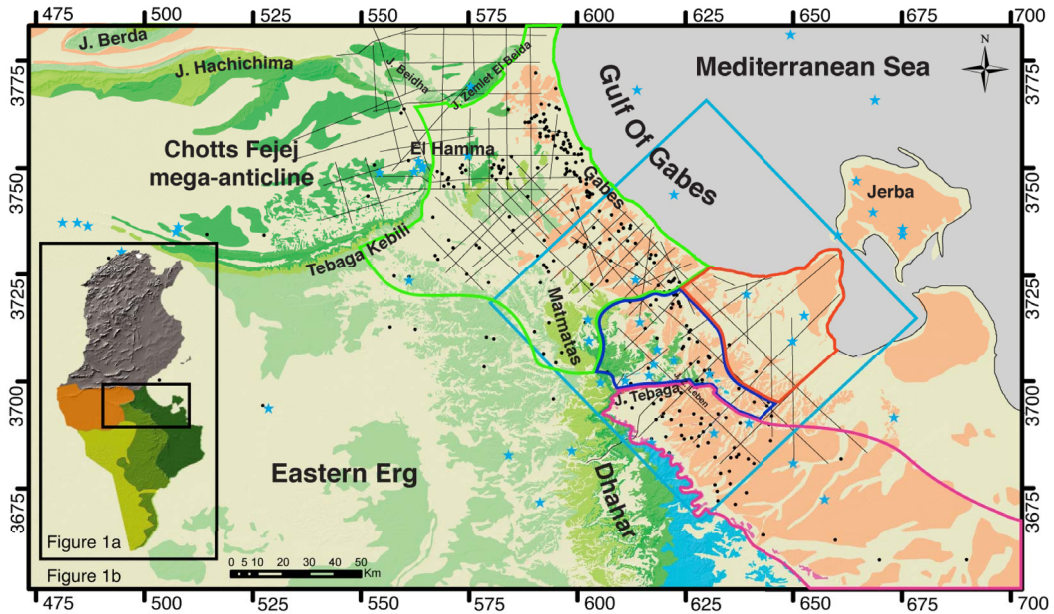
Geostatistics can be used to quantify and reduce uncertainties through (i) its basic tool the variogram that underlies the variability structure and shows sharp changes in the rock properties (bounding surfaces elevations, rock types, hydraulic parameters . . .); and (ii) additional suitable algorithms that allow integration and homogenization of multisource data, and more importantly in-

roducing geological knowledge and concepts in the modeling procedures [Ravenne, 2002a,b, Linde et al., 2015, Caumon et al., 2016, Chihi, 2021].

The main motivation of our work is to provide efficient tools to describe, at both regional and local scales, the geometry and extent of the main reservoir formations of the Jeffara basin in southeastern Tunisia (Figure 1). The strata located in the Jurassic-Quaternary interval of the Jeffara basin show a significant degree of variability in sedimentary facies, leading to the occurrence of local, either carbonate or siliciclastic, reservoirs (of water resources and hydrocarbons) or seals.

The structural geology of the Jeffara basin in southeastern Tunisia is described in detail in the literature, mainly, based upon outcrops [Busson, 1967, Bishop, 1975, Bodin et al., 2010, Bouaziz et al., 2002] and/or geophysical data [e.g. Gabtni et al., 2012, Chihi et al., 2013, Soua and Chihi, 2014]. Most of these preceding studies have been devoted to the structural style and evolution of the main structural features and domains. However, much less is known about the stratigraphic/structural architecture, the regional extent and the facies evolution of the reservoir formations. Our investigations required to pay attention to two important issues. The first is the high complexity of the tectonic environment and the regional fault network over all the Jeffara. Actually, the rocks are outcropping on the western side and are buried at a depth of some 3500–4000 m in the eastern side. Another challenge is the difficulty of predicting the short-distance evolution of lithostratigraphic formations from one structural bloc or domain to another one, mainly within low data density areas. Accordingly, a multiscale modeling was carried out within the Tunisian Jeffara. This large study area provides an excellent example of the trade-off between structural complexity, input data, geostatistical methods and uncertainty of model outputs.

To address these challenges, we propose a multiscale modeling workflow that takes into account data heterogeneity and scarcity, geological concepts and interpretations; while leveraging the potential of geostatistical algorithms to provide adequate and reasonable procedures for integrating multisource data, optimizing the prediction of geological and hydrogeological parameters and quantifying and mitigat-



**Figure 1.** Major domains of the study area (a). Regional geological map, showing the main geological outcrops around the Jeffara basin and the data location (b). Synthetic regional lithostratigraphic chart and major petroleum and water reservoir formations (c).

ing uncertainty. The geostatistical tools, including stationary and non-stationary algorithms (such as ordinary kriging, kriging with external drift and 3D potential field/based cokriging interpolation), were adapted according to the modeling scale and objectives.

This paper presents a part of the results of a comprehensive regional study on the Jeffara basin (Figure 1), undertaken under the research project “Geomodelling and Natural Resources Characterization” framework. It summarizes the recent interpretation efforts carried out to promote the exploration of this challenging area and to develop a better definition of its structural setting and geological evolution. It will be demonstrated that the use of the stratigraphic analysis at the regional scale of sedimentary sequences, based on outcrops, cores and logs, and geophysical data, leads not only to a higher degree of details in the regional fault network and the 3D architectural models but also, to various significantly different interpretations concerning compartmentalization and connectivity of the faulted reservoir units hosted within the multilayered aquifer systems.

The rest of the paper is organized as follows. Section 2 provides a brief geographical and geological overview of the study area. Section 3 addresses the regional modeling of the extended sedimentary sequences to identify the spatial arrangement of regional structural domains and to better understand the relative positioning of the reservoir systems. Section 4 describes the integration of geophysical data to construct a local structural model of the coastal multilayered aquifer systems. Section 5 explains how the various findings and interpretations, presented in Sections 3 and 4, constituted the most imperative input involved in the implementation of a 3D geological model. The explanation is then extended to reveal how such combined modeling approaches can provide effective tools for (i) a better understanding of the hydrostratigraphic framework and (ii) the description of the aquifer systems distribution, (iii) the characterization of the reservoir connectivity within and between all aquifer compartments and (iv) the identification of the groundwater flow pattern, and (v) the regionalization of potential zones for aquifer recharge and aquifer system management. In Section 6 we summarize the major challenges faced during the different modeling processes, discuss the re-

sults and findings and propose solutions for future work and applications for both the present study area or other regional and/or local case studies. Section 7 outlines the concluding remarks, summarizing the main findings and results of the study.

## 2. Overview of the study area

The study area (Figure 1) is extended over a large domain of southeastern Tunisia that covers the Jeffara basin and the surrounding Dhahar-Matmata Mountains to the south and the west, the Chotts Fejej mega-anticline to the north, the Gabes region to the northeast, and Jerba island and a part of the Gulf of Gabes within the Mediterranean Sea to the east. It is characterized by an arid climate exposed to both Mediterranean and continental influences, with high annual average evapotranspiration of around 2700 mm, low average rainfall reaching about 200 mm/year and an average temperature ranging between 12 °C and 30 °C.

The complex geological configuration is inherited from several successive tectonic events that affected the sedimentary units distribution and controlled their lateral facies and thickness variation. This geological architecture is mainly featured by several normal major NW–SE and minor NE–SW faults. The most prominent one is the Medenine normal fault, oriented NW–SE. It subdivides the Jeffara basin into (i) the southwest uplifted domain marked by the outcropping of the Albo-Aptian and Lower Cretaceous terrains and (ii) the northeast collapsed domain covered by a Mio-Plio-Quaternary (MPQ) layer. The tectonic evolution has resulted in the sculpting of the major geological structures in the study area [Busson, 1967, Bouaziz *et al.*, 2002]. These are namely (Figure 1a,b) (i) the northern and the southern Chotts chains surrounding the Chotts Fejej mega-anticline, (ii) the Dhahar-Matmata high mountains corresponding to a large monoclinical plunging westward until it finally dives under the dune formations of the Oriental Erg, (iii) the Jeffara coastal plain constituting the collapsed eastern flank of the Dhahar, which was buried under the continental deposits of the MPQ sediments, and (iv) other structural highs such as Jebel Tebaga of Medenine, Jebel Zemlet Leben and Jebel Tejra.

The study area, including the Jeffara basin and surrounding domains, is distinguished by an extensive stratigraphic series that extends from the Permo-Triassic to the Quaternary period (Figure 1c). Several regional and local unconformities dated as Early Jurassic, Late Aptian–Early Albian, intra-Aleg unconformities, Late Cretaceous and Oligocene–Miocene, in conjunction with thickness and facies variations, have been recorded as a result of continuous tectonic activity involving the reactivation of deep faults [Letouzey and Trémolières, 1980, Ben Youssef and Peybernes, 1986, Klett, 2001, Bodin *et al.*, 2010, Chihi *et al.*, 2013]. The major potential oil and water reservoir formations are hosted within (1) the Permian represented by Tebaga and Zoumit formations, (2) the Triassic marked by the Kirchaou formation, (3) the Jurassic consisting of Techout and M'Rabtime formations, (4) the Lower Cretaceous represented by Sidi Aich, Merbah Lasfer and Orbata formations, (5) the Cenomanian materialized by the Zebbag formation, (6) the Turonian composed of Bireno and Beida formations, (7) the Senonian represented by Abiod and Aleg formations, and (8) the Miocene hosting the Ain Ghrab formation [Ben Ferjani *et al.*, 1990, Peybernes *et al.*, 1993, Chihi *et al.*, 2013, Jabir *et al.*, 2020, Kraouia *et al.*, 2022].

Important multilayered aquifer systems are extending along the Jeffara basin, they are largely interconnected through the dense fault network [Ben Baccar, 1982, Mammou, 1990, Chihi *et al.*, 2015, Hammami *et al.*, 2018a, Mezni *et al.*, 2022a]. The main are (Figure 1b) (i) Gabes aquifer system, hosted mainly in the Mio-Pliocene in northern Gabes, in the Carbonate Senonian and Turonian in Southern Gabes, and in the Continental Intercalary in western Gabes, (ii) Zeuss-Koutine aquifer embedded in the Jurassic, Albo-Aptian, Turonian and Senonian intervals, (iii) the Sahel El Ababsa aquifer enclosed in the Triassic sandstone series, and (iv) the Mio-Pliocene Jorf aquifer containing the detrital sediments of the MPQ. The Jeffara basin is featured by a highly dense hydrographic network articulated around the major “wadis” (or non-perennial rivers) that drain rainwater from the high Dhahar-Matmata chain to be discharged to the northeast, particularly towards the Gulf of Gabes and the Sebkhass.

### 3. Regional scale modeling

Regional-scale modeling was conducted to enhance geologic knowledge, promoting a necessary update of the stratigraphic/tectonic architecture all over the Jeffara basin, and a better understanding of the reservoir systems arrangement and the regional extent of reservoir formations.

#### 3.1. Methodology

We present an original multi-approach methodology to build a regional 3D architectural model of the Jeffara basin, using geostatistical algorithms. It is based on a sequential analysis and modeling of the bounding surfaces and the thickness variations, of the regional sequences belonging to the Cretaceous intervals. To make this paper somewhat shorter and more concise, we will focus only on three regional Cretaceous sequences, namely, the Lower Cretaceous, the Albian–Cenomanian and the Upper Cretaceous sequences (Figure 1c).

##### 3.1.1. Data management

A dense dataset (Figure 1, Table 1), consisting of (i) 12,000 km seismic lines and 40 exploration petroleum wells provided by the Tunisian Company of Petroleum Activities (ETAP), (ii) 370 water boreholes, mainly deriving from the General Directorate of Water Resources (DGRE), that were carried out over the past 60 years across the entire Jeffara basin, and (iii) 8 assembled geological maps acquired from the National Office of Mines (ONM), forms the foundation for the architectural model construction. This database represents the most detailed and extensive sampling ever considered in such a modeling procedure in the region. However, this significant amount of data is of variable quality and frequency. The configuration shows that in some areas data are very *clustered* and can provide too much information to take into consideration (for some local kriging estimation and for visualization on the geologic maps and later, on calculated models). On and next to outcrops settings, mainly in the southwestern side area, data are typically *scarce*. Moreover, the alternative of using seismic-reflection data is not a practicable option. The seismic data are very much degraded next to outcrops and faulted zones, elsewhere because of sub-seismic resolution. In fact, the low

**Table 1.** Datasets providing inputs for the development of the regional, local and 3D geological models

Data	Notes	Source
Satellite Mission—SRTM with 1 arc-second (approximately 30 m) resolution	Digital Elevation Model (DEM)	USGS database (earthexplorer.usgs.gov/)
Geological maps	<ul style="list-style-type: none"> <li>• 8; 1/100,000</li> <li>• Gabes, Mareth, Ajim, Oglet Mertba, Medenine, Tamezret, Ghomrassen and Tataouine,</li> <li>• A surface of approximately 43,750 km<sup>2</sup></li> </ul>	ONM
Digitized outcrops	8.958 points with a resolution of about 1 point/500 m <sup>2</sup>	Carried out within the framework of the Jeffara research project
Seismic data	18 Seismic sections	ETAP
Exploration Petroleum/Well logs	40	ETAP
Boreholes	370	DGRE
Geological cross-sections	60, the major were created for calibrating the 3D geological model	Established within the framework of the Jeffara research project
Piezometric Data	68	DGRE

USGS: United States Geological Survey, ONM: National Office of Mines, ETAP: Tunisian National Oil Company, DGRE: General Directorate of Water Resources.

vertical seismic resolution greatly affects the quality of seismic interpretation in certain areas. Some seismic markers may not be continuously available, and some others are below vertical seismic resolution, making them difficult to identify. Consequently, the database needed to be augmented with stratigraphic details derived from alternative sources, namely the numerous *geological cross-correlation sections* that were established and all outcrops information derived mainly from the geological maps and previous field studies covering the whole area. Nevertheless, the *seismic* interpretation helped to map the major crosscutting faults.

On the other hand, an assembled geological map (Figure 1) covering Gabes, Mareth, Ajim, Oglet Mertba, Medenine, Tamezret, Ghomrassen and Tataouine, with a surface of approximately 43,750 km<sup>2</sup>, was imported into ArcGIS software [ArcGIS, 2012]. All landforms, the surrounding outcrops and Jebels were sampled to obtain a regular distribution of data points. The gridding technique was performed in a way that enables the resampling of the geologic features taking into account the structural details such as lithological boundaries, elevation variation of the topographic surface for each

geologic outcropping formation, fault traces, etc. In fact, the grid's dip and azimuth were defined to provide an optimal surface sampling. The size of the grid cells was defined so that the frequency of the sample points could be large enough to cover rightly the outcropped geological units but not exceed the borehole's frequency. A point cloud of 8.958 points with a resolution of about 1 point/500 m<sup>2</sup> was generated.

Taking into account the details needed for the modeling procedure and the further applications of the 3D model, the multisource data described above were harmonized. The different datasets derived from geological maps, field information, water boreholes and petroleum exploration wells were integrated in such a way they fit together with respect to both geometry and semantics. In fact, we have elaborated a unified stratigraphic scheme that defines the lithostratigraphic sequences, their units, and their regional boundaries which are the sequence unconformity/relative conformity surfaces (Figure 1c). Furthermore, all datasets were organized in a GIS environment using ArcGIS software to edit and generate any geological and hydrogeological information that would be useful during the different modeling pro-



cesses.

### 3.1.2. *Methods overview*

The regional architectural model typically involves two main stages, that were carried out simultaneously.

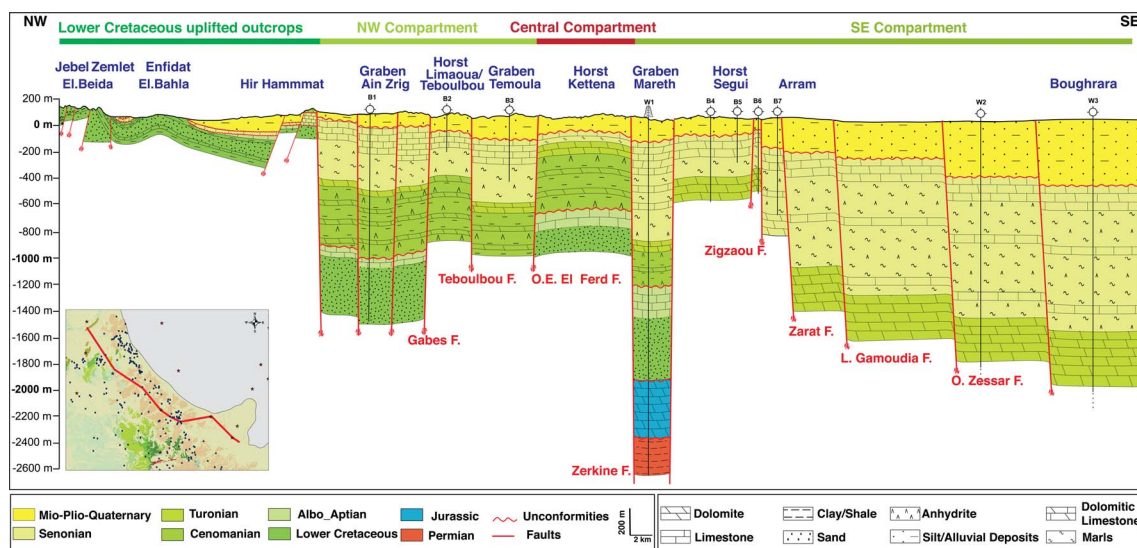
**Geological investigation.** A detailed geological investigation was carried out, focusing mainly on an analysis of the stratigraphic/tectonic framework of the Cretaceous-MPQ sequences of the Jeffara basin, based on the unified stratigraphic scheme (Figure 1c), and using the multisource data described above. However, the modeling will concern only three stratigraphic intervals. The first is the *Lower Cretaceous sequence* that is delimited at its upper boundary, by the Aptian–Albian unconformity surface and at its lower boundary, by the Upper Jurassic surface. It includes the Merbah Lasfer, Sidi Aich and Orbata formations. The sediments are composed of clastic alluvial materials from the Saharan Platform [Bishop, 1975]. The Merbah Lasfer formation (Berrasian to Hauterivian) is mainly composed of sandstones and dolostones. Barremian rocks consist of Sidi Aïch sandstones. Aptian rocks are mainly composed of limestones and dolostones of the Orbata formation. Bishop [1975], Ben Ferjani *et al.* [1990]. The second, and the third sequences belong both to the Upper Cretaceous interval. The *Albian–Cenomanian* sequence is bounded by the Late Aptian–Early Albian unconformity surface at the bottom and the Upper Cenomanian at the top; it includes the Albian to Cenomanian Zebbag formation. It was deposited sealing unconformably the Orbata formation and is composed primarily of carbonate rocks. The Turonian–Senonian sequence extends in between two stratigraphic limits, the Upper Cenomanian on the bottom and the Late Cretaceous unconformity on the top. The Turonian to Campanian sediments overly the Zebbag formation. They consist of the Beida formation composed mainly of anhydrite and minor amounts of dolomite, the Bireno formation which consists of limestone and marls and the Aleg formation including mudstone, limestone, and marls. The Abiod formation is usually composed of fractured chalky limestones but is fairly *eroded* on a large domain of the study area [Ben Ferjani *et al.*, 1990].

To investigate the structural geometry and lateral facies variations, we constructed two sets of numerous parallel NE–SW and NW–SE oriented geological cross-sections, perpendicular to structural trends. One of them is illustrated in Figure 2. We emphasize here that, for purposes of understanding the stratigraphic/tectonic framework and subsequent geometric modeling needs, we have plotted vertically for each well, the hiatuses and depositional sequences, matching their appropriate ages. The correlation of the studied intervals, mainly the Lower to Upper Cretaceous stratigraphy, was complicated first by, obviously, the tectonic activities but also by these unconformities because in many areas they were erosive and removed a large part of the underlying sequence and even the entire formation (e.g. the Abiod formation). It was important in this study to pay special attention to the detailed occurrence of each sequence within each well and borehole, and identify eventual hiatuses by graphic correlation based on outcrops, as described on each of the eight geologic maps, field studies, seismic sections and previous works within the study area [Chihi *et al.*, 2013, 2016, Mezni *et al.*, 2022a].

**Geostatistical modeling.** The stratigraphic–structural model, within the Jeffara basin, was constructed using the stack-unit mapping approach. Stratigraphic sequence boundary maps were estimated by kriging interpolation using ISATIS geostatistical algorithms [ISATIS, 2020]. The sequences' thickness maps were then calculated by subtraction from kriged lower and upper boundaries. However, the subtraction process was carefully performed so that the resulting maps appear geologically plausible.

We recall that approximately 370 control water boreholes and 40 petroleum exploration wells were augmented with more than 8.958 ancillary data generated by digitizing the surrounding outcrops in the study area (Table 1, Section 3.1.1). Furthermore, the established cross-correlation profiles in addition to the derived pseudo-wells were used to supplement the database. Elevation data (in meters) were imported into the ISATIS software as a data point file “X, Y, elevations”.

An adequate understanding of the regional properties of the lithostratigraphic sequences and of the limitations imposed by the data is absolutely necessary to adapt the geostatistical approaches



**Figure 2.** NW–SE regional geological cross-section drawn sub-parallel to the shoreline within the Jeffara coastal plain. It shows globally a horst and graben structuring that is installed through NE–SW striking normal faults, with lateral facies and thickness variation. Note at the extreme NW the raised fold of Zemlet El Beida featuring the uplift and outcrop of the Lower Cretaceous series. It also illustrates, as indicated on the top, the impact of the NE–SW normal faults on the geographical pattern and the extent of aquifers and the various compartments, defined in this study as Lower Cretaceous, NW, Central and SE compartments.

and their input requirements in order to construct robust architectural models, minimize the associated uncertainty and honor the geological realism. However, in carrying out the data integration, we should take into account a series of factors that may be crucial for the quality of the end results. The first factor concerns the degree of detail aimed for the geological objects to be modeled and their interpretation. For us, the actual aim of the study is to model the regional stratigraphic–structural sequences using mainly a cross-section-based approach. The second factor consists in performing an adequate integration of the various types of all information described above to conduct the regional modeling. This could generally allow the set of major faults running through the study area to be reconstructed with sufficient accuracy and to draw easily the structural elements easily at a large scale. In our previous work on the Jeffara basin [Chihi et al., 2013, 2014, Hammami et al., 2018a], we tend to include in our modeling procedure all those discontinuities which can be identified from the seismic survey and established lithostratigraphic cross-correlation profiles (regardless of whether these have an impact on

fluid flow), the purpose was to optimize the reservoir modeling procedures. This could not be adequate for the regional modeling, in this section, as it may influence the draw of isovalues' curves. In fact, this could disturb the major lineament continuity and modify the structural elements' shape and extent.

As an effect, we decided to elaborate an adequate simplified modeling procedure that allows optimizing the mapping at a regional scale, to highlight the regional major faults which turn out to have a significant impact on the structuring of the Jeffara basin and on the geometry and extent of its most important sequences.

### 3.2. Modeling the stratigraphic sequences

Geostatistical methods were used to study the spatial variability and the geological continuity of the sequences' boundaries [Caumon et al., 2016, Chihi, 1997, Chihi and de Marsily, 2009, Chihi et al., 2016]. First, experimental variograms of the different "elevation" variables were calculated along the principal structural directions (i.e. normal faults strik-

ing northwest–southeast and northeast–southwest), to identify the underlying variogram components. Figure 3 displays the elementary variograms in both directions N45 and N135 for the different geological horizons.

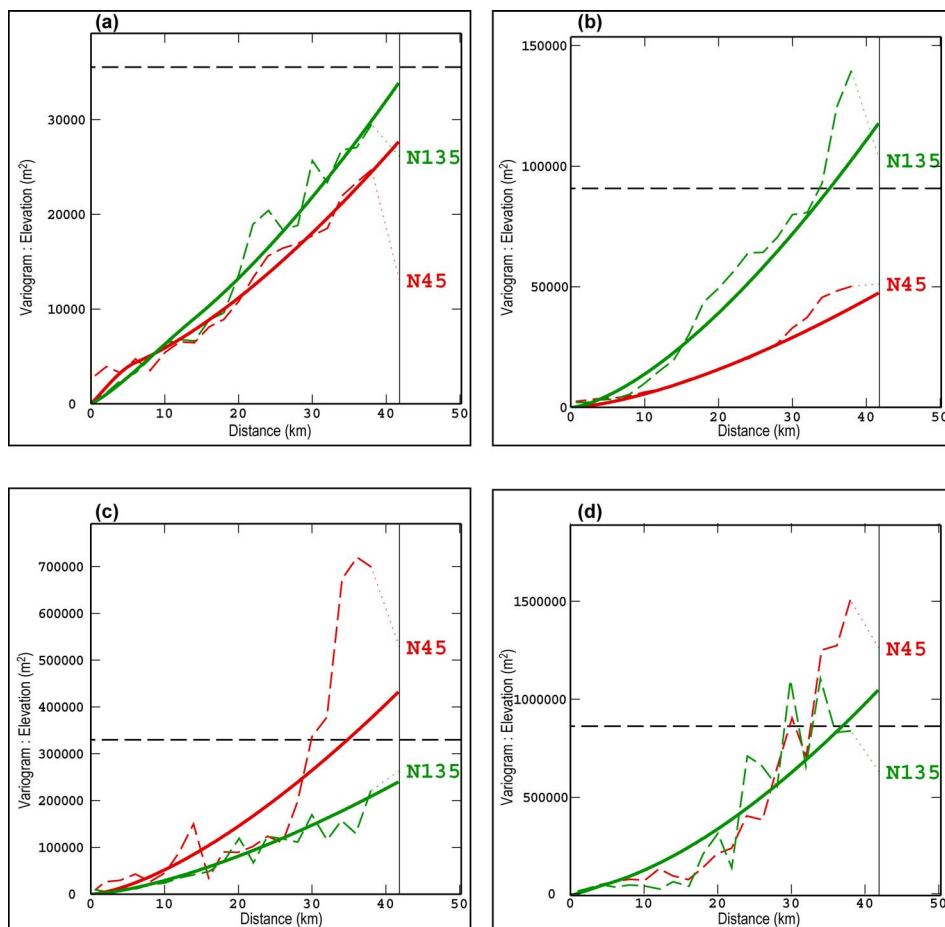
The Upper Cretaceous variable shows an anisotropic behavior (Figure 3a). We observe (i) a stationary structure describing the local depth variability and (ii) a non-stationary structure expressed along both the northeast–southwest and northwest–southeast directions. These drift structures reveal a downtilting of the Upper Cretaceous horizon that is more important in the northwest direction. The directional variograms of the Cenomanian horizon display a more pronounced anisotropic and non-stationary behavior with a higher variability along the northwest direction (Figure 3b). The Aptian variograms (Figure 3c) exhibit (i) an isotropic stationary component that describes the depth variability at a local scale and a nonstationary anisotropic drift that is more clearly expressed along the northeast. The variability of the Jurassic horizon shows an isotropic behavior and reveals a complex structure consisting of a local stationary component and a drift (Figure 3d). Table 2 summarizes the different model parameters (the range and the sill) obtained for each “elevation” variable.

The variogram models were used to estimate successively the “elevation” variable for each sequence, at all unsampled locations by ordinary kriging on a grid covering an area of 175 km × 150 km discretized into 1 km square cells. A moving window with 8 angular sectors procedure was used to select the optimal number and configuration of the neighboring samples of the point to be estimated. This ensured to sample the most appropriate number of points along all directions as consistently as possible, the closest ones to infer local variability and the farthest ones to respect the non-stationarity of the variable. For each point estimation, we calculated the associated uncertainty defined by the standard deviation. Standard deviation maps were then generated for all estimated horizons to assess the estimation efficiency. We display, as an example, that of the Upper Cretaceous horizon in Figure 4. The values vary from 1 to 5 m around the observed points, up to 15 m where the data density is relatively lower. Accordingly, the estimation is considered sufficiently accurate. At the northwest and

the southwest edges of the study area, the kriging standard deviation is much higher because data are scarce. We emphasize that during the modeling process, we have to consider comprehensively the influence of missing data in order to provide a means for improving the estimation all over the studied domain, i.e. (i) in areas with high but uneven data density and (ii) more importantly along the northwest and the southwest sides of the study area showing a few data. In fact, the first estimation trials required supplementing the database with some secondary data, derived from outcrops and geological cross-sections, which constituted an important part of the modeling approach and had a major impact on enhancing the kriging results. This would provide a better confidence for both the geological interpretation of the kriged surfaces and the generated thickness maps.

We, therefore, consider our interpretation in the reliable areas of the estimated maps while paying close and specific attention to the shaped geological features within sectors showing a relatively low data density. The thicknesses of the studied sequences that were estimated by subtracting the kriged elevations of the boundary surfaces, were validated systematically through a thorough comparison with the established geological cross-sections, the available seismic profiles in some areas and the general geological configuration as displayed on the geological maps and on outcrops.

On the other hand, it should be emphasized here that the thickness calculation resulted in negative values at some specific locations. This reveals undesired crossings between successive horizons which are geologically implausible, in these locations, and need to be investigated and corrected. Inconsistent thickness values can occur in the extrapolation zones (Figure 4) where few data are available, particularly in areas of low thickness, such as in the northwest of the study area. These inconsistencies were addressed by adding dummy data generated from the geological correlation profiles as stated above. The same can, occasionally, be observed in areas where the data density is relatively high but where there is a hiatus, the successive geological interfaces are naturally overlapping, for example in the southwest of the study area. As the surfaces are estimated by different variograms, they will have different but very close estimated values. At these points, if one of the thick-



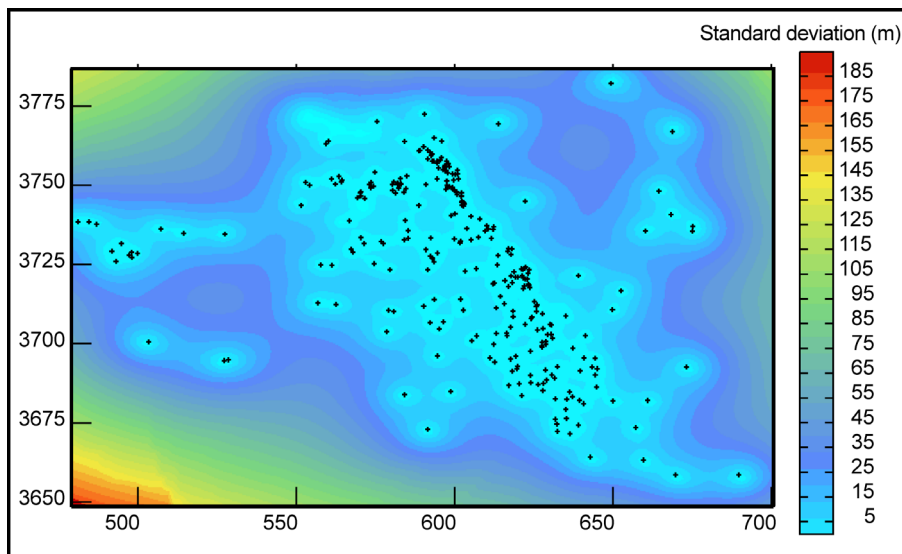
**Figure 3.** Directional experimental (---) and theoretical (—) variograms of the Upper Cretaceous top (a), the Cenomanian top (b), the Aptian top (c) and the Jurassic top (d). The variograms display a drift component that indicates a notable increase in depth, revealing a downward tilting surface boundary.

**Table 2.** Theoretical models of “elevation” variables for each horizon

Variables	Structure	Range (m)		Sill (m <sup>2</sup> )	Equation
		N45	N135		
Cretaceous horizon	Spherical	6,000	14,000	3,000	$\gamma(h) = 3000 \text{ Spherical } (6000_{N45} + 14,000_{N135}) + 20,000 \text{ power } (36,000_{N45} + 31,000_{N135})^{1.5}$
	Power	36,000	31,000	20,000	
Cenomanian horizon	Power	30,000	55,000	7,200	$\gamma(h) = 7200 \text{ power } (30,000_{N45} + 55,000_{N135})^{1.5}$
Aptian horizon	Spherical	2,800	2,800	3,000	$\gamma(h) = 3000 \text{ Spherical } (2800) + 33,000 \text{ power } (35,000_{N45} + 12,000_{N135})^{1.5}$
	Power	35,000	12,000	33,000	
Jurassic horizon	Spherical	5,000	5,000	25,000	$\gamma(h) = 25,000 \text{ Spherical } (5000) + 220,000 \text{ power } (16,000)^{1.6}$
	Power	16,000	16,000	220,000	

nesses is calculated as a negative number, it is set to zero. This ensures that the thickness of the differ-

ent geological sequences is accurately represented in the model. A comparison of all the calculations with



**Figure 4.** Kriging standard deviation map of the Upper Cretaceous top. The estimation is reasonably accurate around the observed points. At the northwest and southwest boundaries of the study area, the estimation error is higher due to the scarcity of available data.

the measured data points and the geological cross-sections was undertaken to further check the accuracy of the generated maps.

### 3.3. Visualization and interpretation

The calculated maps (Figure 5a–d) display the elevation variability of the sequences' bounding surfaces defined by the Jurassic, the Aptian, the Cenomanian and the Upper Cretaceous horizons. The geological interpretation is derived from a detailed examination of the geometry spatial variation of these elevation maps that is also supported by the thickness maps of the Lower Cretaceous, Albian–Cenomanian and Upper Cretaceous sequences (Figure 5e–g). The structural interpretation is based on trend analysis and local slope variation, perceived on the elevation maps in order to ultimately delineate the structural ensembles. These results are verified through comparison with geological cross-sections, seismic lines' interpretations and previous modeling studies in the study area [Chihi et al., 2013, 2014, Hammami et al., 2018b].

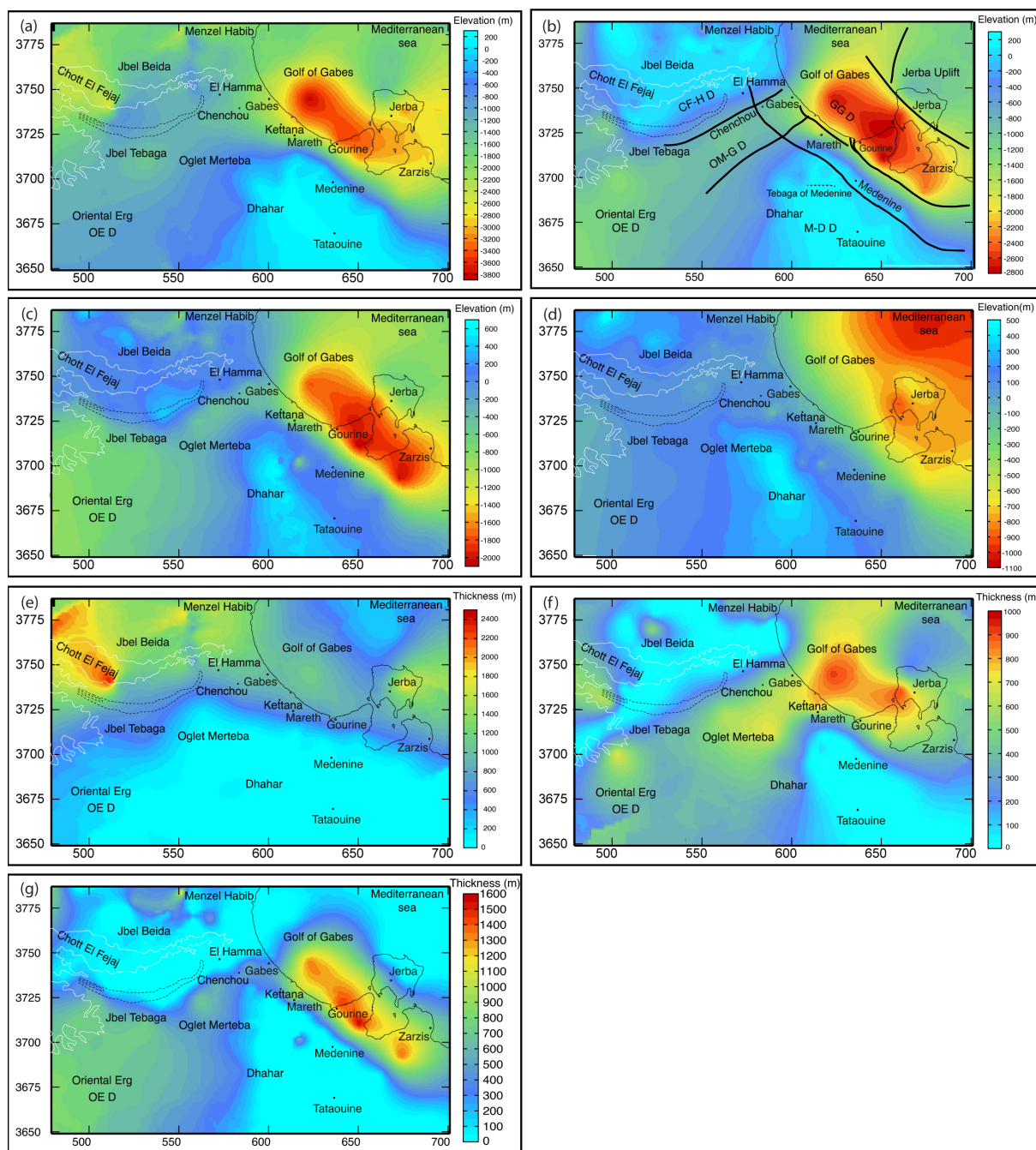
The structural feature alignment and the inherited surface morphology probably played an important role in the extent and thickness of the sedimentary sequences. Accordingly, for each sedimentary

sequence, we will describe the lower boundary and the thickness map. Globally, elevation and thickness maps (Figure 5), show that the Jurassic–Cretaceous strata exhibit an irregular topography and complex geometry while dipping toward the Gulf of Gabes.

On the other hand, seven major prominent domains are distinguished. They remain relatively persistent and exhibit similar patterns across the calculated maps, but their geographic locations shifted at different times. They are marked on Figure 5b: (i) the Medenine-Dhahar Domain (M-D D) in the southern sector of the study area; (ii) Oglet Mertba–Gabes (OM-G D) and (iii) Chotts Fejej-El Hamma Domain (CF-H D), in the northwestern sector; (iv) the Oriental Erg (OE D) on the west; (v) Gulf of Gabes Domain (GG D), along the coast from Kettana to Jorf and shallow part of the Mediterranean, (vi) the Mediterranean Domain (MD) representing the rest of the Gulf of Gabes, and (vii) the Jeffara plain.

#### 3.3.1. The Lower Cretaceous sequence

Sediments of the Lower Cretaceous series, deposited on the top of the Jurassic formations, regionally show a deepening trend emanating from all around Medenine-Dhahar (M-D D) structure



**Figure 5.** Estimated maps of the Jurassic (a), Aptian (b), Cenomanian (c) and Cretaceous (d) horizons; derived thickness maps of the Lower Cretaceous (e), Albian–Cenomanian (f) and Upper Cretaceous (g) sequences. Seven major structural domains were distinguished, namely: Medenine-Dhahar, Oglet Mertba, Chotts Fejej-El Hamma, Oriental Erg, Jeffara plain, Gulf of Gabes and Mediterranean. They are portrayed in Figure 5b, but remain relatively persistent and exhibit similar patterns across the kriged maps. It should be noted, however, that their geographical positions shifted at varying intervals.

(Figure 5a), albeit with variable slope intensity and a main deepening trend along the northeast direction.

The Jurassic upper surface is tilted northeastward from about 200 m in the south of (M-D D) up to 2–4 km in the northeast. This deepening is even more pronounced offshore of Gabes, along a NW–SE elongated depression, we named the Gulf of Gabes domain (GG D). Towards the northeast, in the Mediterranean Sea, the dipping gradient is locally reversed in favor of a considerable increase of elevation, up to –1400 m, within the Mediterranean Domain (MD).

This surface is also significantly tilted northward of (M-D D). The deepest northern area constitutes an extended relatively flat-lying area covering the Chotts Fejej-El Hamma Domain (CF-H D), with elevations ranging between –1200 m to –1600 m. These trends converge along a NE–SW lineament that we will be referring to as Oglet Mertba–Gabes (OM-G D) downlift. It is observed along Oglet Mertba, Chenchou and Gabes extended area, with a relatively uniform elevation ranging between –1300 m and –1700 m. In the southwestern part of the basin, a gentler and steadier gradient is observed towards the Oriental Erg (OE D) in an ESE–WNW direction.

The calculated thickness map of the Lower Cretaceous (Figure 5e) reveals two sedimentary zones, namely a shallow and a deep domain. The shallow domain is characterized by relatively little to no presence of Lower Cretaceous formations; these areas are located on the Dhahar dome and its continuity in the continental Jeffara. The paleogeographic evolution of this region indicates that it had been uplifted, since the end of the Jurassic. On the periphery of this region and around the Matmatas, the Lower Cretaceous sediments are relatively thin and hardly exceed 500 m of thickness. The deep domain extends within two main subsiding areas: northward in the Chotts region and eastward within the maritime Jeffara and along the offshore within the Gulf of Gabes Domain and Jerba to the east.

### 3.3.2. *The Albian–Cenomanian sequence*

The Aptian surface elevations range between 300 and –2800 m (Figure 5b) and exhibit globally a similar elevation distribution as the Jurassic surface roughly within the same domains. A similar dipping, but with a clear steeper gradient, is observed orthogonally to the M-D D and heads towards the

Mediterranean Sea (GG D) on the east and to the Oriental Erg (OE D) on the west. This surface, also, significantly deepens northward and portrays a clearer NE–SW elongated “Gabes Oglet Mertba” depression (–1000 m and –1200 m), before reaching a high area, covering the Chotts Fejej-El Hamma Domain (CF-H D). Elevation values increase, considerably above mean sea level over 200 m. The Oriental Erg is getting distinguished as a depression with well-drawn contours. On the eastern side, the downlifted Gulf of Gabes Domain (GG D) is also recognizable but with a significant change in its shape, with a larger NE–SW extent, it is deeper in its central zone, around Jorf and Jerba island reaching values down to –2800 m. The deepening of the Aptian is reversed towards the Mediterranean. Through correlation, the elevated Mediterranean uplift for the Jurassic map exists in the Aptian map as well, albeit not peaking in the same area of the Gulf of Gabes but in Jerba instead.

The *thickness of the Albian–Cenomanian sequence*, varying from 0 m to 1000 m, confirms multiple stated observations (Figure 5f). The Thickening trend seems to conform with the deepening trend in the Lower Cretaceous and, the lineaments observed on the elevation map can be recognized in the thickness map as well. In this case, the elevated areas of the Dhahar and of Chotts Fejej correspond to the areas where the Albian is almost inexistent or with very low thickness ranging between 0 to 100 m. The structural domains of (GG D), Oglet Mertba and Oriental Erg and their limits are recognized in the thickness map showing the most important values. The maximums, 800 to 1000 m, are within the Gulf of Gabes Domain.

### 3.3.3. *The Upper Cretaceous sequence*

Sediments of the Upper Cretaceous lay directly on the *Cenomanian surface* (Figure 5c) which obeys the same structural trends observed in the previous maps and also shows three major deepening directions. The same structural domains are recognized, but with some variations. In fact, the NE–SW deepening that radiates from the Dhahar has a gentler slope in the continental Jeffara of Gabes and of Medenine area. The Gulf of Gabes downlift which represents the peak of the deepening towards the Mediterranean Sea, is more prominent. The progressive shifting observed of the Jurassic and the Lower Cretaceous se-

ries continues for the Cenomanian series as this domain now covers the entirety of Jorf and Zarzis where the dipping reaches  $-2100$  m. The elevated Jerba uplift, more to the northeast, seems to be of lesser importance than in the underlying layers. In the NW–SE direction, the two opposite dipping trends seen in the Aptian map are also present and converge to the deep lineament in the area of Oglet Mertba, Gabes and Chenchou. The northernmost Chotts Fejej uplifting domain is more extended and shows a spectacular rising as elevations range from 0 to 600 m. To the West, the Oriental Erg domain shows a better expressed shape with a greater extent and a softer slope that rotates towards the southwest.

However, the *Upper Cretaceous* elevation map behaves differently when compared to the underlying sequences (Figure 5d). The Medenine-Dhahar domain remains present albeit with much gentler deepening trends than those of the Cenomanian and the Lower Cretaceous with elevation ranging between 200 and 500 m. The Fejej-Hamma domain is also more attenuated and retracts to the northwest. The Oglet Mertba–Gabes and Gulf of Gabes domains are no longer distinguishable. However, a noticeable elevation dipping can be perceived in northern Gabes and the Oum Zessar areas down to  $-800$  m in Jerba. In fact, the Upper Cretaceous top seems to conform with the topographic surface, the exceptions being northern Gabes and the Oum Zessar–Jerba areas where important Neogene layers are deposited (Figures 5d and 1). Unlike the lower sequences where elevation increases within the Mediterranean domain, this trend is reversed in the Upper Cretaceous top elevations dip down to  $-1100$  m towards the northeast. The Oriental Erg deepening is much softer with a slower but more uniform slope that seems to further rotate towards the southeast similarly to the Cenomanian top surface (Figure 5c).

The *thickness map* (Figure 5g) *actually* brings more information on the Upper Cretaceous structural setup as the trends observed in the underlying sequences can be seen once more due to the inherited morphological setup of the Cenomanian layer. The Upper Cretaceous is thin to non-existent in the Medenine-Dhahar domain starting from where the thickening begins. Towards the north and the northeast, the thickness increases reaching up to about 600 m in the same Oglet Mertba, Chenchou and Gabes corridor, and 800 m in the Gulf of Gabes do-

main. It decreases once more in the Fejej-Hamma domain where it coincides with the outcropping of the Lower Cretaceous to the East of Hamma and the important Neogene deposits replacing the Upper Cretaceous in northern Gabes. Important Upper Cretaceous deposits can be observed again in the area of Menzel Habib in the northwest.

#### 3.3.4. *Insights on the hydrogeological setting*

In light of the above *interpretations, correlations with the hydrogeological setting* of the region can support these results. Indeed, the setting of the aquifer systems of the region is closely related to the structural evolution and paleogeography of the area. Considering that the Jeffara plain is a result of the Atlasic orogeny at the limit of the Saharan platform, its tectonic collapse contributed to the obstruction of the violent Atlasic movements still observed on its limits, causing the division of the region into uplifted and downlifted domains. These tectonic events resulted in a specific structural arrangement of sedimentary sequences and fault networks within each domain, making the aquifer systems they host very distinct in nature. Regionally, the main trends interpreted from the elevation and thickness maps are (i) the uplifted compartments of Chotts Fejej–El Hamma and Medenine-Dhahar as well as Jerba to a lesser extent; (ii) the downlifted NE–SW compartments of Oglet Mertba–Gabes and the NW–SE Gulf of Gabes; and (iii) the relatively stable Oriental Erg domain. The collapse of the coastal Jeffara and its individualization from the rest of southern Tunisia was the result of post-Cretaceous tectonic compression that reactivated and reworked the fault networks that once controlled the Cretaceous and Jurassic deposition [Bodin *et al.*, 2010]. The resulting downlifted and uplifted domains can be delineated on all elevation and thickness maps. However, they are becoming increasingly prominent within the upper layers (Figure 5a,e). The uplifted Medenine-Dhahar domain contains two important aquifers included in the Jurassic and the Lower Cretaceous, which belong to the so-called Zeuss-Koutine aquifer system located in the Medenine region. This system extends to the northeast downlifted coastal Jeffara domain by the Turonian aquifer and joins the aquifer system of Gabes in the northwest (Figure 1b). Although the Chotts Fejej–El Hamma domain is also an uplifted domain, it exhibited a different behavior



during the Jurassic, with a thickness of up to 1600 m, which indicates that this domain was a subsiding zone (Figure 5a). The uplifting trend becomes progressively more visible in subsequent maps, starting from the Aptian. This exposed high domain faced an intense erosion of the ante-Miocene deposits, on the Lower Cretaceous outcrops in particular [Bishop, 1975]. The eroded sediments were deposited in the northern Gabes region where the Neogene aquifers lay directly on the Upper Cretaceous formations. Neogene deposits can also be found within Oum Zessar and Jerba areas and are a testament of the extreme erosion that occurred in the post-Cretaceous and led to the flattening of the Jeffara (Figure 5d,g). This open-cut erosion also contributed to the karstification of carbonates outcropping in the Cretaceous and Jurassic formations (i.e. Zeuss-Koutine and Gabes aquifer systems). The collapsing has also contributed to the preservation of formations hosting the Jeffara aquifer system within the grabens. Indeed, the two most important subsidence basins, the Chenchou-Gabes corridor and the Gulf of Gabes contain thick layers of Jurassic and Cretaceous deposits (Figure 5b) not affected by erosion. However, these structural movements have played a major role in the compartmentalization and karstification of some of the Jeffara aquifer systems.

In addition, the El Hamma Fault functions as a conduit between the “Continental Intercalary” and Jeffara aquifer systems, allowing water to emerge naturally along the faults. This phenomenon is a direct consequence of the ability of the fault network to promote the discharge of these waters from the western Lower Cretaceous “Continental Intercalary” aquifer (covering the entire western domains of the Fejej, Dhahar and Oriental Erg) to the eastern Upper Cretaceous “Jeffara” aquifer (within El Hamma and Gabes).

#### 4. Local scale modeling

Regional mapping provided a numeric and reasonable foundation for understanding the spatial extent, the stratigraphic pattern and the structural factors delimiting the neighboring multilayered aquifers. To achieve a meticulous evaluation of water resources within the coastal aquifer systems, we constructed an architectural model of the Jeffara of Medenine area (Figure 1b). Extending the database by geophysical

investigations allowed a geostatistical mapping of all geological features of interest at the required level.

##### 4.1. Methodology

An architectural model of the Jeffara of Medenine area was constructed based on a sequential estimation of the Cretaceous sequences’ boundaries using well logs, core and seismic data. The local modeling workflow involves four essential processing steps: seismic interpretations, fault network construction, building the horizons models in time domain, and depth conversion using an external drift kriging. The thicknesses of the different reservoir units were sequentially estimated by subtracting their respective kriged upper and lower limits.

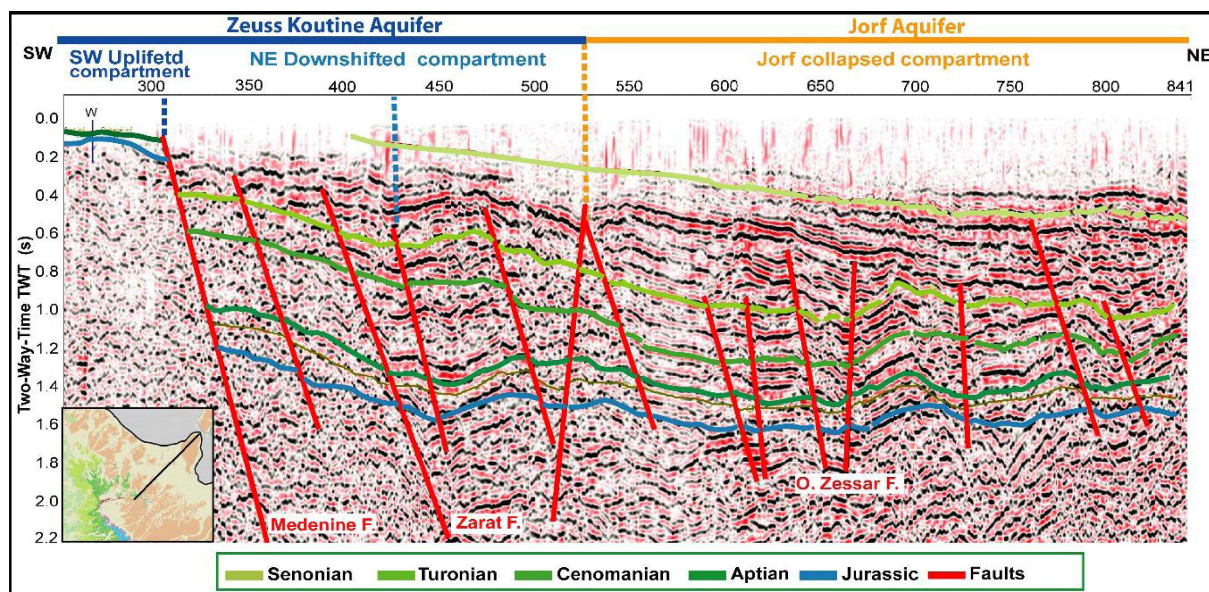
###### 4.1.1. Seismic investigation

The architectural modeling workflow involved eighteen time migrated 2D seismic lines, eleven well stratigraphic logs and associated petrophysical parameters (especially sonic log), and stacking velocities obtained from seismic lines processing, using SMT Kingdom software [Seismic Micro-Technology (SMT), 2015].

The harmonization process, which enables the different datasets to be processed so that they conform to each other in both geometric and semantic stratigraphic terms, is guided by the unified stratigraphic chart (Figure 1c) concordant with that adopted for large-scale characterization. The scheme defines the lithostratigraphic units and their bounding surfaces (reservoir unit tops or relative conformity/unconformity surfaces). This scheme constitutes the fundamental referential for interpreting seismic and well data, while also providing the details necessary for subsequent 3D modeling procedures.

Figure 6 shows an example of a seismic line crossing the study area with a NE–SW trend, it displays the interpreted seismic horizons which are the Jurassic, the Aptian, the Cenomanian, the Turonian and the Cretaceous tops. They are affected by a cluster of NW–SE oriented normal faults. The most significant are dipping northeast at approximately a 70° down-dip angle. The other faults dip to the southwest at a dip angle of about 60°. They define a series of Horst and Graben and tilted block structures.

These major faults divide the region into structural blocks, some of which define the limits of



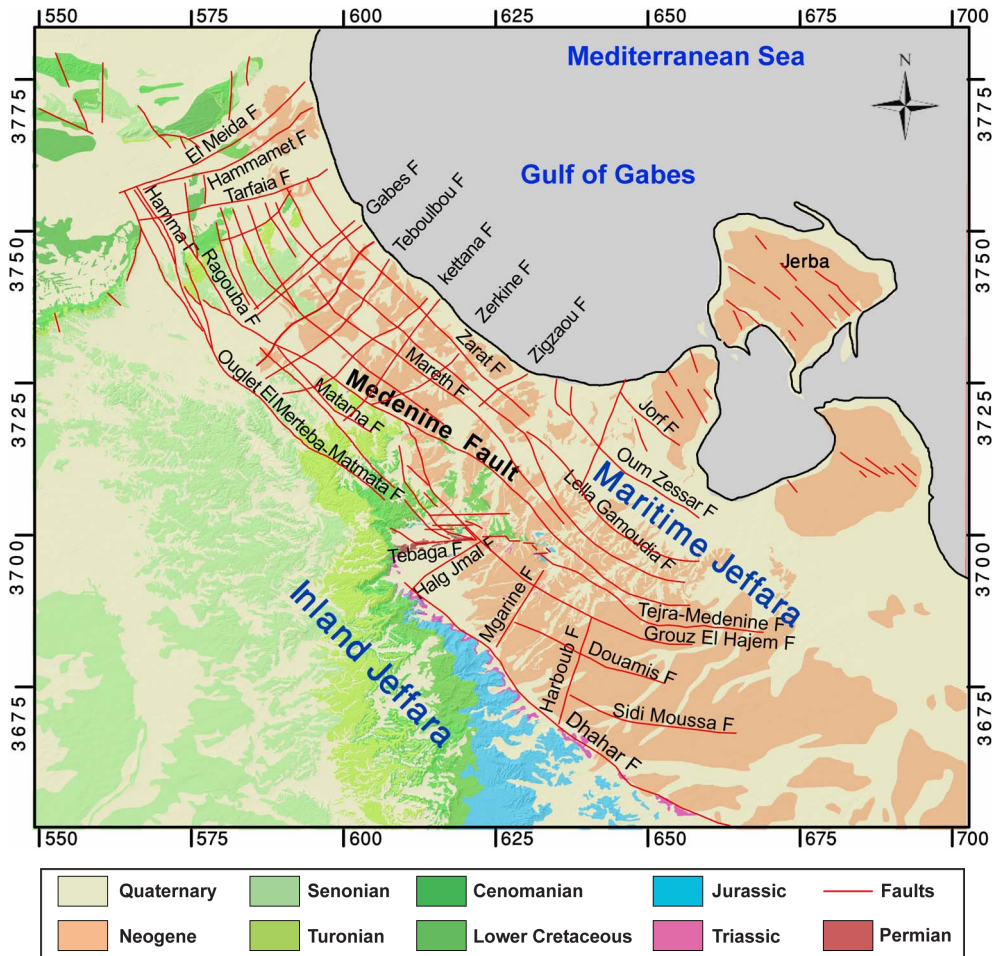
**Figure 6.** Example of a SW–NE oriented seismic line interpreted using SMT-Kingdom. It shows normal faulting, Horst and Graben, and tilted block structures. It also illustrates, as indicated on the top, the impact of the NW–SE normal faults on the geographical pattern and the extent of aquifers and the various compartments, defined in Chihi et al. [2015] and confirmed in this study, as the SW uplifted and the NE downshifted compartments of the Zeuss-Koutine aquifer, and the collapsed compartment of the Jorf aquifer.

aquifer systems (Figure 6). The included formations are affected by the repetitive faults' movement resulting in lateral extent and thickness variation.

#### 4.1.2. Building the fault network

The fault network strongly affects the arrangement, continuity and functioning of any such aquifers. It is, therefore, necessary to construct the fault network, to be integrated into the modeling process and to be used as a basis for understanding its impact on the various water units' functioning. Accordingly, faults interpreted from the seismic data and from the constructed regional geologic cross-sections were all plotted on the assembled geological map (Figure 1) using GIS tools. A thorough correlation analysis of all fault segments while taking into account locations of outcrops and their limits, bedding attitude, and fault geometric criteria (strike, throw magnitude, dip and dip direction), as well as topographic data and fieldwork validation. This analysis allowed to build the fault network, within the Jeffara basin, displayed on Figure 7.

The obtained structural map shows that the Jeffara basin is subdivided into two zones: the inland Jeffara represented by a series of continental terraces dipping gently westward and southward from the Medenine fault to the Dhahar foothills (Figures 1 and 5a–d). The maritime Jeffara extends along the coastal plain characterized by a thick sedimentary pile covered by Quaternary deposits. It is divided by a network of conjugate normal faults. The most prominent of these are the deep “*en echelon*” normal faults, which strike in the NW–SE direction and run parallel to the coast. These faults caused the sediments to shift downward towards the Gulf of Gabes. The most important is the Medenine fault, along which sediments may be displaced up to 1000 m. The NE–SW faulting system splits the study area into Horst and Graben blocks. The major aquifer systems, considered in this study, are situated along the Maritime Jeffara. The “Medenine Fault” divides all of them into two major domains: a northeast subsiding domain and a southwest high domain (Figures 6 and 7). The NE–SW faulting system splits the major domains into sub-compartments such as Teboulbou, Kettana,



**Figure 7.** Structural map of the study area, as specified in current structural studies, showing outcropping terrains and major normal faults mapped from interpreted seismic profiles, established geological sections and field studies.

Zerkine, Temoula, Ain Zrig, Koutine, Zeuss and Ksar Charif (Figure 2).

#### 4.1.3. Modeling of primary horizons in time domain

The interpreted fault pattern and structural horizons bounding the stratigraphic sequences are basically the main components to construct a structural model. In the first step, a horizon model is constructed in time domain based on (i) the picked points, from interpreted seismic profiles (Figure 6) referred to each geological unit top, and (ii) the spatial continuity analysis of the “elevation variable” as it is detailed below. For the sake of simplicity, we have chosen to describe the modeling of the Turonian reservoir, bounded by the Cenomanian on its

bottom and the Turonian on its top. Accordingly, in the following, we will detail the modeling procedure of the Cenomanian horizon through a time-to-depth conversion by means of kriging with external drift algorithm using the ISATIS geostatistical software. The thickness map of the Turonian reservoir was generated by subtracting the Lower Cenomanian and Upper Turonian kriged horizons.

The experimental variograms were calculated for the different “time” variables along the principal structural directions defined by the major normal faults trending northwest–southeast and northeast–southwest. The directional variograms of the Cenomanian horizon are displayed on Figure 8a, shown as an example. The variogram curves are con-

tinuously increasing, indicating a regional trend in both directions. They are also marked by a significant variability due to the structural faulting, inducing variable displacements of the considered surfaces. They also exhibit a slightly anisotropic behavior, which proves that the spatial variability of the Cenomanian horizon changes with directional orientation. The maximum variability is globally SW–NE. This indicates, in concrete terms, that the surfaces are tilted, and that the tilting is more accentuated towards the northeast rather than towards the southwest, as we would reasonably expect from seismic data and geological cross-sections.

The estimated fitted model for the “Cenomanian horizon” variable implies an anisotropic power function (Figure 8a):

$$\gamma(h) = 0.05 \text{ power } (15,000_{N45} + 17,500_{N135})^{1.6}.$$

Ordinary kriging estimation was performed on an estimation model grid covering 5330 elementary estimation cells of 500 m size. The kriging procedure was constrained by the structural complexity of the studied domain that has been taken into account, during the interpolation. In fact, the main faults affecting the bounding surfaces were used as barriers (Figures 7, 6 and 2) delimiting sub-areas of statistical consistency of the “elevation in time domain”. Furthermore, to deal with the uneven data distribution the neighborhood configuration was inferred by implementing a moving neighborhood using a large number (6) of angular sectors.

The surface estimation has given successful outcomes. In fact, the kriging standard deviation is low (0.00–0.01 s) evidently along the seismic profiles and around their intersections where the data density is the highest. Even, in between the seismic profile grid meshes the standard deviation values are ranging from 0.01 to 0.04 s. At locations where the seismic data are less dense, particularly in the southwest domain the estimation error increases up to 0.14 s. Figure 8b displays as an example the time depth map of the Cenomanian horizon.

#### 4.1.4. Modeling in “depth” domain

The cross-plot of elevation (m) versus seismic times TWT (s) for the Cenomanian horizon demonstrated a clear correlation between these variables at

the well locations. This would suggest that the seismic times might be used to help predict horizons’ elevations between wells. This time-to-depth conversion is possible through the *kriging with external drift* (KED) procedure, where the seismic time data guide the kriging process to introduce a component of its trend into the estimated depth. The seismic data are introduced as an additional constraint to optimize the faulted horizons.

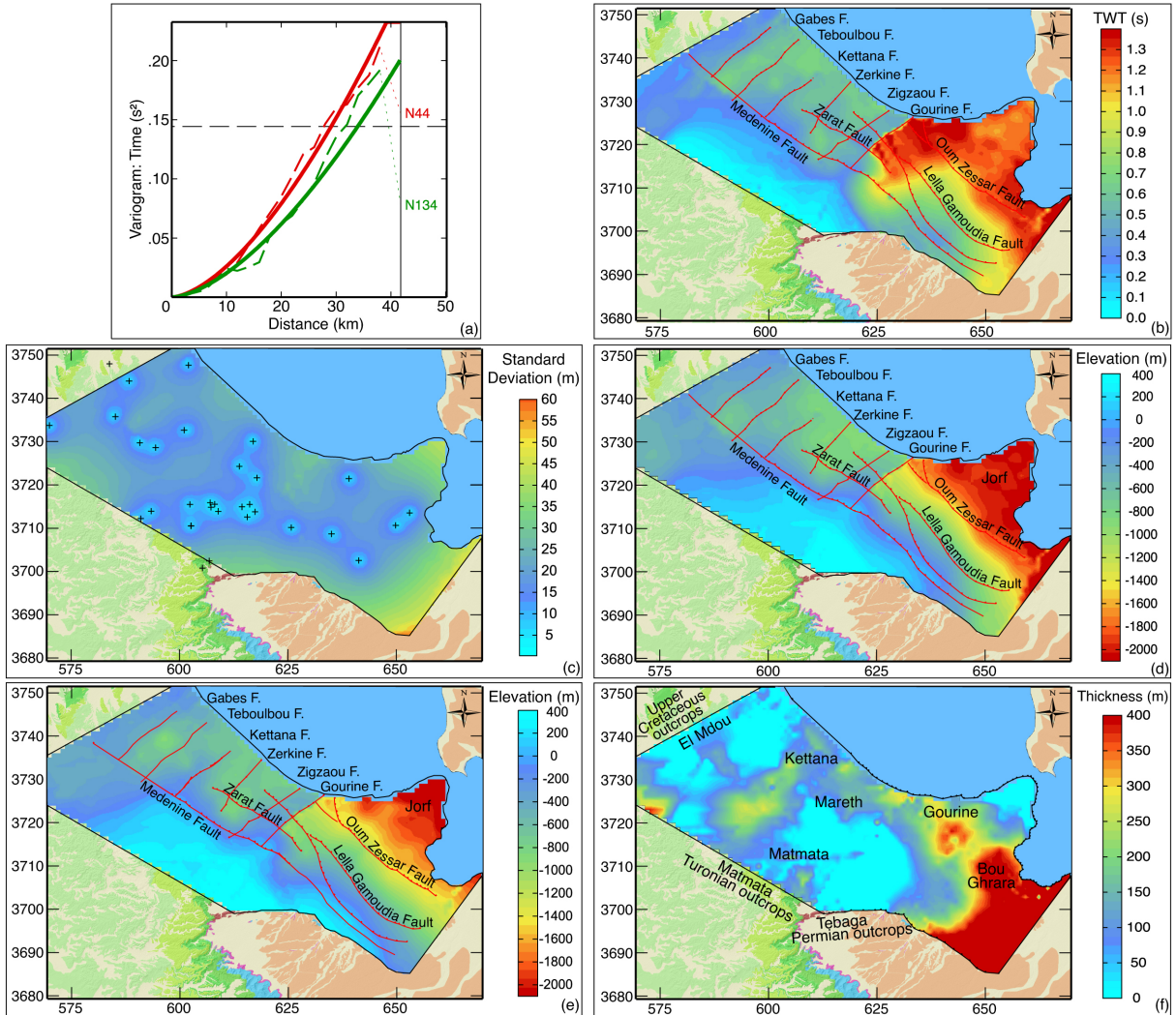
First, the model of the target variable (elevations of the Cenomanian measured at the wells) was defined, taking the seismic information into account as a drift, according to the theory of intrinsic random functions of order  $k$  (IRF-K) and using generalized polynomial covariances [Chilès and Delfiner, 2012, Chihi, 1997, Chihi *et al.*, 2000, 2007]. The estimated generalized covariance model is composed of two structures: a generalized covariance of order 1 and a linear drift part (1 fl). Then kriging was performed using elevation at wells as main data and the time map (Figure 8b) as an external drift. The quality of the estimation result was evaluated through the kriging standard deviation map (Figure 8c). The uncertainty is low near the wells, 0 to 5 m, and increases up to 12 m in between. Uncertainty rises significantly only along the eastern edges of the study area.

The geostatistical KED allowed, without integrating a velocity model, to generate the Cenomanian map (in depth) achieving a relatively small absolute error, honoring well data and reproducing the geological reality (Figure 8d). The same procedure was applied to calculate successfully the Turonian horizon (Figure 8e). Therefore, an isopach map of the Turonian reservoir was calculated by subtraction from the mapped Lower Cenomanian and the Upper Turonian boundaries (Figure 8f).

## 4.2. Visualization and interpretation

The geostatistical analysis suggests a complex spatial structure on the scale of the model grid (Figure 8). Globally, the study domain is characterized by two geological compartments that are separated by the Medenine fault (Figures 8 and 6).

Additionally, both the Cenomanian and the Turonian surfaces show general increasing elevations from southwest to northeast indicating that the whole area is downtilted towards the Mediterranean Sea (Figure 6). Maximum depth is observed along



**Figure 8.** Time-to-“depth” conversion using kriging with external drift: Directional experimental (---) and theoretical (—) variogram of the Cenomanian time horizon showing a drift component that indicates a notable increase in depth, revealing a downward tilting surface boundary (a). Time map of the Cenomanian horizon (b). Kriging Standard deviation map of the Cenomanian horizon indicating accurate estimation, the uncertainty is low close to the wells and in between, it rises only along the eastern edges of the study area (c). Elevation map of the Cenomanian horizon (d). Elevation map of the Turonian horizon (e), they reproduce the two regional compartments that are separated by the Medenine fault and display a general increase in elevation indicating that the region is down-tilted towards the Mediterranean Sea. The tilting is more pronounced in the southeastern sector, situated to the east of the Gourine fault. Thickness map of the Turonian reservoir, significantly affected by the crossing NW–SE and the NE–SW normal faults, the higher thickness values are observed on the southeast side within Gourine and Bou Ghrara regions. The reservoir is much thinner on the northwestern and southern sides, where it outcrops in the regions of El Mdou and Matmata.

the southwestern and northeastern sides within Jorf area and Oum Zessar Graben (Figure 8). A more

detailed trend analysis shows that the tilting is not uniform and presents different intensities. Indeed,

the tilting is higher in the southeast sector, located east of the Gourine fault, compared to that on the northwest sector, within Mareth and south Gabes regions (Figure 8b,d,e). Meanwhile, the calculated surfaces clearly depict, there, several Horst and Graben structures, generated by the normal NE–SW faults: Gourine, Zigzaou, Zerkine, Kettana and Teboulbou.

This dense fault network, dramatically, affects the reservoir thickness as it is illustrated in Figure 8f. The higher thickness values, greater than 400 m, are observed on the southeast side within Gourine and Bou Ghrara regions (Figures 8f and 2). Turonian thickness varies from 100 to 200 m in the regions of Kettana and Mareth. However, the reservoir is much thinner on the northwestern and southern sides, where it outcrops, respectively, in the regions of El Mdou and Matmata.

The stratigraphic/structural configuration, of the local studied domain, although generally consistent with that advanced in regional mapping (Figures 8d and 5c), shows important improvement to the mapping of lateral extent and regional/local dip of the reservoir formation. It particularly captures diverse, distinctive features that make it attractive from a hydrogeological point of view. In this regard, special attention was paid to the reservoir formations' extent and connection and more generally the compartmentalization and its impact on fluid circulation. This is described more in detail in Section 5.

## 5. Multi-scale modeling results—Aquifer systems characterization and management

The various findings and interpretations presented previously in this paper constituted the most imperative input involved in the implementation of a 3D geological model, and covering a wide domain in the Jeffara of Medenine area that includes four multilayered aquifer systems: Zeuss-Koutine, Sahel El Ababsa, Jorf and Mareth aquifers. This huge aquifer system will be designated as the “Jeffara of Medenine aquifer system” (JMAS). The outputs of the (3D) geological modeling were then used as a basis for the JAMS aquifer system management. This has been developed in three steps that will be clarified in the following sections: (i) description of the aquifer systems distribution, (ii) characterization of the reservoir connectivity within and between all aquifer compartments and (iii) development of a conceptual model

describing the groundwater flow pattern, and identifying potential zones for aquifer recharge.

### 5.1. 3D geological modeling

The 3D geological modeling was conducted, using GeoModeller [Intrepid Geophysics GeoModeller 3D modeling, 2020], to characterize the multi-layer reservoir systems involving the Triassic-Quaternary intervals and to enhance the understanding of their functioning at regional and local scales. It is based on a huge database (Section 3.1.1, Table 1), namely: (i) the digitized geological maps, (ii) a Digital Elevation Model (DEM) with 1 arc-second (approximately 30 m) resolution, derived from Satellite Mission – SRTM (Shuttle Radar Topography Mission), available on the USGS database ([earthexplorer.usgs.gov/](http://earthexplorer.usgs.gov/)); (iii) structural data related to the geological interfaces deduced from time-to-depth conversion and the fault network (Section 4); (iv) boreholes' data and wells' logs; (v) interpreted seismic profiles and geological cross-sections; and (vi) geological interpretations and assumptions.

The process of computing consists in a specific 3D interpolation method based on the potential field theory, that defines a geological interface as an implicit surface, namely a particular isosurface of a scalar field defined in the 3D space [Lajaunie *et al.*, 1997, Mcinerney and Guillen, 2005]. It is implemented through a co-kriging geostatistical technique involving two main variables: (i) the “geological contact” defining the upper boundary elevation, of a given geological sequence, treated as an iso-potential surface, and (ii) the “orientation data” (strike, dip and facing) that constitute the gradients of the potential field. The interpolation was constrained by some geological knowledge and hypothesis that are derived from the combined geological–geophysical investigations (presented above in Sections 3 and 4). The first and most important geological constraints are those concerning the stratigraphic layering pattern, defining the chronological order of the geological sequence, to define correctly the “geological contact” variables. Furthermore, we have to specify the complex relationships (erosive and/or onlap) between geological series, for this, multiple potential fields can be combined in the same model [Chilès *et al.*, 2004, Calcagno *et al.*, 2008, Caumon *et al.*, 2016]. The second geological constraint concerns

the faults affecting the continuity of the studied sequences [Chihi *et al.*, 2015, Hammami *et al.*, 2018a,b]. Specifically, we defined the links (i) between geological formations and faults and (ii) between the faults themselves. Intersection relationships between fault pairs are identified by examining which faults stop on other faults. This assumption allows an age relationship to be considered since the terminal fault is assumed to be the oldest. Similarly, the estimation of the fault surfaces is based on the interpolation of the potential field by considering the fault traces as data points and the fault orientation data as gradients. The geological boundaries and the fault surfaces were drawn as iso-values of the interpolated scalar field [Lajaunie *et al.*, 1997, Caumon *et al.*, 2016]. The 3D model of the JMAs, encompassing the four multilayered aquifer systems, covers an area of 2275 km<sup>2</sup> (65 km × 35 km) and extends down to a depth of 1.8 km. The Digital Elevation Model (DEM) was input to confer the terrain upper surface of the model.

The constraints defined above make it possible to ensure the reliability of the model, and deal effectively with major uncertainties [de Marsily, 1994, Wellmann and Caumon, 2018]. The main challenges, for the present study, are defined by (i) the complexity of the structural configuration defined by several geological faulted and downtilted layers showing lateral thickness variation and (ii) the dataset, although considered large, has an uneven distribution. The uncertainty is taken into consideration through (i) the data coherence, (ii) the robustness of the interpolation procedure; and (iii) the iterative computing/reevaluation process.

The position and coherence of *the different datasets* were thoroughly checked during the different previous (regional and local) modeling processes. However, their integration as input data in the GeoModeller software allowed through the 3D visualization a careful examination of the coherence and the relative position of the surface topography, the geological outcrops settings, boreholes and petroleum wells, and also the seismic profiles and the digitized geological cross-sections. This was performed to guarantee the correctness of the data within and between different datasets and to avoid inconsistencies in the conceptual geological model and the data sources.

*The robustness of the interpolation procedures* is as

indicated above based on a cokriging method. The geological interfaces were estimated by cokriging involving the “contacts” and the “orientation” data, derived from the various datasets. These are guided by the imposed constraints that define the priority between the layers “erode” and “onlap”, as well as a stratigraphic column. All this would make it possible to reproduce the geometry and the relative 3D position of the geological interfaces.

On the other hand, *including faults* that were modeled by introducing discontinuous external drift functions in the cokriging system to calculate the potential field, allowed to reproduce the effect of faults on the geological interfaces. The management of the faults themselves constituted a key issue in the present case study. As presented above (Sections 3 and 4), the faults show a change in the structural orientation, from major faults striking southwest–northeast to minor northwest–southeast and also east–west faults all along the study area (Figure 7). They are also characterized by variable displacement and depth extent. Taking into account their paramount importance, a thorough analysis was systematically carried out for all major faults, in order to define the geometric parameters allowing to integrate them into the surface modeling process, while respecting the imposed constraints namely their position in relation to all affected surfaces and their relative relationships to neighboring faults.

Involving an *iterative computing/reevaluation process* allowed a continuous validation of the model. Parameters, including contact and orientation data and faults characteristics, were systematically varied or supplemented by integrating more geological cross-sections until a satisfactory model honoring the data and the geological reality was obtained. Practically, this was achieved by comparing visually the observed data with the modeled geological interfaces and faults while checking the structural configuration and the stratigraphic pattern within each compartment.

All these manual, visual, cognitive and technical processes helped to optimize the estimation procedure, reduce uncertainty and reproduce the geometry and the relative 3D position of the geological interfaces and the crossing faults.

The relevance of the 3D model was evaluated in terms of misfits that express quantitatively the match

between model outputs and the observed data. A model is considered adequate when the percentage of the total assessed mismatch is below a tolerance threshold defined according to the complexity of the geological environment.

The implicit 3D geological modeling approach that we performed allowed the geological model to properly match with the existing data. In fact, 70% of the total data, show a misfit lower than 3%, these data are in general located within the eastern side of the study area characterized by a higher data density. The remaining 30% indicate a misfit of about 5%, these data concern the western region.

The 3D geological modeling outputs were displayed in diverse, impressive representations illustrating three-dimensional geometries, subsurface stratigraphic patterns and internal lithological variations, in 3D visualizations, 2D map views and modeled cross-sections were generated [see Mezni *et al.*, 2022a, for more details]. Here, we present a south view 3D visualization of the model (Figure 9). Insightful descriptions were provided concerning (i) the reservoir/aquifer systems distribution, (ii) the reservoir formation geometry, the thickness variation and the extent according to the fault network, and (iii) reservoir connectivity within and between all aquifer compartments. This helped to develop a conceptual understanding of groundwater flow in aquifer systems. This will be detailed in the following sections.

### 5.2. Reservoir/aquifer systems distribution

The 3D regional geological model (Figure 9) reproduces the spatial arrangement of aquifers against each other and in conjunction with the major fault structures. At a lower scale, it accurately renders the layouts of geological bodies in space: i.e. stratigraphic series, fault patterns and the compartments' distribution within each aquifer system.

At the center lies the Zeuss-Koutine aquifer system, which is partitioned by the major Medenine fault into: (i) a southwest uplifted compartment, comprising the karstified limestones of the Jurassic and Albo-Aptian and (ii) a northeast collapsed compartment including the Turonian and Senonian carbonates, overlain by thick MPQ sediments (Figures 9 and 6). To the northeast of Zeuss-Koutine aquifer, we depict the Jorf aquifer containing the

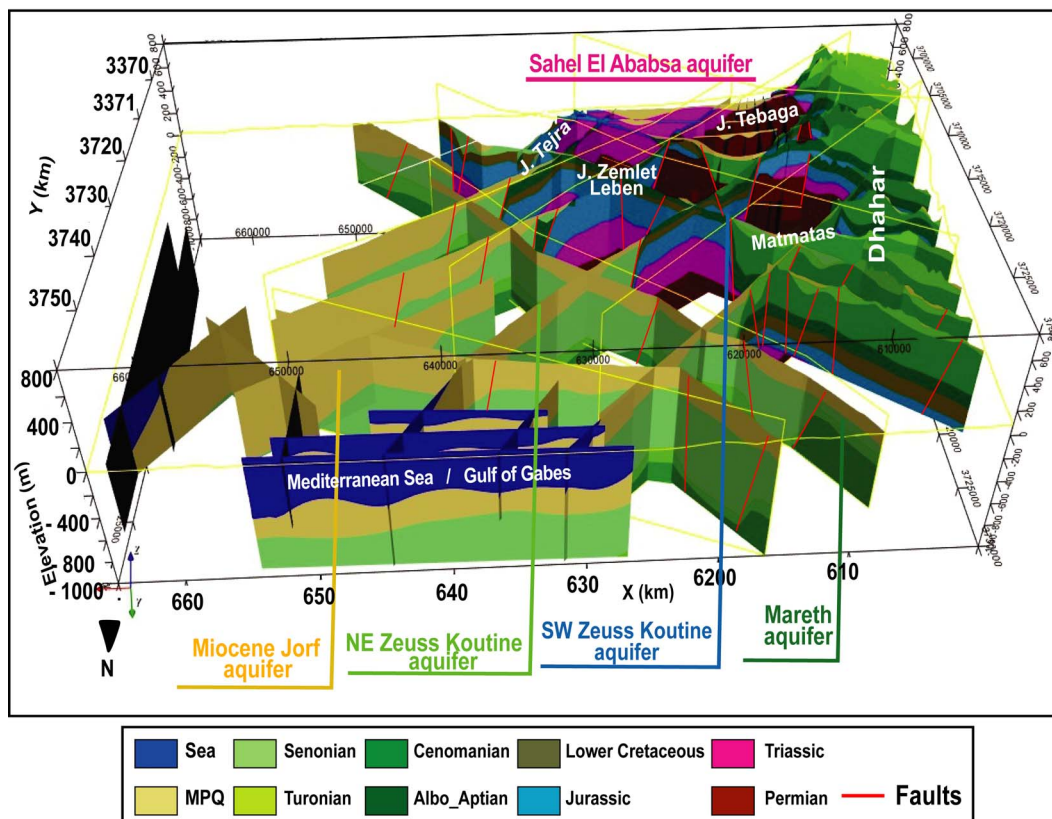
detrital sediments of the MPQ. It is delimited to the south by the Lella Gamoudia fault. To the south, the Triassic sandstone aquifer is extended along Sahel El Ababsa region. It is separated from the Dhahar outcrops by the Dhahar fault and is connected to the Zeuss-Koutine aquifer system by the Tebaga-Tejra faults. To the west, we recognize the southern Gabes region hosting Mareth aquifer, constituted of Lower Senonian and Turonian carbonates, and the Jebel Dhahar outcrops. In the southern domain, the Permian sedimentary unit hosting the reef reservoirs can be discerned, even on the surface.

### 5.3. Reservoir connectivity analysis within aquifer systems

Water flow paths are dependent on the continuity of reservoir formations and the degree of connection within and between all aquifer compartments [de Marsily *et al.*, 2005]. This is largely governed by the internal stratigraphic structural pattern of each compartment and on the hydrodynamic properties of its bounding faults, deduced from lateral facies variation on both sides of each fault, as well as the sources of recharge and outflow locations. Accordingly, a thorough analysis of reservoir connectivity was conducted along the Jeffara of Medenine area on the basis of (i) a piezometric map of the JMAS that was kriged taking into account the major faults (Figure 10), (ii) several restituted modeled cross-sections, (iii) a thorough correlation of the potentiometric gradient/lithologic properties along aquifer compartments and (iv) the hydraulic behavior of tectonic faults, such the major "Medenine fault", allowed to highlight the flow exchange between reservoir units and between major aquifers and to deduce a conceptual model of the water flow path.

Taking into account the regional groundwater flow within the JMAS (Figure 10), in general, from southwest to northeast, and the significant water level changes across the different compartments showing shorter and multidirectional flow paths, we revealed that the tectonic structures and the stratigraphic pattern govern largely the behavior of the flow units and the water flow path. First, we noticed a direct recharge of water from outcropping formations, mainly within Dhahar and Matmata Mountains. A high amount of water recharge is provided



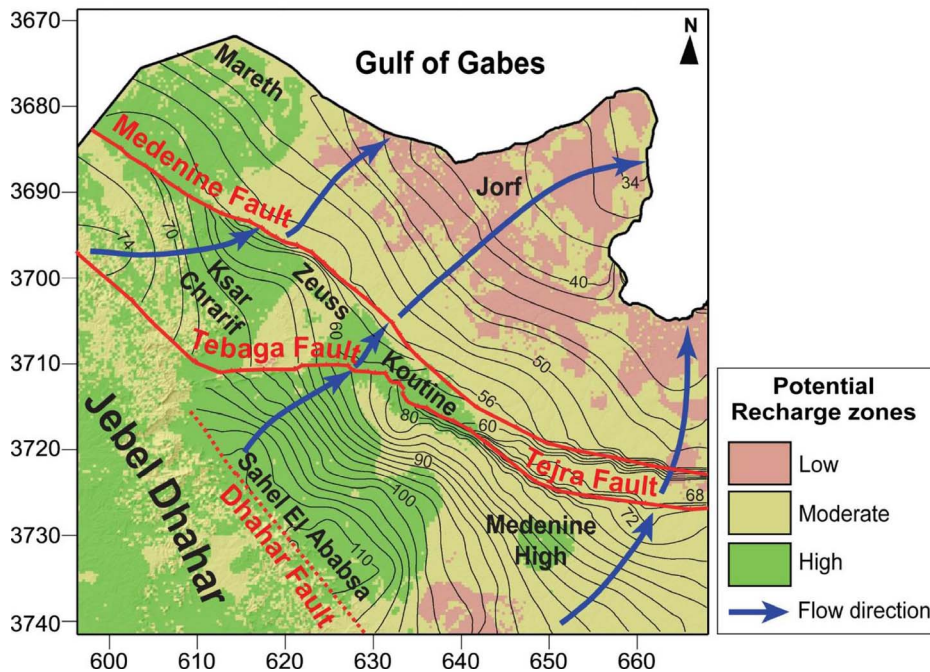


**Figure 9.** Visualization of the 3D geological model looking toward the south. It shows the uplifted series within the Dhahar and Matmatas mountains and their downshifting toward the Jeffara plain. It reproduces accurately the spatial arrangement of aquifers against each other and in conjunction with the major fault structures. Note as well how the collapsing of the Upper Cretaceous and the MPQ series is portrayed beneath the Mediterranean Sea.

through the uplifted karstified formations. The most important are those of Koutine and Ksar Charif compartments (Figures 9 and 10) that allow rapid percolation of water. In addition, groundwater in the aquifer system flows through a network of a steeply dipping normal NW–SE faults [Chihi et al., 2015]. The most important is the Medenine fault showing variable hydraulic behavior, according to the throw amount that increases from the southwest to the northeast. Actually, along its southwest segment, the fault juxtaposes the permeable Lower Cretaceous formations against the Upper Senonian impermeable formation, acting as an impermeable seal. In Zeuss area, the Medenine fault presents a conduit behavior. In fact, the throw amount is higher, involving the junction of two permeable carbonate units,

namely the Jurassic and the Lower Senonian against each other. More to the northwest, the same fault adjoins the Turonian and the Senonian carbonates, thus ensuring water movement.

On the other hand, NE–SW normal faulting developed Horst and Graben structural features. However, the throw amount is relatively less important, within the northeast and the southwest compartments (Figure 2), engendering the same water-bearing formations to be juxtaposed, i.e. Senonian vs. Senonian carbonates or Turonian vs. Turonian carbonates. Within the central compartment, the fault displacement is more important, allowing to place the Senonian and Turonian carbonates close to each other. In any case, the fault block displacements preserve the connectivity of the aquifer formations (Figure 2). Ad-



**Figure 10.** An overlay of the potential recharge zones map and the piezometric map within Jeffara of Medenine aquifer system. Potential recharge zones map shows that high recharge potentials cover 31% of the total area and are located within Koutine, Ksar Chararif Sahel El Ababsa, Medenine High, and Mareth. These locations are characterized by the highest infiltration rates enhanced by geologic factors, in particular the underlying thin vadose zone composed of permeable facies and in particular, karstified layers in Koutine and Ksar Chararif compartments.

ditionally, it is noteworthy that these faults alongside the *rivers* have played an important role in their *development* and their evolution. These rivers actually play a very important role in the replenishment of aquifer formations.

Subsequently, studying in the same way all the flow units behavior along the different aquifer compartments, we concluded that this architecture ensures a good hydraulic continuity all along the Jeffara of Medenine area. The water flow path could be drawn as following (Figures 10, 9 and 6) (i) from the southwest to the northeast compartments of the Zeuss-Koutine aquifer through Medenine fault, (ii) from the northeast Zeuss-Koutine compartment to the Jorf aquifer via the Lella Gamoudia fault, (iii) from the Sahel El Ababsa aquifer to the Zeuss-Koutine aquifer system via the Tebaga and Tejra faults, (iv) and from the Mareth aquifer to the Jorf aquifer by means of the Zerkine fault.

#### 5.4. Contribution to the aquifer system management

The results of the digital data processing within the geological model were also quite successful and opened new prospects for the production of thematic maps. Accordingly, we committed to develop a multidisciplinary methodology that enables to identify favorable areas for natural groundwater recharge (GWR) and to suggest suitable recharge structures [Mezni et al., 2022b]. A range of factors including, mainly, infiltration capacity and spatial and temporal distribution of rainfall and water runoff, govern GWR [Jaafarzadeh et al., 2021]. The infiltration capacity depends on the thickness and permeability of the soil and its evaluation is an important prerequisite, especially in case of lack or absence of measured data and, moreover, in a complex underlying geology. In order to overcome the lack of measured data related to the infiltration capacity, we conducted an ef-

ficient method based on a knowledge-driven modeling approach, using GIS, Remote Sensing, geostatistical interpolation and Analytical Hierarchy Process (AHP) to optimize the infiltration capacity distribution. This allowed the delineation of the potential GWR zones.

The originality of this approach [detailed in Mezni *et al.*, 2022b] compared with that proposed in the preceding studies lies mainly in the integration of the structural features characterizing the complex geological environment. These features are the aquifers architecture, the stacking pattern of the aquifer units, the structural compartmentalization, the lateral facies variation, the faults network, and fractures occurrence on geologic outcrops. We demonstrated that all these factors played a crucial role on the distribution of infiltration capacity and the waterways from the surface towards aquifer formations, via the vadose zone. This presents an asset to optimize the mapping of the potential aquifer recharge zones.

The multicriteria AHP-overlay method was applied, aided by ArcGIS to integrate various thematic layers comprising (i) topographic and hydrogeomorphological factors, including slope, drainage density, effective rainfall and land-use/land-cover maps; and (ii) geological and structural factors involving lineament density, soil, permeability distribution of the vadose zone and aquifer formations, and vadose zone thickness maps [Mezni *et al.*, 2022b]. All these factors were derived from all the previous work described above, primarily from (i) the combined geological–geophysical investigation, (ii) the geomodeling results concerning the spatial extent of geological formations, their connectivity in relation with the faults’ network and (iii) the potentiometric gradient and the deduced conceptual functioning of the aquifer systems in the whole region.

The generated map delineating the GWR potential zones (Figure 10) was evaluated by the Receiver Operating Characteristic (ROC) curve and the Area Under the ROC Curve (ROC AUC) score tools, and thereafter used to optimize the suitable sites for Managed Aquifer Recharge (MAR) facility installation. Accordingly, the GWR zones map shows that high recharge potentials cover 31% of the total area, they are located within (i) Koutine and Ksar Chararif compartments, (ii) Sahel El Ababsa area and Medenine High, and (iii) Mareth region. They are characterized by

the highest infiltration rates where water percolation is enhanced by many geologic factors, such as (i) the underlying thin vadose zone composed of permeable facies and (ii) in particular, karstified layers, as those of the Koutine and Ksar Chararif subcompartments, and more importantly with similar fractured multilayer aquifer systems. Therefore, these MAR structures built at different locations would promote sustainable groundwater supply and, crucially, allow groundwater withdrawals to be controlled. The overall results indicate that the proposed approach is reliable and can be applied within similar fractured multilayered aquifer systems. More importantly, it can be adapted to any semi-arid/arid environment.

## 6. Discussion

Multiscale modeling was carried out within the Tunisian Jeffara. This large study area provides a valuable example of the balance and interplay between structural complexity, input data, geostatistical modeling algorithms and procedures for dealing with uncertainty of model outputs. In the following, we summarize the major challenges and describe how we addressed these challenges while proposing solutions for future work and applications in other regional and/or local case studies. A crucial guiding point to consider is that a successful model depends on the density and, in particular, the availability of data according to the needed scale and the specificity of model purposes.

First, a *regional structural model* was constructed based on geological data, i.e. boreholes and well data, geological maps, geological cross-sections and field information.

The first challenge concerned *the uneven data distribution*, with a big difference in data density between the eastern and the western side of the study area (Figure 1). This prompted us to incorporate secondary information from outcropping sequences and geological cross-sections on the western side (Section 3), to control mapping results far from the wells.

Another challenge encountered was the *geological complexity* that was addressed through the adjusted theoretical variograms showing globally local stationary and drift anisotropic structures. These nested structures allowed to reproduce the complex

geometry of the sequences' boundaries while highlighting the local "elevation" variability within each structural bloc and the larger scale variability describing the downshifting of the surfaces from one bloc to another one.

The rate of the "elevation" variability was also handled through the adopted "*moving neighborhood selection*" procedure. In fact, we implement an octant search while optimizing the size and adjusting the relative spread within each angular sector. This ensured to sample the most appropriate points along all directions as consistently as possible, the closest ones to infer local variability and the farthest ones to respect the non-stationarity of the variable.

Based on the adjusted variogram and the neighborhood parameters, the kriging procedure provides the best estimates of the elevation variables, for each point, and an approximation of the uncertainty defined by the kriging standard deviation. The visualization and the analysis of the geological horizons and the standard deviation indicated a higher estimation error on the northwestern and southwestern sides of the study area, having fewer input data. Therefore, the kriging method can effectively indicate where additional measurements need to be made to reduce uncertainty and improve the estimation of surfaces. The needed additional information was generated from geological cross-sections, the only possible source for supplementary data in these regions.

Thereby, imposing such constraints within the modeling approach, we managed to produce highly realistic basal structures, even in areas of sparse data within the western side of the study area and complex structures along the eastern side. We emphasize that the northwest and southwest sides of the study area although having less input data, have much simpler geological structure, so it can therefore be inferred that the overall uncertainty is not necessarily significantly high and that the modeling results can be accepted.

We should note that completing the database with secondary data improves the results of the mapping away from the wells. However, they should not be integrated as hard data but as soft data generated through a reasonable approximation of the "elevation" values, taking into account the geological knowledge, the local variability and regional structural trends. There is another way for integrating

different sources of data to improve the mapping of the geological horizons. *Multivariate geostatistical* interpolation can also be used to supplement some measurements from the upper, generally more sampled, surfaces to better estimate the lower ones, where data is limited. Accordingly, it may be possible to improve the estimation of a given horizon using *cokriging* interpolation, which takes into account the correlation between successive horizons. This method is, however, complex and requires further work, as the relationships between successive geological horizons are significantly different across the geological compartments.

Second, a *local structural model* was built on the basis of the detailed seismic information with limited geological data, using geostatistical algorithms. We chose to apply the KED method for time-to-depth conversion and constructed accurate depth maps of the successive geological horizons. The generated maps are in accordance with the general expectation deduced from the regional geological model while exhibiting new detailed features that are captured and thus represented in the model.

Nevertheless, it should be noted that the time-to-depth conversion methodology, as used in a *sequential framework*, may eventually lead to a propagation of the errors from top to bottom. To address this uncertainty problem, a more efficient *global (joint) approach* can be applied. It consists in processing all horizons at once while optimizing the use of the available data, i.e. data from a given horizon are used in the estimation procedure of that horizon as well as the others. This procedure takes advantage of all available data, and thus reduces the uncertainty of the prediction and improves the horizons mapping while reproducing the geological reality. However, a clear correlation between the processed horizons must be evidenced. This is not the case in our current study. Such structural models could be used to simulate facies and petrophysical properties. This provides crucial 3D support for volume calculations, flow simulation and reservoir behavior. This can have a significant effect on consequent reservoir management.

*3D geological modeling* of the Jeffara of Medenine was based on 3D cokriging interpolation involving geological contact, and orientation data and taking into account faults as external drift. This advanced estimation method is also constrained, as

explained in Section 5, by (i) “erode and onlap” and (ii) the stratigraphic column that defines the chronological order of the interfaces. Inferring all these interpolation processes make it possible to solve all uncertainty issues and to avoid the crossing of the geological interfaces mentioned above (Section 3.2). The generated 3D geological model is coherent with the general geological knowledge, it reproduces the spatial distribution of the aquifer systems, the geometric features and the fault network, the thickness variability of the faulted geologic units, their lateral extent and their continuity to the land surface.

The outputs provide a sound basis for a better understanding of the geological configuration, the heterogeneity and connectivity of the reservoir units, in both local and regional scales. This effectively helped to identify the *regional pattern of groundwater flow* and the *exchange of water between the different aquifer compartments* and to propose rational solutions for *water recharge* for the planning of sustainable water resources management in the Jeffara. The model could serve as a support for hydrodynamic modeling, to investigate processes associated with inter-aquifer mixing, including groundwater contamination [Hammami *et al.*, 2021, 2022a,b, 2023].

The present work demonstrates the potential of such modeling multiscale approach to assist decision-makers in their exploration or research programs and resources management. Large regional and, more specifically, transboundary aquifer systems require a new way of thinking that can help to better understand the complexity of the subsurface structure and its associated uncertainties. This approach can be improved by implementing conditional simulation to generate multiple realizations of the subsurface, which can then be used to make more informed decisions about exploration and resources management.

## 7. Conclusions

In this work, we present original multiapproach methodologies to build regional and local architectural models in the extensive Jeffara basin, lying on the southern shore of the Mediterranean, using advanced geostatistical algorithms. *The workflow*

offers a level of approach combination and accuracy to reconstruct and interpret the regional/local stratigraphic/tectonic setting of the main sedimentary sequences belonging to the Cretaceous-Mio-Plio-Quaternary interval. The developed methodology allowed to implement efficient procedures to deal with two major constraints the data scarcity and their unequal distribution and the complex geological setting. It enabled integrating a large variety of multisource data, i.e., outcrops data, boreholes, oil wells, satellite images, geological and structural maps, lithostratigraphic cross-sections, seismic data and more importantly geological reasoning. Actually, this large database represents the most detailed and extensive sampling ever considered in such a modeling procedure in the region. On the other hand, the joint application of geological, hydrogeological and geostatistical methods shows that it is possible to significantly improve the integration and the reliability of existing and extended data and to enhance the characterization of the architecture and the functioning of aquifer systems at large and small scales.

At *regional scale*, we produced a new comprehensive map of the fault network that affects the extensive area of the Jeffara. In addition, seven major structural domains were distinguished, namely: Medenine-Dhahar, Oglet Mertba, Chotts Fejej-El Hamma, Oriental Erg, Jeffara plain, Gulf of Gabes and Mediterranean. Accordingly, a thorough correlation with the hydrostratigraphic framework allowed us to get a better understanding of the structural arrangement of the multilayered aquifer systems that extend in the region, i.e. the aquifers of Gabes, Zeuss-Koutine, Sahel El Ababsa and Jorf. We revealed that this specific layout of these aquifer systems was significantly conditioned by the tectonic events that occurred in the region.

At *local scale*, the integration of 2D seismic profiles and well-logging data allowed building a 3D geological model to highlight the existing relations between the coastal aquifer system units. The time-to-depth conversion, using kriging with external drift, revealed a complex structuring of the studied coastal aquifers with numerous NE–SW oriented Horsts and Grabens, combined with NE-dipping bedding planes triggered by major NW–SE faults. Downdip tilting prevails in the Jorf area, bringing the aquifer formations to greater depths. The extended data allowed to accurately reproduce the geological bodies

in space: i.e. stratigraphic series, fault patterns and the compartments' repartition within each aquifer system. The newly revealed fault system, at reservoir scale, contributed to a thorough re-evaluation of the structural/stratigraphic configuration, geometry and spatial extent of the subtle structural features, seals and reservoirs included within the coastal Jeffara of Medenine aquifer systems. A comprehensive correlation of the potentiometric gradient/lithologic properties along aquifer compartments and the hydraulic behavior of tectonic faults allowed to identify the flow exchange between reservoir units and between aquifer compartments. This brings out the preferential flow paths at large and small scales.

The geological model outputs constituted fundamental basis to conceive a reasoning approach based on AHP overlaying methods to accurately predict infiltration capacity, to regionalize potential GWR zones, and to propose adequate types and location for MAR systems. This constituted one of the major contributions of this study as it provided a sound basis for *planning sustainable water resources* in such arid environment where surface water supply is inadequate.

*Uncertainty* is an inherent aspect of geological modeling, and this study provides an illustration for each modeling process. Although we have integrated a vast amount of data, there are still limitations and uncertainties in some of the data sources, such as the quality of seismic data. The use of geostatistical algorithms fortunately allows uncertainty to be addressed, particularly in areas where data is scarce. Inferring the adequate interpolation processes and geological constraints makes it possible to solve the major uncertainty as mentioned above. We also recommended several solutions that could refine and improve the modeling processes and to further mitigate uncertainty, so as to increase the accuracy and realism of geological models. In this respect, and for example, the proposed geomodeling approaches could be improved by implementing conditional simulation to generate multiple realizations of the subsurface heterogeneities, which can then be used to make more informed decisions about exploration and resources management.

*As for future work*, there are several avenues to explore. One area of interest is the potential impact of climate change on the hydrogeological system of the Jeffara basin. Considering the region's arid environ-

ment and the significant importance of its aquifers, it is important to investigate how shifts in precipitation patterns and rising temperatures may affect water resources availability and quality in the future.

Another important direction for the application of such multiscale modeling is to support *decision-makers* in exploring and managing resources in regions with similar geological and hydrogeological settings, particularly in arid or semi-arid areas. This could help improve our understanding of aquifer systems and their behavior, which could ultimately lead to a more effective and sustainable management of water resources. It is also of major strategic importance to address large regional and *transboundary aquifer systems*. These are often shared by two or more countries and may trigger conflicting demands for water resources. By using such multiscale modeling in these areas, we can broaden the understanding of aquifer systems and their functioning across national boundaries, thereby helping to identify potential areas of concern and to develop strategies to prevent or resolve conflicts between different stakeholders. This can ultimately contribute to more equitable and sustainable water resource management practices.

## Conflicts of interest

Authors have no conflict of interest to declare.

## Acknowledgments

This work was part of the project "Geomodeling and Natural Resources Characterization", undertaken in the Georesources Laboratory of the Center for Water Research and Technologies and supported by the Ministry of Higher Education and Scientific Research. The authors would like to thank the Tunisian National Oil Company (ETAP) for providing seismic and petroleum well data and the General Directorate of Water Resources (DGRE) for providing borehole data.

## References

- ArcGIS (2012). Release 10.1, Environmental Systems Research Institute (ESRI), Redlands, CA. <https://www.arcgis.com/index.html>.

- Ben Baccar, B. (1982). *Contribution à l'étude hydrogéologique de l'aquifère multicouche de Gabès sud*. PhD thesis, University of Paris-Sud, Orsay, France.
- Ben Ferjani, A., Buroillet, P. F., and Mejri, F. (1990). *Petroleum Geology of Tunisia*. Tunis, Memoir No 22. Tunisian National Oil Company (ETAP).
- Ben Youssef, M. and Peybernes, B. (1986). Données micropaléontologiques et biostratigraphiques nouvelles sur le Crétacé inférieur marin du Sud-Tunisien. *J. Afr. Earth Sci.*, 5(3), 217–231.
- Bishop, W. F. (1975). Geology of Tunisia and adjacent parts of Algeria and Libya. *Assoc. Pet. Geol. Bull.*, 59(3), 413–450.
- Bodin, S., Petitpierre, L., Wood, J., Elkanouni, I., and Redfern, J. (2010). Timing of early to mid-Cretaceous tectonic phases along North Africa: New insights from the Jeffara escarpment (Libya–Tunisia). *J. Afr. Earth Sci.*, 58(3), 489–506.
- Bouaziz, S., Barrier, E., Turki, M. M., Tricart, M., and Angelier, J. (2002). Tectonic evolution of the Southern Tethyan margin in Southern Tunisia. *Tectonophysics*, 357, 227–253.
- Busson, G. (1967). *Le Mésozoïque saharien 1ère partie : l'extrême-Sud tunisien*. Service Géologique, 8. CNRS, France.
- Calcagno, P., Chilès, J. P., Courrioux, G., and Guillen, A. (2008). Geological modeling from field data and geological knowledge Part I. Modelling method coupling 3D potential-field interpolation and geological rules. *Phys. Earth Planet. Inter.*, 171(1–4), 147–157.
- Caumon, G., Jessell, M., de Kemp, E., Nemeth, B., Peron, G., and Schetselaar, E. (2016). Introduction to special section: Building complex and realistic geological models from sparse data. *Interpretation*, 4(3), 1A–Y1.
- Chihi, H. (1997). *Modélisation 3-D des unités stratigraphiques et simulation des faciès sismiques dans la marge du golfe du Lion*. PhD thesis, Sorbonne University, UPMC-Paris VI, France.
- Chihi, H. (2021). Building 3D Geomodels for Water Resources Management. In *The 5th International Conference on "Natural Resources and Sustainable Environmental Management"*, November 8–12, 2021, Turkey.
- Chihi, H., Alain, G., Ravenne, C., Tesson, M., and de Marsily, G. (2000). Estimating the depth of stratigraphic units from marine seismic profiles using non-stationary geostatistics. *Nat. Resour. Res.*, 9(1), 77–95.
- Chihi, H., Bedir, M., and Belayouni, H. (2013). Variogram identification aided by a structural framework for improved geometric modeling of faulted reservoirs: Jeffara basin, Southeastern Tunisia. *Nat. Resour. Res.*, 22(2), 139–161.
- Chihi, H. and de Marsily, G. (2009). Simulating Non-stationary seismic facies distribution in a prograding shelf environment. *Oil Gas Sci. Technol. – Rev. d'IFP Energies Nouv.*, 64(4), 451–467.
- Chihi, H., de Marsily, G., Belayouni, H., and Yahyaoui, H. (2015). Relationship between tectonic structures and hydrogeochemical compartmentalization in aquifers: Example of the “Jeffara of Medenine” system, south–east Tunisia. *J. Hydrol. Reg. Stud.*, 4(part B), 410–430.
- Chihi, H., de Marsily, G., Bourges, M., and Sbeaa, M. (2016). A constrained geostatistical approach for efficient multilevel aquifer system characterization. *J. Water Resour. Hydraul. Eng.*, 5(3), 80–95.
- Chihi, H., Jeannée, N., Yahayoui, H., Belayouni, H., and Bedir, M. (2014). Geostatistical optimization of water reservoir characterization case of the Jeffara de Medenine aquifer system (SE Tunisia). *Desalin. Water Treat.*, 52(10–12), 2009–2016.
- Chihi, H., Tesson, M., Alain, G., de Marsily, G., and Ravenne, C. (2007). Geostatistical modelling (3D) of the stratigraphic unit surfaces of the Gulf of Lion western margin (Mediterranean Sea) based on seismic profiles. *Bull. Soc. Géol. Fr.*, 178(1), 25–38.
- Chilès, J. P., Aug, C., Guillen, A., and Lees, T. (2004). Modeling the geometry of geological units and its uncertainty in 3D from structural data: the potential-field method. In *Proceedings of "Orebody Modeling and Strategic Mine Planning"*, Perth, WA, pages 22–24.
- Chilès, J. P. and Delfiner, D. (2012). In *Geostatistics: Modelling Spatial Uncertainty*. John Wiley & Son Interscience Publication, New-York.
- de Marsily, G. (1994). Quelques réflexions sur l'utilisation des modèles en hydrologie/On the use of models in hydrology. *Rev. Des Sci. l'eau/J. Water Sci.*, 7(3), 219–234.
- de Marsily, G., Combes, P., and Goblet, P. (1992). Comment on 'Ground-water models cannot be validated', by L.F. Konikow and J.D. Bredehoeft. *Adv. Water Resour.*, 15(6), 367–369.

- de Marsily, G., Delay, F., Gonçalves, J., Renard, P., Teles, V., and Violette, S. (2005). Dealing with spatial heterogeneity. *Hydrogeol. J.*, 13(1), 161–183.
- Gabtni, H., Alyahyaoui, S., Jallouli, C., Hasni, W., and Mickus, K. L. (2012). Gravity and seismic reflection imaging of a deep aquifer in an arid region: case history from the Jeffara basin, southeastern Tunisia. *J. Afr. Earth Sci.*, 66, 85–97.
- Hammami, M. A., Brahem, S., Chihi, H., Mezni, I., Baba Sy, M. O., and Cau, P. (2022a). Web-based deployment of a hydrogeological database using the observations data model 2 (ODM2). In *Soil and Water Resources Management for Combating Desertification in Drylands under Climate Change (SWDCC 2022), June 14–17, 2022, Djerba Island, Tunisia*.
- Hammami, M. A., Chihi, H., Ben Mammou, A., and Yahyaoui, H. (2018a). Aquifer structure identification through geostatistical integration of geological parameters: case of the Triassic sandstone aquifer system (SE Tunisia). *Arab. J. Geosci.*, 11(11), article no. 248.
- Hammami, M. A., Chihi, H., Brahem, S., and Baba Sy, M. O. (2023). Characterization of the hydrogeological interfaces of Gabes aquifer system: An integrated geological modeling approach. In *The 4th Atlas Georesources International Congress: Geoscience Innovations For Resource Management: Socio-Economic Challenges In An Environmentally Constrained World, March 18–20, 2023, Hammamet, Tunisia*.
- Hammami, M. A., Chihi, H., and de Marsily, G. (2018b). Building constrained (3D) geostatistical models case of the triassic sandstone aquifer system (SE Tunisia). In Kallel, A., Ksibi, M., Ben Dhia, H., and Khélifi, N., editors, *Recent Advances in Environmental Science from the Euro-Mediterranean and Surrounding Regions*, Proceedings of “Euro-Mediterranean Conference for Environmental Integration (EMCEI-1)”, Tunisia. Springer, Cham.
- Hammami, M. A., Chihi, H., Mezni, I., and Baba Sy, M. O. (2021). Hydrodynamic modeling of a multilayer aquifer system: Effect of the geological architecture. Case of the Gabes aquifer system, SE Tunisia. In *The 5th International Conference on “Natural Resources and Sustainable Environmental Management, NSREM-2021”, November 8–12, 2021, Turkey*.
- Hammami, M. A., Chihi, H., Mezni, I., and Baba Sy, M. O. (2022b). Impact of complex geological setting on the Hydrodynamic modeling Case of the Gabes aquifer system, SE Tunisia. In *The 21st Annual Conference of the International Association for Mathematical Geosciences, August 29 to September 03, 2022, Nancy, France*.
- Intrepid Geophysics GeoModeller 3D modeling (2020). Commercial platform developed by the BRGM (French Geological Survey) France, <https://www.brgm.fr/fr/logiciel/geomodeller-logiciel-modelisation-geologique-3d-sous-sol> and Intrepid Geophysics Melbourne, Australia, <https://www.intrepid-geophysics.com/product/geomodeller/>.
- ISATIS (2020). Geovariances, France, <https://www.geovariances.com/en/isatis/>.
- Jaafarzadeh, M. S., Tahmasebipour, N., Haghizadeh, A., Pourghasemi, H. R., and Rouhani, H. (2021). Groundwater recharge potential zonation using an ensemble of machine learning and bivariate statistical models. *Sci. Rep.*, 11(1), article no. 5587.
- Jabir, A., Cerepi, A., Loisy, C., and Rubino, J.-L. (2020). Stratigraphy, sedimentology and paleogeography of a Paleozoic succession, Ghadames and Jefarah basin, Libya and Tunisia. *J. Afr. Earth Sci.*, 163, article no. 103642.
- Klett, T. R. (2001). Total petroleum systems of the Pelagian Province, Tunisia, Libya, Italy, and Malta, the Bou Dabbous–Tertiary and Jurassic–Cretaceous Composite. U.S. Geological Survey Bulletin, 2202-D, version 1.0.
- Kraouia, S., Saidi, M., Soussi, M., El Asmi, K., and Mabrouk El Asmi, A. (2022). Petroleum basin modelling of the Middle Jurassic series in the Chotts Basin-Dahar areas (Southern Tunisia): Burial History, Thermal Maturity and timing of generation. In *Proceedings of “AAPG-EAGE Medina Mediterranean and North African Conference and Exhibition”, Tunisia*.
- Lajaunie, C., Courrioux, G., and Manual, L. (1997). Foliation fields and 3D cartography in geology: principles of a method based on potential interpolation. *Math. Geol.*, 29, 571–584.
- Letouzey, J. and Trémolières, P. (1980). Paleostress fields around the Mediterranean since the Mesozoic derived from microtectonics: Comparisons with plate tectonic data. In *Géologie des chaînes alpines issues de la Téthys. 26ème C.G.I., Paris. Mém. Bur. Rec. Géol. Miniè.*, 115, 261–273.



- Linde, N., Renard, P., Mukerji, T., and Caers, J. (2015). Geological realism in hydrogeological and geophysical inverse modeling: A review. *Adv. Water Resour.*, 86(Part A), 86–101.
- Mammou, A. (1990). *Caractéristiques et évaluation des ressources en eau du Sud Tunisien*. PhD thesis, University of Paris-Sud, Orsay, France.
- Mariethoz, G. and Lefebvre, S. (2014). Bridges between multiplepoint geostatistics and texture synthesis: Review and guidelines for future research. *Comput. Geosci.*, 66, 60–80.
- Mcinerney, P. and Guillen, A. (2005). Building 3D geological models directly from the data? A new approach applied to Broken Hill, Australia. In *Proceedings of Digital Mapping Techniques, Baton Rouge, Louisiana*, volume 05, pages 119–130.
- Mezni, I., Chihi, H., Bounasri, M., Ben Salem, A., and Ayfer, S. (2022a). Combined geophysical-geological investigation for 3D geological modeling: case of the jeffara reservoir systems, medenine basin, SE Tunisia. *Nat. Resour. Res.*, 31, 1329–1350.
- Mezni, I., Chihi, H., Hammami, M. A., Gabtni, H., and Baba Sy, B. (2022b). Regionalization of natural recharge zones using analytical hierarchy process in an arid hydrologic basin: a contribution for managed aquifer recharge. *Nat. Resour. Res.*, 3, 867–895.
- Peybernes, B., Kamoun, F., Ben Youssef, M., Trigui, A., Ghanmi, M., Zarbout, M., and Frechengues, M. (1993). Sequence stratigraphy and micropaleontology of the Triassic series from the southern part of Tunisia. *J. Afr. Earth Sci.*, 17(3), 293–305.
- Ravenne, C. (2002a). Stratigraphy and oil: a review. Part 1: Exploration and seismic stratigraphy: observation and description. *Oil Gas Sci. Technol. - Rev. d'IFP Energies Nouv.*, 57(3), 211–250.
- Ravenne, C. (2002b). Stratigraphy and oil: a review. Part 2: characterization of reservoirs and sequence stratigraphy: quantification and modeling. *Oil Gas Sci. Technol. - Rev. d'IFP Energies Nouv.*, 57(4), 311–340.
- Seismic Micro-Technology (SMT) (2015). KINGDOM software, Houston, United States, <http://www.seismicmicro.com>.
- Soua, M. and Chihi, H. (2014). Optimizing exploration procedure using oceanic anoxic events as new tool for hydrocarbon strategy in Tunisia. In Gaci, S. and Hachay, O., editors, *Advances in Data, Methods, Models and Their Applications in Oil/Gas Exploration*, pages 25–89. Cambridge Scholars Publishing (CSP), New York.
- Wellmann, F. and Caumon, C. (2018). 3-D Structural geological models: Concepts, methods, and uncertainties. Cedric Schmelzbach. *Adv. Geophys.*, 59, 1–121.





Research article

Geo-hydrological Data & Models

# Cokriging of transmissivity from head measurements revisited in the case of two quasi steady state flows

Chantal de Fouquet<sup>Ⓢ,\*,a</sup> and Jérémy Nos<sup>Ⓢ,b</sup>

<sup>a</sup> Mines Paris, PSL, Centre de géosciences, 77300 Fontainebleau, France

<sup>b</sup> Mines Paris, PSL, France

*E-mails:* chantal.de\_fouquet@minesparis.psl.eu (C. de Fouquet),  
jeremy.nos@mines-paris.org (J. Nos)

**Abstract.** The theoretical relationships between simple and cross-covariances of the transmissivity and the hydraulic head, deduced from the linearisation of the flow equation are used for cokriging transmissivity field with data from one then two quasi steady-state flows. First a non-exhaustive historical review of geostatistical modelling based on flow equations in a simplified case is given. Then original results show that the estimation of the logtransmissivity by cokriging with head data from two flows is improved compared with the case of head data from a single flow. The estimation in the presence of a transmissivity barrier is thoroughly investigated. Practical conclusions are drawn and perspectives based on the recent numerical variogram are proposed.

**Keywords.** Inverse problem, Flow equation, Steady state, Cokriging, Geostatistics.

**Funding.** The presented study was funded by Schlumberger Water Service, France.

*Manuscript received 26 April 2023, revised 18 October 2023, accepted 24 November 2023.*

## 1. Introduction

In *four decades on inverse problem in hydrogeology*, de Marsily et al. [2000, see also de Marsily et al., 1999] observe that “in hydrogeology, there are a great number of competing methods for solving the inverse problem”. Kitanidis [1995] notes: “although the consensus is that the linear method is appropriate only when the variance of log conductivity is less than 1, in my experience the method may work reliably for larger variances provided that the variation happens to be gradual or in the direction perpendicular to the streamlines”. However, referring to the comparison of methods for solving the inverse problem carried out by several teams on four synthetic data sets,

Zimmerman et al. [1998] point out “the evident superiority of the nonlinear methods as compared with linear ones”. In a more recent review of the inverse problem, Illman [2014] presents the evolution of approaches for mapping heterogeneity.

In this paper, we come back to the early cokriging of the transmissivity from transmissivity and head measurements, based on the linearisation of the flow equation linking head and transmissivity fluctuations. First a non-exhaustive historical review of geostatistical modelling based on the linearisation of flow equations is given. The relationships between simple and cross-covariances of transmissivity and head fluctuations for two quasi-steady state flows established by Zhang and Neuman [1996] in another context are applied to the inverse problem. On synthetic cases original results show the interest of crossing two quasi-steady-state flows to improve the estimation of transmissivity. The restrictive hy-

\* Corresponding author.

potheses used for the linearisation of the flow equation are then discussed, as well as the practical application of cokriging of log-transmissivity.

The present paper does not give an extensive review of “geostatistical inversion”, even for methods based on the linearisation of the flow equation. For a substantiated analysis of the difficulties posed by the spatial variability or the heterogeneity of subsurface reservoir properties for flow and transport modelling, the reader is invited to refer to de Marsily et al. [2005]. Neuman’s feedback [2020] also states the importance of spatial variability of parameters. A recent non-exhaustive review of methods for constraining geostatistical estimation by flow physics can be found among others in de Fouquet et al. [2023].

## 2. Steady state linear flow

### 2.1. Linearisation of the flow equations

For steady state flow without recharge, the diffusivity equations in two dimensions is written

$$\operatorname{div}(T \cdot \operatorname{Grad} H) = 0 \quad (1)$$

where  $T$  represents the transmissivity and  $H$  the hydraulic head.

Delhomme and de Marsily [2005] observe that Equation (1) can be written

$$\Delta H = -\operatorname{Grad}(\ln T) \cdot \operatorname{Grad} H \quad (2)$$

making “it clear that the gradient of  $\ln T$  was intrinsically involved in the flow equation, regardless of whether  $T$  is lognormally distributed or not”.

Let now  $x = (x_1, x_2)$  denote a point in the plane. In the simplified case with a macroscopic head oriented in the  $\vec{Ox}_1$  direction with constant modulus  $J$ , and constant scalar transmissivity  $T(x) = T_0$ , the hydraulic head  $h$  is linear:

$$h(x) = h_0 - Jx_1$$

where  $h_0$  is an arbitrary constant (since the hydraulic head is a potential).

For a transmissivity field with spatial variability, for example when the transmissivity is lognormal  $T(x) = T_0 e^{\Theta(x)}$ , let us denote  $\Phi$  the associated head perturbation:  $H(x) = h(x) + \Phi(x)$ .

When only the first order perturbations are retained, Equation (1) becomes:

$$\Delta \Phi = J \partial \Theta / \partial x_1. \quad (3)$$

The hypothesis that the transmissivity is lognormal is here not necessary, Equation (3) being based on the first order expansion of the transmissivity in Equation (1).

The initial models of Mizell et al. [1982] and Kitanidis and Vomvoris [1983] linking the covariances of the perturbations  $\Theta$  and  $\Phi$  were based on stationarity assumptions. Dagan [1985] calculated a second-order approximation of the covariance of the hydraulic head, observing that “the first-order approximation is very robust and even for a log-conductivity variance equal to unity, the second-order correction of the head variances is smaller than 10% of the first-order approximation”.

Dong [1990] wrote the solution of Equation (3) as follows, since transmissivity cannot generally be assumed to be differentiable:

$$J\Theta = \Delta Z \quad \text{and} \quad \Phi = \partial Z / \partial x_1. \quad (4)$$

This equation shows that head and transmissivity do not have the same degree of stationarity. Indeed, Matheron [1973] showed that, if  $\Theta$  is a stationary random function of order two, the solution  $Z$  of the Poisson equation

$$J\Theta = \Delta Z$$

is an Intrinsic Random Function of order 1 (IRF-1), twice differentiable. More generally, if  $\Theta$  is an IRF- $k$ ,  $Z$  is an IRF- $(k+2)$ .

Matheron also showed that if  $Z$  is an IRF-1, the solution  $\Phi$  of the partial differential equation

$$\Phi = \frac{\partial Z}{\partial x_1}$$

is an anisotropic intrinsic random function (IRF-0). According to these results, if the transmissivity is stationary of order two (i.e. mean and spatial covariance are stationary), the head perturbation  $\Phi$  is an anisotropic IRF-0, and the head is an IRF-0 with a linear drift (following the first coordinate in this particular case).

The generalised simple and cross covariances between transmissivity and head perturbations which are solution of Equation (4) verify the following relationships, with  $h = (h_1, h_2)$ :

$$J^2 C_\Theta = \Delta^2 C_Z; \quad C_\Phi = -\frac{\partial^2}{\partial h_1^2} C_Z; \quad J C_{\Theta\Phi} = \Delta \frac{\partial}{\partial h_1} C_Z. \quad (5)$$

The simple (generalised) covariances are even. The cross-covariance is odd, and thus equal to zero for

$h = 0$ . The anisotropic variogram of the head perturbation is generally unbounded.

Dong [1989, 1990] and Roth [1995] give the explicit expressions of the generalised covariance of head perturbations as well as the cross-covariance between transmissivity and head perturbations, for usual various models of the variogram of transmissivity. These expressions allow an estimate of the head consistent with Equation (4). Above all they allow estimating transmissivity by co-kriging from head and transmissivity measurements, providing a linearised approximation of the solution of the inverse problem.

In this model, boundary conditions (fixed head, zero flow) are rejected to infinity and occur via the macroscopic gradient of the head. In the cokriging these conditions are introduced as data only at the locations where they are known, and not necessarily all around the modelled field.

The literature presents several applications of this cokriging for solving the inverse problem. Rubin and Dagan [1987] consider that the spatial simple and cross covariances of transmissivity and head depend on a parameter vector  $\theta$  which cannot be determined with certainty from the data. They develop the analytical expressions for the conditional covariances when  $\theta$  is determined by a maximum likelihood procedure.

Ahmed and de Marsily [1993] perform the cokriging with an approximated but consistent model (ensuring the positivity of variances) of the simple and cross-covariances of head and transmissivity, calibrated to the sample simple and cross-covariances. On a synthetic case approaching a real one, they show that cokriging improves transmissivity estimation compared with kriging. Taking into account the local head gradient rather than the regional one improves the estimation of the transmissivity.

Hernandez et al. [2003] confirm that geostatistical inversion by co-kriging hydraulic head and log-conductivity improves predictions of head and flux compared with conditioning on conductivity or head data alone.

Moreover the expressions (5) of simple and cross covariances between transmissivity and head perturbations can be used for the direct geostatistical simulation of transmissivity, head and flow when linearisation assumptions are valid [de Fouquet, 2000, Mariotti et al., 2000].

## 2.2. Case of two flows

Zhang and Neuman [1996] observed that “even locally uniform gradients tend to show seasonal fluctuations in magnitude and direction”. An example is given in Verley et al. [2003] for the *nappe de la Beauce*, locally.

Zhang and Neuman examine the effects, on the longitudinal and transverse dispersion respectively, of these temporal fluctuations in the magnitude and direction of the mean velocity by developing the auto and cross-covariances of velocity, head and log conductivity. The previous first-order approach is generalised considering quasi steady-state flow at two times. These expressions of auto and cross covariances of head and log conductivity can also be used for geostatistical inversion by “crossing” two flows, in order to improve the estimation of transmissivity.

Let us revisit their developments. Two macroscopically linear flows, indexed by 1 and 2, are first considered in the same transmissivity field, following two orthogonal directions  $Ox_1$  and  $Ox_2$ . To simplify the presentation, the module  $J$  of the macroscopic gradient is assumed to be the same in the two cases. However this hypothesis can be relaxed. Based on the relationships (4), the solutions  $\Phi_i$  are written:

$$J\Theta = \Delta Z; \quad \Phi_i = \frac{\partial Z}{\partial x_i}. \quad (6)$$

The simple and cross covariances between transmissivity and head perturbations are deduced from (6)

$$\begin{aligned} J^2 C_{\Theta} &= \Delta^2 C_Z; \quad C_{\phi_i} = -\frac{\partial^2}{\partial h_i^2} C_Z; \\ J C_{\Theta\phi_i} &= \Delta \frac{\partial}{\partial h_i} C_Z \quad \text{and} \quad C_{\Phi_1\Phi_2} = -\frac{\partial^2}{\partial h_1 \partial h_2} C_Z. \end{aligned} \quad (7)$$

For a flow with the same macroscopic head gradient  $J$  and any angle  $\alpha$  with the first coordinate axe, the head perturbation is deduced from the linearity of Equation (6), putting

$$\Phi_{\alpha} = \phi_1 \cos \alpha + \phi_2 \sin \alpha. \quad (8)$$

The simple and cross covariances of transmissivity and head perturbations are derived from Equations (7).

However it is simpler to choose a base such that  $\overline{Ox_1} = \overline{Ox_{\alpha}}$ . Denoting  $\beta$  the angle between the two

macroscopic flows, the expressions of the simple and cross covariances are simplified:

$$\begin{aligned} C(\Phi_1, \Phi_\beta) &= \cos \beta C(\Phi_1, \Phi_1) + \sin \beta C(\Phi_1, \Phi_2); \\ C(\Theta, \Phi_\beta) &= \cos \beta C(\Theta, \Phi_1) + \sin \beta C(\Theta, \Phi_2) \\ \text{and } C(\Phi_\beta, \Phi_\beta) &= \cos^2 \beta C(\Phi_1, \Phi_1) \\ &\quad + 2 \cos \beta \sin \beta C(\Phi_1, \Phi_2) \\ &\quad + \sin^2 \beta C(\Phi_2, \Phi_2). \end{aligned} \quad (9)$$

Because of the linearisation, when the head data points are identical for the two flows (assumed not to be parallel) the cokriging results (transmissivity or head) do not depend on the angle  $\beta$  between the two flows.

Always because of the linearity of the equations, the head of a third flow is expressed as a linear combination of the first two. The co-kriging system with head data from three flows is therefore singular, unless measurement errors are introduced in the variogram of the head data, typically *via* nugget effects.

### 3. Example: comparison of cokriging the transmissivity from one or two flows

#### 3.1. Construction of synthetic cases

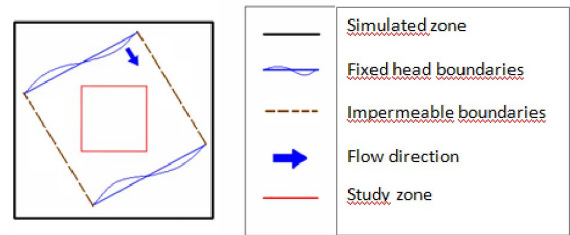
Cokriging tests are classically performed as follows: (i) geostatistical simulation of a transmissivity field; (ii) flow calculation on this field (using a flow simulator, here Visual Modflow), the fixed macroscopic gradient is imposed via conditions on the boundaries of a larger field; (iii) data extraction far enough from these boundaries; (iv) kriging and cokriging with data from one or two flows.

Several areas are specified (Figure 1):

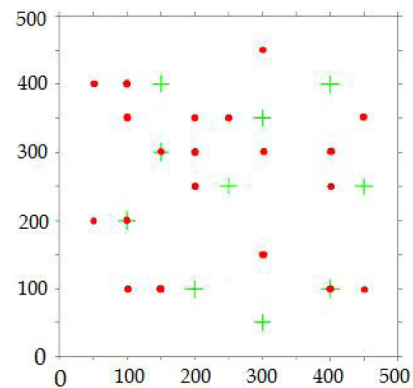
- the transmissivity is simulated on a “simulated area” of 1600 m × 1600 m;
- the “study area” is 500 m × 500 m from which the data are taken.

The first zone is large enough to fix the boundaries conditions of the various flows. These boundaries conditions define an “intermediate area” of 1100 m × 1100 m, in which the heads are calculated. However, in order for the head data to be far enough from the boundaries, the “study area” is smaller than the “intermediate area” (Figure 1).

The log-transmissivity is simulated with a spherical variogram with a range equal to 70 m and a sill of 0.5. Three flows with a macroscopic gradient with unit module are calculated



**Figure 1.** Definition of the three areas: simulated area for transmissivity simulation, intermediate area for setting boundary conditions, and study area. The arrow indicates the main flow direction for the represented boundary conditions.



**Figure 2.** Location of the transmissivity (+) and head (•) data.

- North to South (NS)
- Northwest to Southeast (NW–SE)
- West to East (W–E)

Initial tests have shown that isolated head data reduce only slightly the estimation variance of transmissivity, while in the vicinity of two close head data, this estimation variance decreases significantly. This is consistent with the known observation that “head variations due to heterogeneity are small, whereas those of velocities and travel time are large” [Mathéron and de Marsily, 1980].

The co-kriging by two flows is therefore examined with a significant number of close head data (20), cf. Figure 2.

The different estimations of transmissivity are compared according to the following four criteria,

where  $N$  denotes the number of meshes

- the mean absolute error MAE:

$$(1/N) \sum_{i=1}^N |(\log(T_i))^* - \log(T_i)|$$

- the mean absolute standardised error MASE:

$$(1/N) \sum_{i=1}^N (|(\log(T_i))^* - \log(T_i)| / \sigma_i^K)$$

- the root-mean-square error RMSE:

$$\sqrt{(1/N) \sum_{i=1}^N ((\log(T_i))^* - \log(T_i))^2}$$

- and the root-mean-square standardised error RMSSE:

$$\sqrt{(1/N) \sum_{i=1}^N (((\log(T_i))^* - \log(T_i))^2 / (\sigma_i^K)^2)}$$

where  $(\log(T))^*$  denotes the estimate of  $\log(T)$  and  $\sigma^K$  the standard deviation of the estimation (kriging or cokriging) error. The RMSSE criterion based on the standardised error allow verifying the expected improvement of accuracy by cokriging.

### 3.2. Comparison between one and two flows

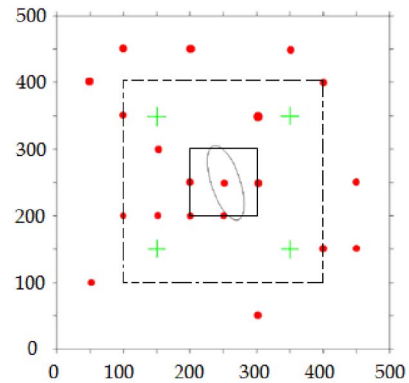
The values in Tables 1 and 2 are given on the decimal logarithm.

The mean deviation MAE and the RMSE show that cokriging improves the estimation accuracy of transmissivity compared with kriging, and even more when crossing two flows. However, the standardised deviations are rather stable between kriging and cokriging with one flow, and slightly increase for two-flows cokriging. The RMSSE are close to unity, which shows that the uncertainty is correctly quantified, and somewhat overestimated by kriging or cokriging with one flow.

### 3.3. Detection of a transmissivity barrier

The second test field presents a low transmissivity barrier of  $10^{-6} \text{ m}^2 \cdot \text{s}^{-1}$  compared to a previous average value of  $10^{-4} \text{ m}^2 \cdot \text{s}^{-1}$ . The variogram model of transmissivity remains unchanged, because no transmissivity data is located in the barrier.

The test is performed with only four transmissivity data. The quality criteria are first calculated on a small area of  $100 \text{ m} \times 100 \text{ m}$ , centred on the anomaly



**Figure 3.** Location of the heterogeneity, represented by an ellipse. Central square: local area; dashed square: extended area.

(Figure 3), in order to verify if cokriging helps detect the heterogeneity.

The MAE and the RMSE (Table 2a) show that the N–S or NW–SE flows, which are yet little impacted by the transmissivity barrier, improve the estimation. Logically the W–E flow, which crosses the barrier, informs the best on the transmissivity field: the mean deviation MAE and the RMSE decrease by almost 60%. Note that, with a mean deviation of 0.347, the combination of the N–S and NW–SE flows gives an estimation accuracy comparable to that of the W–E flow, with a mean deviation of 0.365. The lowest errors (MAE, RMSE) are obtained by cokriging using the N–S and W–E flows, which both cross the barrier.

The standardised deviations are large: the method allows detecting the heterogeneity, but with only four transmissivity data located outside it, the estimation accuracy remains poor. However, the increase in the variance of the transmissivity field was not taken into account in the modelling.

These results are consistent with Kitanidis remark [1995] cited in the introduction according to which “the method may work reliably [...] provided that the variation happens [...] in the direction perpendicular to the streamlines”.

The results are less satisfactory on a larger scale. In the extended area of  $300 \text{ m} \times 300 \text{ m}$ , bounded by dashes lines on the figure, the head data improve or on the contrary deteriorate the MAE, depending on the configurations, but in all cases the RMSE de-

**Table 1.** Estimation of log-transmissivity on the 500 m × 500 m study area (2601 meshes)

		MAE	MASE	RMSE	RMSSE
Kriging		0.238	0.760	0.306	0.971
Cokriging 1 flow	N-S	0.223	0.760	0.288	0.970
	NW-SE	0.220	0.754	0.286	0.967
	W-E	0.223	0.766	0.290	0.983
Cokriging 2 flows	N-S; NW-SE	0.218	0.805	0.285	1.034
	N-S; W-E	0.214	0.790	0.281	1.015
	NW-SE; W-E	0.218	0.804	0.284	1.028

Values are given on the decimal logarithm.

creases compared with kriging. The estimations are poorly improved by the head data of the N-S flow, which is visible on the estimation maps (Figure 4): the low transmissivity barrier is detected, but the surrounding area shows large transmissivities that do not exist. Once again, the head data from the W-E flow improve greatly the precision, and to a lesser extent, it is the same for the NW-SE flow. The MAE and RMSE criteria are minimal when crossing these two flows (Table 2b). In all cases, the RMSSE remains greater than one, for the same reason as on the local area.

The results remain similar on the study zone (Table 2c): with head data from the N-S flow, the MAE and the RMSE increase compared with kriging and they decrease sharply with head data from the W-E flow only. Once again, MAE and RMSE are minimal by crossing W-E and NW-SE flows. For all configurations, the RMSSE is lower than for the homologous case for the extended area.

However on these three heterogeneous areas, the available data do not allow estimating the transmissivity accurately.

### 3.4. Interpreting the results

The improvement of the precision of the transmissivity with head data from two flows can be explained by returning to the stream lines. It is known [Emsellem and de Marsily, 1971, de Marsily et al., 2000] that the solution of the inverse problem for the steady-state

flow equation, including a recharge or pumping, is written

$$\ln \frac{T(s)}{T_0} = - \int_{h_0}^{h(s)} \frac{\Delta h}{|\nabla h|^2} dh,$$

where the integration is made along a stream-tube,  $T_0$  is the transmissivity and  $h_0$  the hydraulic head at point  $s_0$  of the tube, and  $T(s)$  the transmissivity at point  $s$  where the head  $h(s)$  is measured.

The transmissivity at any point of this stream-tube can thus be deduced from the transmissivity at  $s_0$ . Formally, knowing the transmissivity along a line intersecting all the stream-tubes allows deducing the transmissivity everywhere in the field. In the case of two flows with nonparallel stream lines, knowing the transmissivity in a single point allows deducing the transmissivity everywhere in the field; see for example Castelner [1995].

In practice, the head is not known exhaustively but estimated, and the same for the stream-tubes. The estimate of transmissivity is therefore uncertain.

### 3.5. Practical conclusions from these synthetic cases

This synthetic example leads to practical conclusions on estimating transmissivity field:

- (i) sufficiently close head measurements allow estimating the head gradient, which is more useful for the accuracy than isolated head values;
- (ii) “crossing” two steady-state flows improves the accuracy of transmissivity estimation;



**Table 2.** Estimation of the log-transmissivity in the heterogeneous case, (a) in the local area (121 meshes), (b) on the extended area (961 meshes) and (c) in the study area

(a) Local area 100 m × 100 m		MAE	MASE	RMSE	RMSSE
Kriging		0.916	2.667	1.077	3.138
Cokriging 1 flow	N-S	0.703	2.434	0.870	3.013
	NW-SE	0.612	2.184	0.758	2.716
	W-E	0.365	1.281	0.450	1.576
Cokriging 2 flows	N-S; NW-SE	0.347	1.486	0.416	1.781
	N-S; W-E	0.305	1.296	0.374	1.579
	NW-SE; W-E	0.326	1.393	0.406	1.747
(b) Extended area 300 m × 300 m					
Kriging		0.300	0.905	0.462	1.372
Cokriging 1 flow	N-S	0.317	1.079	0.436	1.494
	NW-SE	0.289	0.986	0.393	1.358
	W-E	0.260	0.879	0.336	1.134
Cokriging 2 flows	N-S; NW-SE	0.319	1.211	0.398	1.513
	N-S; W-E	0.316	1.193	0.408	1.532
	NW-SE; W-E	0.247	0.938	0.321	1.225
(c) Study area 500 m × 500 m					
Kriging		0.274	0.809	0.379	1.114
Cokriging 1 flow	N-S	0.301	0.964	0.390	1.264
	NW-SE	0.270	0.871	0.351	1.149
	W-E	0.247	0.794	0.314	1.013
Cokriging 2 flows	N-S; NW-SE	0.317	1.114	0.400	1.417
	N-S; W-E	0.301	1.059	0.382	1.352
	NW-SE; W-E	0.239	0.840	0.305	1.078

Values correspond to decimal logarithm.

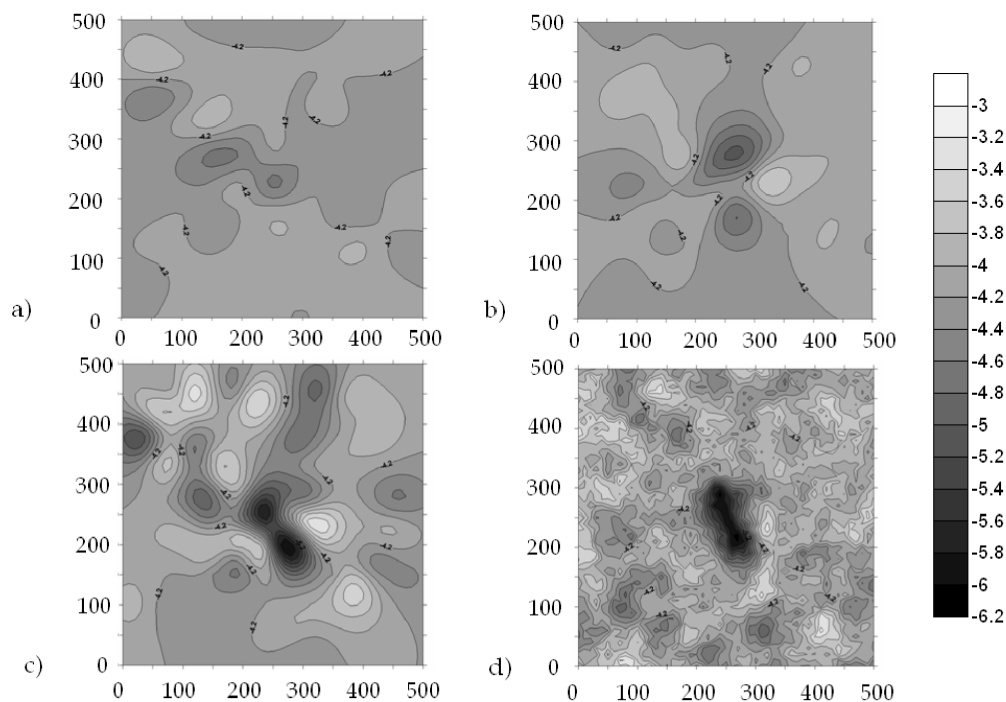
- (iii) this flow crossing is useful for detecting heterogeneities of a transmissivity field. However, transmissivity data have to be numerous enough for precisely delineating the heterogeneities;
- (iv) since the accuracy of the estimation depends on the main flow direction, the same applies to the “optimal” implantation of the data.

## 4. Conclusive remarks

### 4.1. Towards more complex flow conditions

In many cases, the hypothesis of a macroscopically linear two-dimensional flow is unsuitable.

Roth et al. [1998] developed a flexible and general approach in order to avoid excessive simplifications. The principle is to come back to conventional hydrogeological modelling, using a flow simulator. A set of geostatistical (non-conditional) simulated transmissivity fields is used as input of a flow simulator. The



**Figure 4.** Decimal logarithm of Transmissivity field. (a) Estimation from N–S flow; (b) estimation from NW–SE flow; (c) estimation from the two N–S/NW–SE flows; (d) “real” field.

set of associated output results (head, concentration) is used to calculate the explicitly nonstationary covariance  $C(x, x')$  or variogram  $\gamma(x, x')$  of the variable of interest. These nonstationary “numerical” simple and cross covariances or variograms of transmissivity and head are used to estimate by cokriging the transmissivity from transmissivity and head data. However an anamorphosis or at least a logarithmic transformation allows avoiding negative transmissivity estimates. Pannecoucke [2020] generalised this “numerical variogram” to transient flow, for mapping radionuclide activity.

To account for the non-linearity between concentration and hydraulic conductivity field, Schwede and Cirpka [2010] show that a Monte-Carlo approach is preferable for conditioning hydraulic conductivity fields by steady-state concentration measurements.

#### 4.2. Back transformation of estimated log-transmissivity

According to de Marsily et al. [2005] “One also needs to apply Kriging with some rigor on the log-transformed transmissivity values, in order to estimate geometric mean values and not arithmetic means. Back-transforming the kriged  $\ln T$  into  $T$  values must also be done correctly, i.e. simply as  $T = \exp(\ln T)$  without any additional factor using the variance of the estimation error to supposedly correct an assumed estimation of the median rather than of the mean, as is sometimes erroneously done”.

Another question is that of the support of the estimate: is kriging “punctual”, that is to say that the transmissivity has to be estimated on the same support as the transmissivity data, or is a larger support chosen, in which case an equivalent transmissivity is sought? However, except in the very special case of a stratified medium with a flow parallel to the stratification, the permeability is not additive. The equivalent block-permeability lies within the harmonic

and the arithmetic mean of sub-block permeabilities [Matheron, 1993, Delhomme and de Marsily, 2005]. In the special case of lognormal transmissivity with isotropic variogram, and with some additional hypotheses, Matheron [1967] has shown that the equivalent transmissivity is the geometric mean. Coming back to the estimated transmissivity simply consists thus to exponentiate the estimated log-transmissivity.

In practice, the regularisation similarly has to be carried out on the log-transmissivities, and the result is simply exponentiated. The bounds of the uncertainty interval can be calculated in the same way from the bounds of this interval on logarithmic estimates.

In a more thorough approach, tensor transmissivity fields should be estimated.

#### 4.3. *Biased flow predictions computed on estimated transmissivity fields*

However it is known that the (co)kriging “smoothes” map of the estimated values, whose spatial variability therefore differs from that of the real values. The non-linear calculations performed on an estimated transmissivity field (by (co)kriging or any other estimator, linear or not) is biased. A simulation based approach is then required.

The cokriging of log-transmissivity by log-transmissivity and head measurements however allows the direct conditioning of log-transmissivity simulations to these data. When the validity of the linearisation of the flow equation is not verified, especially when the head gradient cannot be assumed to be constant at the macroscopic scale, the simple and crossed “numerical variogram” between log-transmissivity and head (perturbations) provides a consistent bivariate model, which is compatible with the boundary conditions where these are known [Schwede and Cirpka, 2010, Pannecoucke et al., 2020]. However, the results shown here in the simplified case of the linearized flow equation remain useful in the general case: “crossing” two quasi steady-state flows (two seasonal situations for example) should improve characterising transmissivity heterogeneities.

#### Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

#### Acknowledgments

The authors warmly thank Jean-Pierre Delhomme for initiating and contributing to the study of flow crossing. They also thank Marco De Lucia for his initial contribution to review the theoretical part of this study and the reviewers for their comments that helped to better highlight the results.

#### References

- Ahmed, S. and de Marsily, G. (1993). Cokriged estimation of aquifer transmissivity as an indirect solution of the inverse problem: A practical approach. *Water Resour. Res.*, 29(2), 521–530.
- Castelier, E. (1995). *Estimation d'un champ de perméabilité à partir de mesures de charge hydraulique*. Doctorat en géostatistique, Ecole nationale supérieure des mines de Paris.
- Dagan, G. (1985). A note on higher-order corrections of the head covariances in steady aquifer flow. *Water Resour. Res.*, 21(4), 573–578.
- de Fouquet, C. (2000). Geostatistical simulation of transmissivity and flow: simplified case of the linearized diffusion equation in steady state. In Darsargues, A., editor, *Tracers and Modelling in Hydrogeology*, IAHS Publication 262, pages 103–108. IAHS, Wallingford, Oxfordshire, UK.
- de Fouquet, C., Le Coz, M., Freulon, X., and Pannecoucke, L. (2023). Making kriging consistent with flow equations: application of kriging with numerical covariances for estimating a contamination plume. *Hydrogeol. J.*, 31, 1491–1503.
- de Marsily, G., Delay, F., Gonçalves, J., Renard, Ph., Teles, V., and Violette, S. (2005). Dealing with spatial heterogeneity. Invited paper, special issue “The Future of Hydrogeology”. *Hydrogeol. J.*, 13, 161–183.
- de Marsily, G., Delhomme, J.-P., Coudrain-Ribstein, A., and Lavenue, A. M. (2000). Four decades of inverse problem in hydrogeology. In Zhang, D. and

- Winter, C. L., editors, *Modeling and Field Investigation in Hydrogeology: A Special Volume in Honor of Shlomo P. Neuman's 60th Birthday*, Special Paper 348, pages 1–17. Geological Society of America, Boulder, Colorado.
- de Marsily, G., Delhomme, J. P., Delay, F., and Buoro, A. (1999). Regards sur 40 ans de Problèmes Inverses en hydrogéologie. "Point Sur". *C. R. Acad. Sci. Paris*, 329, 73–87.
- Delhomme, J. P. and de Marsily, G. (2005). Flow in porous media: An attempt to outline Georges Matheron's contributions. In Bilodeau, M., Meyer, F., and Schmitt, M., editors, *Space, Structure and Randomness*, volume 183 of *Lecture Notes in Statistics*. Springer, New York.
- Dong, A. (1989). Kriging variables that satisfy the partial differential equation  $\Delta Z = Y$ . In Armstrong, M., editor, *Geostatistics. Vol 1. 3rd International Geostatistics Congress*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Dong, A. (1990). *Estimation géostatistique des phénomènes régis par des équations aux dérivées partielles*. Thèse de doctorat en géostatistique, Ecole Nationale Supérieure des mines de Paris.
- Emsellem, Y. and de Marsily, G. (1971). An automatic solution for the inverse problem. *Water Resour. Res.*, 7(5), 1264–1283.
- Hernandez, A., Neuman, S., Guadagnini, A., et al. (2003). Conditioning mean steady state flow on hydraulic head and conductivity through geostatistical inversion. *Stoch. Environ. Res. Risk Assess.*, 17, 329–338.
- Illman, W. A. (2014). Hydraulic tomography offers improved imaging of heterogeneity in fractured rocks. *Groundwater*, 52(5), 659–684.
- Kitanidis, P. K. (1995). Quasi-linear geostatistical theory for inversing. *Water Resour. Res.*, 31(10), 2411–2419.
- Kitanidis, P. K. and Vomvoris, E. G. (1983). A geostatistical approach to the inverse problem in groundwater modelling (steady state) and one-dimensional simulations. *Water Resour. Res.*, 9(3), 677–690.
- Mariotti, P., de Fouquet, C., and Bruno, R. (2000). Geostatistical characterization of an aquifer by joint simulation of a random function and its partial derivatives. In Kleingeld, W. and Krige, D. G., editors, *Geostats 2000 Cape Town*, volume 1, pages 29–38. Geostatistical Association of Southern Africa, Johannesburg.
- Matheron, G. (1967). *Éléments pour une théorie des milieux poreux*. Masson, Paris.
- Matheron, G. (1973). The intrinsic random functions and their applications. *Adv. Appl. Probab.*, 5(3), 439–468.
- Matheron, G. (1993). Quelques inégalités pour la perméabilité effective d'un milieu hétérogène. In de Fouquet, C., editor, *Cahiers de géostatistique, fascicule 3*, Compte-rendu des Journées de Géostatistique 1993. Ecole des Mines de Paris, France.
- Matheron, G. and de Marsily, G. (1980). Is transport in porous media always diffusive? A counterexample. *Water Resour. Res.*, 16(5), 901–917.
- Mizell, S. A., Gutjahr, A. L., and Gelhar, L. W. (1982). Stochastic analysis of spatial variability in two-dimensional steady groundwater flow assuming stationary and nonstationary heads. *Water Resour. Res.*, 18(4), 1053–1067.
- Neuman, S. P. (2020). Twenty lessons drawn from my subsurface hydrology career. *Perspect. Earth Space Sci.*, 2, article no. e2020CN000131.
- Pannecoucke, L. (2020). *Combinaison de la géostatistique et des simulations à base physique – application à la caractérisation de panaches de contaminants*. Thèse de doctorat, Mines ParisTech, PSL, <https://pastel.hal.science/tel-03135798>.
- Pannecoucke, L., Le Coz, M., Freulon, X., and de Fouquet, C. (2020). Combining geostatistics and simulations of flow and transport to characterize contamination within the unsaturated zone. *Sci. Total Environ.*, 699(2), article no. 134216.
- Roth, C. (1995). *Contribution de la géostatistique à la résolution du problème inverse en hydrogéologie*. Doctorat en géostatistique. Document du BRGM 241. ISBN 2-7159-0805-9.
- Roth, C., Chilès, J.-P., and de Fouquet, C. (1998). Combining geostatistics and flow simulators to identify transmissivity. *Adv. Water Resour.*, 21(7), 555–565.
- Rubin, Y. and Dagan, G. (1987). Stochastic identification of transmissivity and effective recharge in steady groundwater flow: 1. Theory. *Water Resour. Res.*, 23(7), 1185–1192.
- Schwede, R. L. and Cirpka, O. A. (2010). Interpolation of steady-state concentration data by inverse modeling. *Groundwater*, 48(4), 569–579.
- Verley, F., Brunson, F., Verjus, Ph., and Cholez, M.

- (2003). *Nappe de Beauce – Piézométrie hautes eaux 2002*. DIREN Centre et Ile-de-France. 53 p, 7 fig, 5 tab, 6 pl, 6 ann. ISBN N° 2-11-094172-3.
- Zhang, D. and Neuman, S. (1996). Head and velocity covariances under quasi-steady state flow and their effects on advective transport. *Water Resour. Res.*, 32, 77–84.
- Zimmerman, D. A., de Marsily, G., Gotway, C. A., Marietta, M. G., Axness, C. L., Beauheim, R. L., Bras, R. L., Carrera, J., Dagan, G., Davies, P. B., Gallegos, D. P., Galli, A., Gómez-Hernández, J., Grindrod, P., Gutjahr, A. L., Kitanidis, P. K., Lavenue, A. M., McLaughlin, D., Neuman, S. P., RamaRao, B. S., Ravenne, C., and Rubin, Y. (1998). A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow. *Water Resour. Res.*, 34(6), 1373–1413.





Research article

Geo-hydrological Data & Models

# Modelling the coupled heterogeneities of the lacustrine microbialite-bearing carbonate reservoir of the Yacoraite Formation (Salta, Argentina)

Vanessa Teles<sup>Ⓢ,\*</sup>, Youri Hamon<sup>Ⓢ</sup>, Rémy Deschamps<sup>Ⓢ</sup>, Sébastien Rohais<sup>Ⓢ</sup>, Fadi H. Nader<sup>Ⓢ</sup>, Elodie Heckenmeyer<sup>a</sup>, Marta Gasparrini<sup>Ⓢ</sup>, Mickael Barbier<sup>a</sup>, Olivier Lerat<sup>a</sup>, Philippe Joseph<sup>a</sup> and Brigitte Doligez<sup>a</sup>

<sup>a</sup> IFP Energies nouvelles, 1 et 4 avenue de Bois-Préau, 92852 Rueil-Malmaison, France

*Current addresses:* Beicip-Franlab, 232 Avenue Napoléon Bonaparte, 92502 Rueil-Malmaison Cedex, France (E. Heckenmeyer), Department of Earth Sciences, Università degli Studi di Milano, Via Mangiagalli 34, 20133 Milan, Italy (M. Gasparrini), Akkodis, 4 rue Jules Ferry, Pau, France (M. Barbier)

*E-mails:* vanessa.teles@ifpen.fr (V. Teles), youri.hamon@ifpen.fr (Y. Hamon), remy.deschamps@ifpen.fr (R. Deschamps), sebastien.rohais@ifpen.fr (S. Rohais), fadi-henri.nader@ifpen.fr (F. H. Nader), elodie.heckenmeyer@beicip.com (E. Heckenmeyer), marta.gasparrini@unimi.it (M. Gasparrini), mickael.barbier@akkodisgroup.com (M. Barbier), olivier.lerat@ifpen.fr (O. Lerat), philippe-joseph@club-internet.fr (P. Joseph), brigittedoligez@orange.fr (B. Doligez)

**Abstract.** The Yacoraite Formation is a complex lacustrine microbialite-bearing carbonate formation whose heterogeneity is due to different sedimentary facies associations and several microbialite geobodies. The bi-variate plurigaussian geostatistical method is applied to simulate in parallel these two types of heterogeneity in a 3D reservoir-scale geological model. Each variable is simulated by a complete plurigaussian method based on geological interpretation and field observations and quantifications. They are coupled through the occurrence of the microbialites within the sedimentary facies and associated into combined facies giving a one-glance representation of the reservoir heterogeneity, then used to constrain the porosity simulation.

**Keywords.** Modelling, Heterogeneity, Microbialite, Yacoraite Formation, bi PGS.

*Manuscript received 18 May 2022, revised 20 October 2022, accepted 22 November 2022.*

## 1. Introduction

*“If the world were homogeneous, i.e. if the rock properties were constant in space, and/or easy to determine,*

*hydrogeology would be a rather boring job: solving well-known equations in a perfectly identified medium. Fortunately, the world is heterogeneous, with highly nonconstant properties in space, and “dealing with heterogeneity” is what makes the work fascinating.”* [de Marsily et al., 2005]. These authors present the various numerical approaches developed through time to represent subsurface heterogeneity: from stochastic methods and geo-

\* Corresponding author.

statistics which suppose an underlying “structure”, Boolean methods based on specific pre-defined shapes to genetic methods now usually called forward or process-based models which model the physical processes leading to the subsurface architecture and heterogeneity. Michael *et al.* [2010] proposed to combine geological-process models and geostatistical models to condition to 3D subsurface structures. de Marsily *et al.* [2005] first question in their discussion is “*when does heterogeneity matter?*”. In the case of sedimentary rocks, their 3D heterogeneity is certainly governed by the sedimentary architecture and facies distribution as well as potential later diagenetic processes. The resulting petrophysical variability may depend on several superimposed structures and processes. This is particularly true in the case of carbonate rocks, where the fabric depends on the metabolism of the different competing organisms and local environmental factors or controls [Ahr, 2008]. Carbonate rocks are also prone to diverse diagenetic transformations through time. This paper focuses on a particular type of carbonate: microbial carbonates which are important components of carbonate sedimentary systems, both in modern and ancient times. Well-represented in marine and lacustrine settings, they are also found in fluvial, spring or cave environments [Riding, 2000]. The large variety of microbial communities and of mechanisms at play [Dupraz *et al.*, 2009, Riding, 2000, 2002, 2008, Suarez-Gonzalez *et al.*, 2019] may result in high spatial variability within facies belts and in different and complex geobodies such as mats, oncolites, stromatolites, thrombolites *etc.* These geobodies can vary in size from a few centimeters to large meter-scale structures and may contribute to the development of kilometre-scale formations that could be interpretable in seismic data [Wright, 2012, Wright and Barnett, 2015]. This highly variable heterogeneity scale (10’s–100’s of meters) and their spatial distribution make their impact on reservoir properties or fluid flow still a matter of debate. Moreover, it is important to keep in mind, as highlighted by Corbett *et al.* [2015], how the bio-architectural component, when controlling porosity in microbial carbonates, may represent a challenge in terms of sampling, measurement and modelling microbialite bearing carbonate reservoirs.

Several models focus on the growth of microbialite

considering the interactions of intrinsic and extrinsic factors whether at intermediate scales, namely bed- or build-up scales by combining Diffusion Limited Aggregation (DLA) and Cellular Automata (CA) [Dupraz *et al.*, 2006, Johnson and Grotzinger, 2006, Curtis *et al.*, 2021] or at larger scales combined with stratigraphic models [Gallois *et al.*, 2016, Kozłowski *et al.*, 2014, Xi *et al.*, 2022]. Several authors [Arslan *et al.*, 2008, Dirner and Steiner, 2015, Janson and Madriz, 2012, Jung *et al.*, 2012] propose to combine Multiple Points Simulation (MPS) with field observations and digital outcrop techniques to create 3D models.

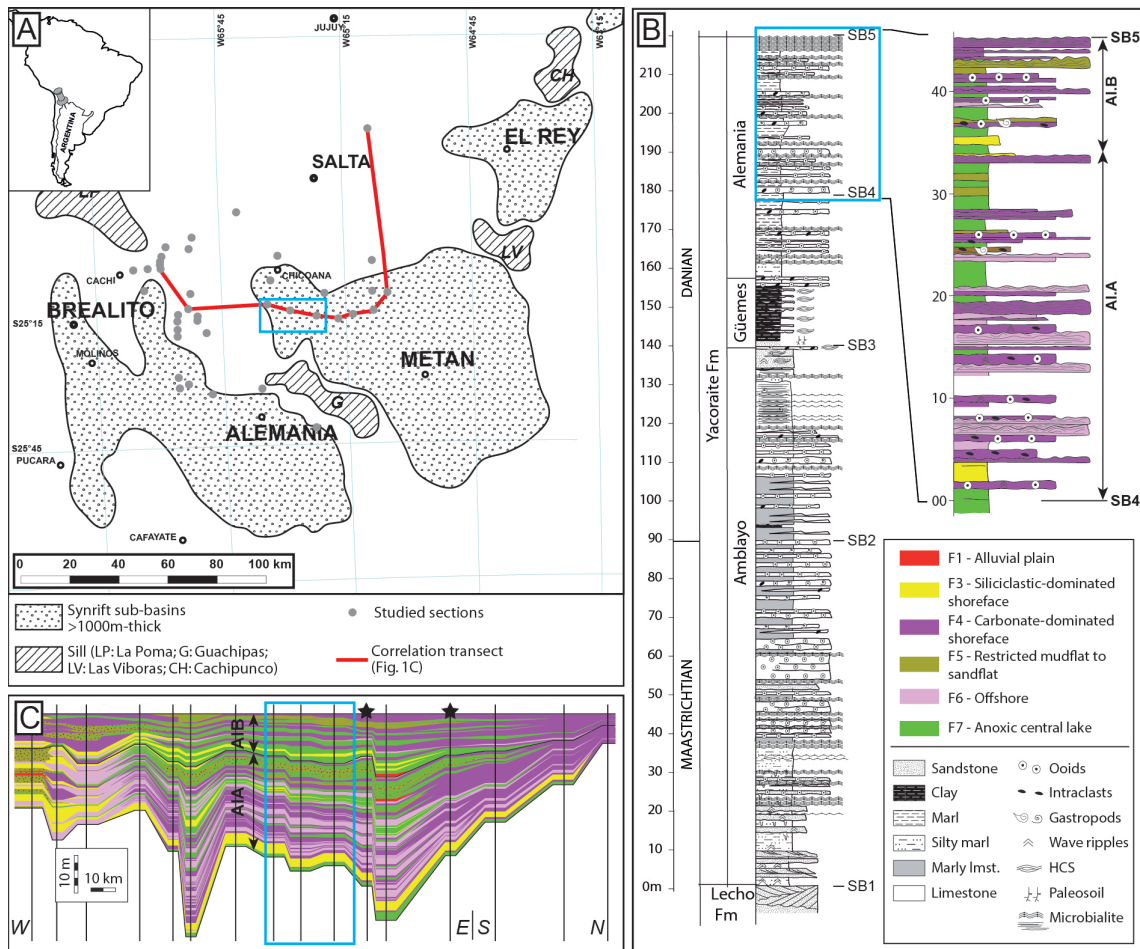
To date relatively few publications have focused on modelling strategies for such geobodies and associated deposits. The aim of this methodological paper is to present how the bi-variate plurigaussian method has been applied to the lacustrine Yacoraite Formation of the Salta rift basin (NW Argentina) whose heterogeneity depends on both sedimentary facies and microbialite geobodies. One of the main focuses is to illustrate how it is possible to integrate geological data and concepts in this modelling geostatistic method to reach a better description of the porosity field.

## 2. Geological context and general description of the sequence of interest

The Salta basin (NW Argentina; Figure 1A) belongs to the Andean Basin system evolving from a rift to a foreland setting [Marquillas *et al.*, 2005, Viramonte *et al.*, 1999]. The sedimentary infill of the Salta Basin is dated from Early Cretaceous to Middle Palaeogene [Reyes, 1972, Salfity, 1982, Marquillas *et al.*, 2007, Rohais *et al.*, 2019]. It can be divided in three main stages: (1) the Pigua subgroup (syn-rift) corresponding to alluvial and fluvial deposits; (2) the Balbuena subgroup (transitional sag) corresponding to aeolian, fluvial and shallow lacustrine mixed deposits; and (3) the Santa Barbara subgroup (post-rift) corresponding to fluvio-lacustrine deposits (Figure 1B).

Several sub-basins exist during the rift evolution of the Salta Basin with two main ones within the study area: Alemania and Metán sub-basins (Figure 1A). In the western edge of the Metán sub-basin, the Yacoraite Formation (formation of interest) is a 220 m thick formation which was deposited during

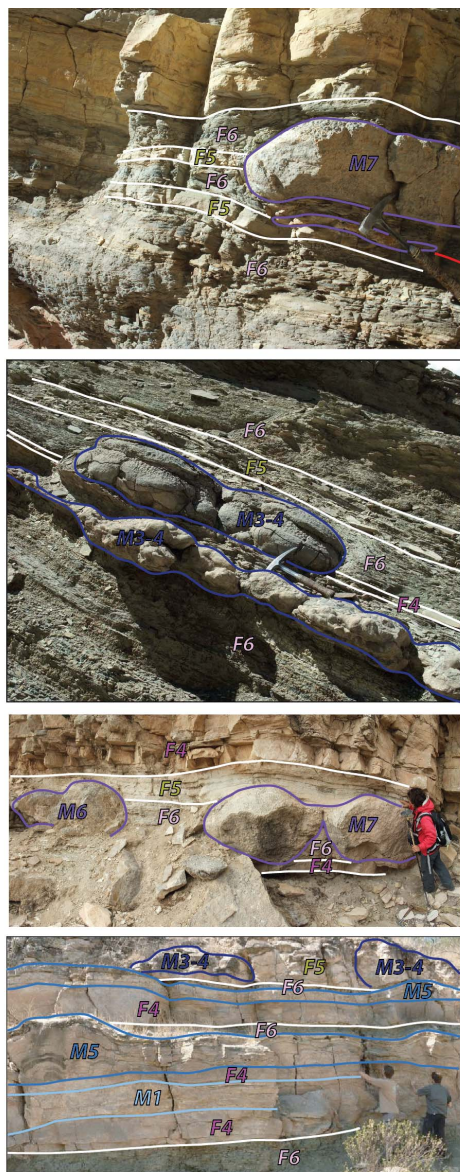




**Figure 1.** (A) Location of the Salta rift basin and the Alemania-Metán sub-basin [modified from Carrera *et al.*, 2006, 2009, Cristallini *et al.*, 1997, Salfity, 1982, Viramonte *et al.*, 1999]. The red line corresponds to the correlation transect presented in Figure 1C. The blue rectangle corresponds to the area modelled in this article. (B) Sedimentological section of the Yacoraite Formation and its three main members (Amblayo, Guemes and Alemania). The blue rectangle corresponds to the stratigraphic interval modelled in this article. A close-up view of this interval is shown, with its subdivision into two units: A1.A and A1.B. (C) Correlation panel of the stratigraphic interval at basin scale modelled in this article. The blue rectangle shows the Cabra Corral area location (the area modelled in this article). The black vertical lines indicate the sedimentary sections and well traces. The two stars indicate the wells informed in terms of petrophysical properties, used in the modelling project.

the sag phase (initiation of thermal subsidence) of the rift evolution after the syn-rift phase characterized by a tectonic subsidence. Its age ranges between Maastrichtian to Danian [Marquillas, 1985, Montano *et al.*, 2022, Moroni, 1982, Rohais *et al.*, 2019] and

it is characterized by a mixed carbonate-siliciclastic sedimentation and a large diversity of microbial carbonates. Based on the observation of constituent facies, vertical stacking, lateral facies change and overall geometry performed during a basin-scale



**Figure 2.** Field photographs showing several examples at various scale, of lateral stratigraphic relationship between microbialite geobodies and host sedimentary facies (hammers and geologists are visible for scale). Note that microbialite geobodies can pass laterally and vertically to several host facies. Facies association and microbialite names and colors refer to Figures 3 and 4.

study of the Alemania-Metán-El Rey sub-basin outcrops, Deschamps *et al.* [2020] defined 10 genetically related facies associations (Table 1): alluvial deposits (FA1), deltaic deposits (FA2), sandflat deposits (FA3), shoreface and low-energy shore (FA4), offshore/profundal deposits (FA5), mixed mudflat deposits (FA6), high-energy carbonate shoreface deposits (FA7), low-energy eulittoral deposits (FA8), oolitic bank (FA9), highly alternating lacustrine deposits (FA10). Based on this sedimentary analysis and on the facies stacking pattern, they also defined a precise stratigraphic framework and divided the succession into four mid-term stratigraphic sequences [numbered 1–4, bounded by sequence boundaries SB1 to SB5; Deschamps *et al.*, 2020] controlled by climate variations. These mid-term sequences can be subdivided into short-term sequences which mainly recorded high frequency climate variations, forced by earth eccentricity variations modulated by obliquity variations [Deschamps *et al.*, 2020]. The short-term sequences are themselves divided into very short-term sequences visible at the scale of alternating facies and included within the facies associations (see for example facies S5 in FA3 defined by Deschamps *et al.* [2020] modelled in FA5 mudflat to sandflat) (Table 1).

It is also characterized by several forms of microbial carbonates or microbialites (Figure 2) which constitute distinctive geobodies with various sizes, macroscopic morphologies, microscopic textures and petrophysical properties. Based both on macrostructure (field observations) and microstructure (thin section observations), these microbialites were grouped in seven categories [Deschamps *et al.*, 2017, Hamon *et al.*, 2012, 2017, Rohais *et al.*, 2011, 2012] following the terminology defined by Riding [2000] and completed by Riding [2008]: (1) planar mats (M1), (2) stromatolite isolated domes (M2), (3) stromatolite coalescent domes and columns (M3-4), (4) agglutinating stromatolites (M5), (5) thrombolite isolated domes (M6), (6) thrombolite coalescent domes (M7), (7) oncoidal rudstone (M8).

**Table 1.** Facies association, description, related facies associations in Deschamps *et al.* [2020], reservoir quality and lateral extension of the facies

Model FAs (this paper)	Description in the model area	Deschamps <i>et al.</i> , 2020 FAs	Deschamps <i>et al.</i> , 2020 facies	Reservoir properties	Porosity data	Lateral continuity
F1: Alluvial plain	<ul style="list-style-type: none"> <li>Reddish silty shales and siltstones with root traces</li> <li>Rare decimeter-scale beds of fine-grained rippled sandstones</li> </ul>	FA1	<p>S1: Massive gravelly sandstones</p> <p>S2: Med. To coarse sandstones with trough-cross bedding</p> <p>S3: Red siltstones with rootlets</p>	Tight clayey material	1 value Mean = 0.15% Max = 0.15%	Less than 1 km
F3: Siliciclastic-dominated shoreface	<ul style="list-style-type: none"> <li>meter-scale, fine, medium to coarse sandstones with plane-parallel to low-angle cross-bedding</li> <li>Rare sandy oolitic grainstone beds</li> </ul>	FA4a FA4b	<p>S7: Medium to coarse sandstones with plane parallel to low-angle cross-bedding</p> <p>S9: Fine to medium sandstones with HCS</p> <p>S8: Fine to medium sandstones with plane parallel to wave ripples</p> <p>M4: Sandy gastropod/oooid rudstone</p> <p>M5: Sandy oolitic grainstone</p>	Theoretical good reservoir properties	8 values Mean = 3.31% Max = 8.7%	Over 1 km
F4: Carbonate-dominated shoreface	<ul style="list-style-type: none"> <li>Coarse-grained peloidal, oolitic and bioclastic (ostracods and gastropods) grainstone to packstone (decametric to metric beds)</li> <li>Centimetric to decametric layers of silty dolomitic marls to mudstones</li> </ul>	FA7 FA10a	<p>C5: Oolitic grainstone</p> <p>C3: Coated ostracod grainstone to packstone</p> <p>M4: Sandy gastropod/oooid rudstone</p> <p>M2: Ostracod wackestone</p> <p>C9: Oncooid rudstone</p> <p>Cl1: Peloid/lithoclast packstone</p> <p>M3: Sandy to silty gastropod floatstone</p> <p>C4: Grapestone-oooids grainstone to packstone</p> <p>C6: Green silty dolomitic marls</p> <p>C8: Intraclastic oncooidal breccia</p> <p>C9: Oncooid rudstone</p>	Theoretical good reservoir properties	52 values Mean = 3.81% Max = 15.7%	Over 1 km
F5: Restricted mudflat to sandflat	<ul style="list-style-type: none"> <li>Thinly laminated siltstones to fine-grained sandstones</li> <li>Bioturbated mudstone to wackestone</li> <li>Wave and current ripples made of medium-grained sandstones and sandy oolitic packstones</li> <li>Frequent desiccation cracks, generally associated with intraclastic breccia made up of stromatolite fragments</li> </ul>	FA3 FA6	<p>S5: Alternating silt and fine-grained sandstones, with ripples</p> <p>S6: Med. To coarse sandstones with sigmoidal cross-bedding</p> <p>M4: Sandy gastropod/oooid rudstone</p> <p>M5: Sandy oolitic grainstone</p> <p>C7: Laminated/bioturbated mudstone to wackstone</p> <p>M1: Green to brown shales</p> <p>M4: Sandy gastropod/oooid rudstone</p> <p>M5: Sandy oolitic grainstone</p> <p>S6: Med. To coarse sandstones with sigmoidal cross-bedding</p>	Tight silty and clayey material	11 values Mean < 2% Max = 5.73%	Good lateral continuity but extension below the kilometer
F6: Offshore	<ul style="list-style-type: none"> <li>Silty dolomitic marls to mudstones</li> <li>Frequent, decimeter-thick intercalations of grapestones and oolites with wave ripples</li> </ul>	FA10b	<p>C6: Green silty dolomitic marls</p> <p>C4: Grapestone-oooids grainstone to packstone</p> <p>C8: Intraclastic oncooidal breccia</p>	Tight marly and clayey material	21 values Mean = 1.25% Max = 15.38%	Up to 3 km
F7: Anoxic central lake	<ul style="list-style-type: none"> <li>Presence of black organic-rich laminated shales to mudstones</li> </ul>	Variant of the FA10b	<p>M7: Black organic-rich shales to mudstone</p> <p>C6: Green silty dolomitic marls</p>	Tight marly and clayey material	1 value Mean < 1%	Up to 3 km

This study focuses on the 8.5 by 5 km<sup>2</sup> area of interest located around the Cabra Corral area (blue square on Figure 1) and on the top of the Alemania Member corresponding to the mid-term sequence #4 defined by Deschamps *et al.* [2020]. This sequence is itself subdivided into two short-term sequences (Al.A and Al.B; Figures 1B and C), by a major exposure surface. Each short-term sequence is characterized by a shallowing upward trend with siliciclastic-dominated lower part and a carbonate-dominated upper part (Figure 1C), associated to climatic variations controlled by eccentricity cycles. The tectonic context does not play a role during this sag phase of the basin evolution. The general architecture and the facies association distribution for the sequence 4 at basin-scale based on chronostratigraphic correlations are proposed on Figure 1C. In this area, this sequence 4 is characterized by a high alternation of proximal and distal facies interpreted as very short-term sequences, recording climatically controlled lake level variations (obliquity cycles; Deschamps *et al.* [2020], Gomes *et al.* [2020] and Magalhães *et al.* [2020]). The impact of the rapid variation of the lake level is enhanced by the shallow depths of the lake prevailing during the deposition of the sequence 4 [Deschamps *et al.*, 2020]. It leads to sharp interbedding of distal facies associations from the lacustrine offshore domain and facies associations from the proximal domain (mudflat/sandflat and shore facies associations) but also of the microbialite forms. Usually, microbial activity occurs in proximal conditions where oxygen, light and nutrients are available. Microbialites can form positive relief geobodies that can be surrounded or capped by facies association from distal domains when the lake level rises bringing into lateral contact microbialite forms with facies associations from offshore domain. Similar observations in other microbialite-bearing formations were reported by other authors [Bunevich *et al.*, 2017, Loucks, 2018]. Figure 2 illustrates at different scales some relationships between microbialite forms and facies associations. It highlights the two types of heterogeneity (sedimentary facies association and microbialite forms) affecting the Sequence 4 of the Yacoraite Formation as well as their geometrical relationships. The heterogeneity coming from microbialite geobodies is not embedded in the facies association heterogeneity but rather spreading across the sedimentary primary

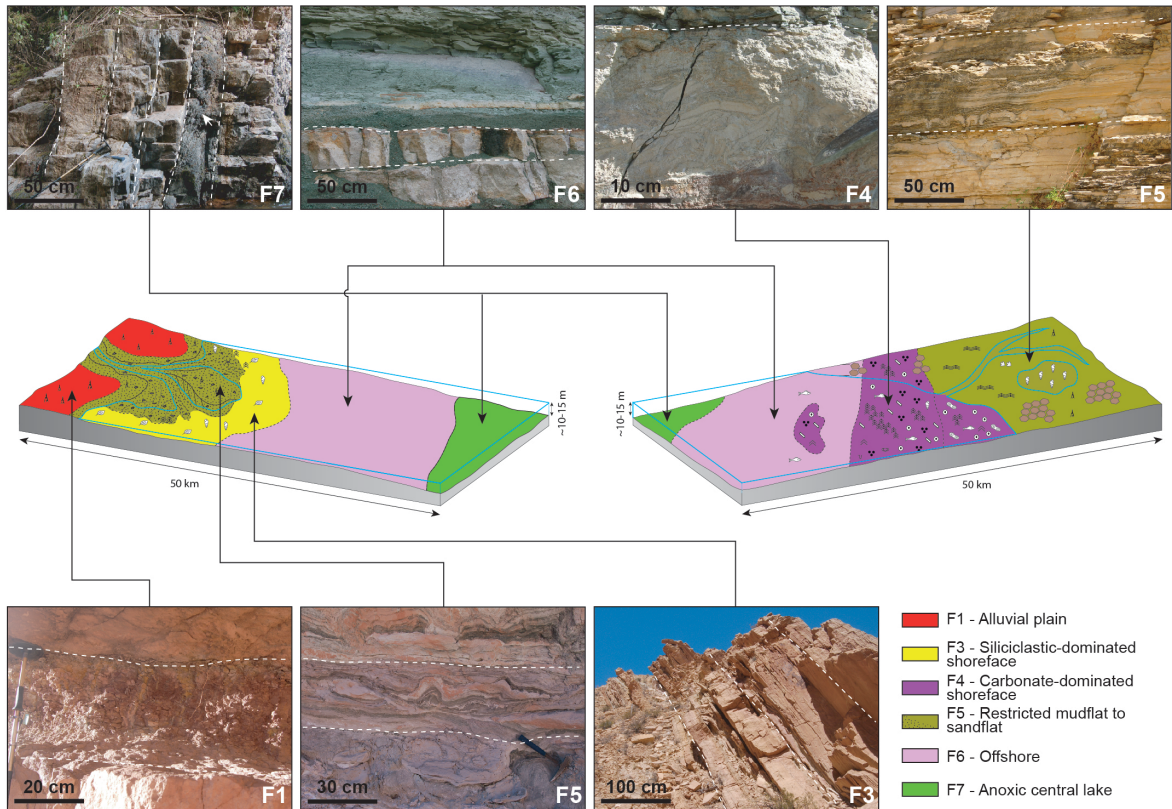
heterogeneity.

### 3. Data

This work is based on the large database (field data, conceptual, sedimentological, diagenesis studies and models) resulting from the Joint Industrial Project COMPAS led by IFP Energies Nouvelles from 2012 to 2015. Both sedimentary and microbialite facies classifications were adapted for this study since some of the ten facies associations defined by Deschamps *et al.* [2020] are not present in the area and in Sequence 4. Others were merged into a single facies as they are similar in terms of reservoir properties (Table 1). For modelling purposes, six facies associations were defined (Figure 3; Table 1). They are presented with their equivalent to the facies association (FA) of Deschamps *et al.* [2020] in Table 1.

These facies associations have been distributed on two separate depositional profiles representative of the Sequence 4 in the Alemania-Metán basin (Figure 3): the first one, a siliciclastic-dominated lacustrine margin is typical of the western part of the basin; the second one, a carbonate-dominated lacustrine margin is characteristic of the eastern part. The studied area is located in the central area of the basin, consequently, both of these conceptual depositional models are relevant at different periods of the Sequence 4.

For modelling purpose, only four simplified microbialites forms were retained (Figure 4) from the original classification [Deschamps *et al.*, 2017, Hamon *et al.*, 2012, 2017, Rohais *et al.*, 2011, 2012]: (1) planar mats (M1) which constitute the most extensive microbial geobodies as they can span horizontally over a kilometer, forming pluridecimeteric to plurimetric thick beds; (2) stromatolite coalescent domes and columns (M3-4) whose spatial isotropic extent is of the order of half kilometer horizontally and 1 m vertically; (3) agglutinating stromatolites (M5) which form geobodies of similar dimensions than the latter; (4) thrombolites, forming coalescent domes (M7), forming meter-scale (height and diameter), generally flattened domes, clustered in discontinuous patches to continuous layers (tens to hundreds of meters of lateral extent). M7 is one of the microbialite forms which exhibits higher porosity. Although it is relatively scarce in the modelled area, it is more present in the western part of the basin.

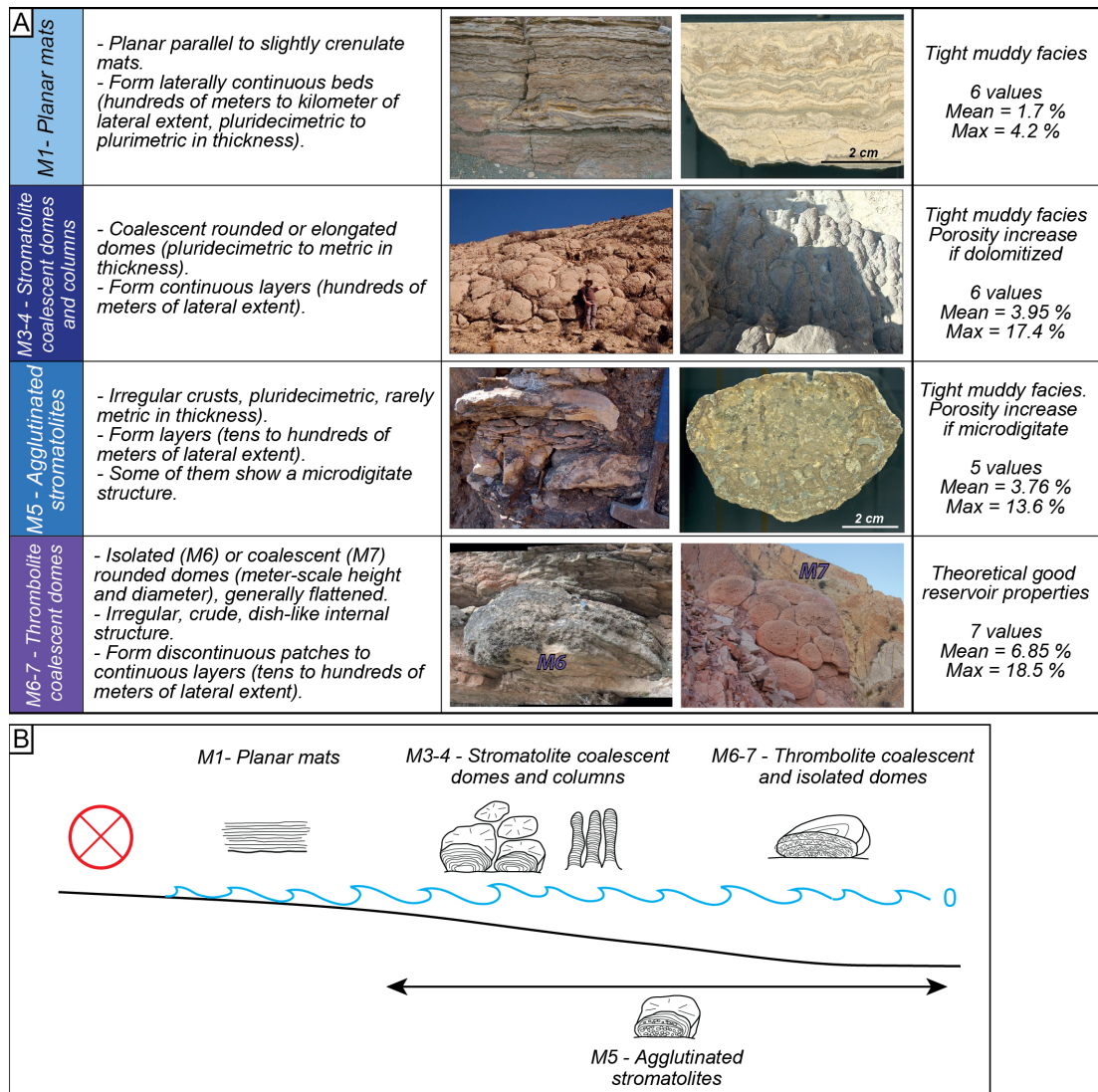


**Figure 3.** Conceptual depositional models for the Alemania Member and associated background facies associations used in this study. Two depositional models are distinguished: on the left, the western siliciclastic-dominated lacustrine margin; on the right the eastern carbonate-dominated lacustrine margin showing the spatial distribution and succession of the different facies associations in each case. On the pictures of the facies, dashed lines indicate the bedding.

This microbialite form was described also in other sites [Muniz and Bosence, 2015, Bosence and Gallois, 2022, Kirkham and Tucker, 2018]. For these reasons, it was kept in the modelling process as it may have an important impact on the reservoir petrophysical properties. A fifth class (M9-No Microbial) is used to represent in the model the absence of microbialite. The stromatolite and thrombolite isolated domes (M2 and M6) have small dimensions (1 m<sup>2</sup> horizontally and 0.5 m vertically) compared to the grid cell dimensions. They were excluded from the modelling process as the impact of these geobodies will not have a significant nor sensible effect on the heterogeneity and petrophysical properties at the considered reservoir scale. The oncoidal rudstone (M8) forms are almost absent of this area, they were

observed in less than 0.25% of the section levels and thus not considered for the model.

In the Cabra Coral area, 17 sedimentological sections have been described at a vertical resolution of 10 cm in terms of both sedimentary facies association and microbialite forms as presented above. During the numerical formatting of the field observations, both categorical data were informed for each sample or levels. When microbialite facies were absent, the description was set to (M9) “No Microbial”. The locations of the described sedimentary sections are quite well distributed in the area (Figure 7), so the facies vertical distribution can be compared in all spatial directions. There is no evidence of any lateral trend in terms of vertical distributions. This stationarity was expected consistently with both the reservoir-scale

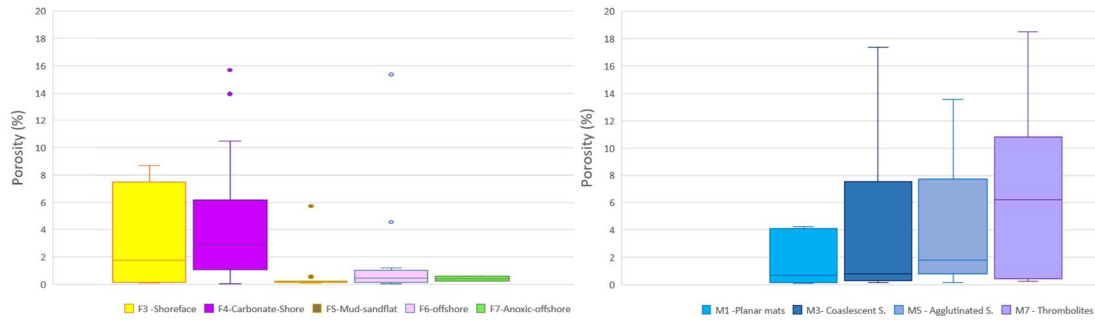


**Figure 4.** (A) Four simplified categories of microbialite forms used for the simulation workflow and their characteristics in terms of geometry and dimensions. The color associated to each category corresponds to the color code used in the simulation workflow. (B) Relative spatial distribution of the microbialite forms on a schematic depositional profile.

dimension and the central position of the grid in the basin. Therefore, we assume that the simulation does not need additional horizontal trend. It should be noted that the alluvial facies association (F1) is marginal in terms of percentage and could have been disregarded in respect to its impact on the porosity mapping. However, this facies type is a useful marker bed that identifies the disconformity (exposure surface) between the units Al.A and Al.B. This facies association has a meaning in terms of stratigraphy with

the occurrence of more proximal environment at this level. It was kept in the modelling process to keep this stratigraphical information that could be a potential local barrier to fluid flow between the 2 units due to the lower permeability of these siltstones and shales deposited in an alluvial/floodplain environment compared the other facies associations.

The relative proportions of microbialites per facies association (Table 2) have been computed from the 17 available sections. Although microbialites grow



**Figure 5.** Porosity distribution per sedimentary facies association (left) and microbialite forms (right). See also Table 1 for mean and max values for each facies.

**Table 2.** Relative proportions in percentiles (%) of microbialite forms per facies association in the 17 observed sections

Sedimentary facies	Microbial form	M1	M3-4	M5	M7	M9
		Planar mats	Coalescent stromatolites	Agglutinated stromatolites	Coalescent thrombolite	No microbial
F1—alluvial	<span style="color:red">■</span>	0	0	0	0	100
F3—siliciclastic shore	<span style="color:yellow">■</span>	1.2	0.7	0	0	98.1
F4—carbonate shore	<span style="color:magenta">■</span>	11.8	7.5	8.9	0.4	71.4
F5—mud-sandflat	<span style="color:brown">■</span>	19.8	11.1	22.3	0.4	46.5
F6—offshore	<span style="color:pink">■</span>	21.7	7.7	4.2	3.5	63
F7—anoxic offshore	<span style="color:green">■</span>	0	0	0	0	100

in proximal environments, they might be laterally in contact with distal facies association due to the sharp interbedding of the facies as discussed above in the section on the geological context. It explains the significant proportion of microbialite forms associated with the offshore facies association (F6).

Porosity quantifications were performed by threshold color extraction and point counting methods on scanned thin sections with the free software “JMicroVision” [Roduit, 2007]. Thin sections from 2 wells studied during the basin-scale study [Deschamps *et al.*, 2017, 2020, Hamon *et al.*, 2012, 2017, Rohais *et al.*, 2011, 2012] and close to the modelled area were considered: Juramento which is at a distance of 6 km east of the modelled area and “CoreLP05” located 25 km away in the NE direction. Overall, 130 estimated values are available, 62 and 68 respectively coming from Juramento and from “CoreLP05” sections. Sedimentary facies and microbial forms were observed via respectively 96 and 34 thin sections. Figure 5 exhibits the distributions of

these porosity values for each facies or microbialite type. It shows that microbialite forms may exhibit higher porosity values than the facies associations except for the planar mats (M1).

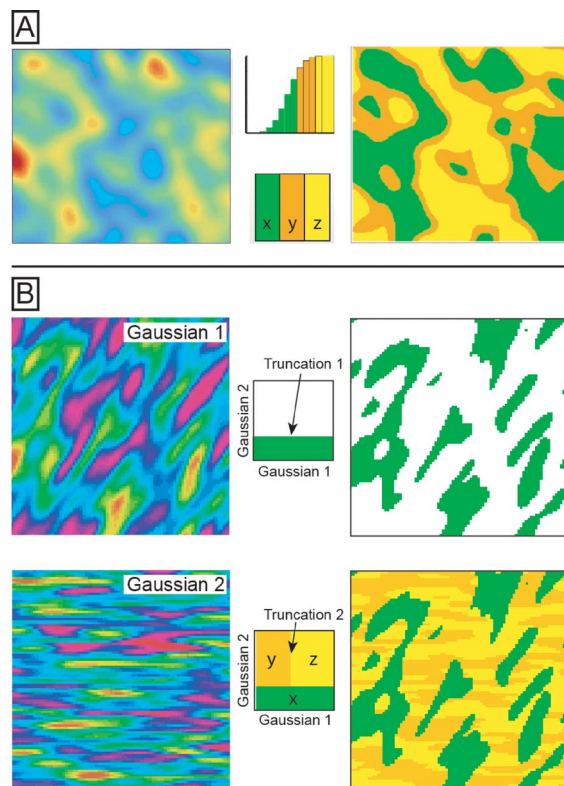
#### 4. Methodology

##### 4.1. Bi-variate Pluri-Gaussian simulation

In order to consider jointly sedimentary facies associations and microbialite forms as two distinct properties, the bivariate plurigaussian (bi-PGS) simulation method [Renard *et al.*, 2008, Hamon *et al.*, 2016] has been applied here. It was developed to process two categorical variables, while accounting for their correlation, their continuity and their potential heterotopy. This bi-PGS method is a sprout in the plurigaussian family tree. The first member of this family was the truncated gaussian methods (TGS) published by the reference papers of Galli and Beucher [1997], Journel and Isaaks [1984], Matheron *et al.* [1987]. It

is based on one spatial gaussian random function truncated by given threshold values that mark off different areas assigned to the different values of the simulated categorical variable (e.g. different geological facies) (Figure 6A). Plurigaussian simulations (PGS) were developed to provide greater flexibility to realistically simulate geological environments [Armstrong *et al.*, 2011, Emery, 2007, Madani and Emery, 2015]. Gaussian random functions (correlated or not) are generated and simultaneously truncated to simulate 3D distribution of geological facies. Each underlying Gaussian function is defined by a variogram type, direction, major and minor ranges. Each one is compliant to a zero mean and a unit variance. They shape the global geometry, the direction, and the scales of the simulated structures. The truncation rule has as many dimensions as the number of gaussian functions. In each direction, it defines the succession of facies used when assigning facies to the different truncated parts of the gaussian. In consequence, it strongly controls the succession of facies and the possible contacts between facies (Figure 6B). Threshold values are computed based on the computed proportions of facies in the data and can vary spatially. In the case of a vertical variability, the threshold values are usually computed based on the Vertical Proportion Curve (VPC) which gives the proportions per layer of the grid. To account for lateral variability in addition to the vertical one, the threshold values can be computed on a matrix of proportion (MP) which holds VPCs as elements. In both cases (VPC and MPC), their computation is usually based on the observed data such as wells or sedimentary sections, but external information can be added to the computation for example palaeogeographical maps or seismic derived information. Thanks to its flexibility and to the large panel of possibilities to account for additional data, PGS is now widely applied in petroleum industry for building realistic geological models. It has been used for handling various applications and geological environments [Albertão *et al.*, 2005, de Galard *et al.*, 2005, Doligez *et al.*, 2007, Gasparrini *et al.*, 2017, Hamon *et al.*, 2016, Pontiggia *et al.*, 2010].

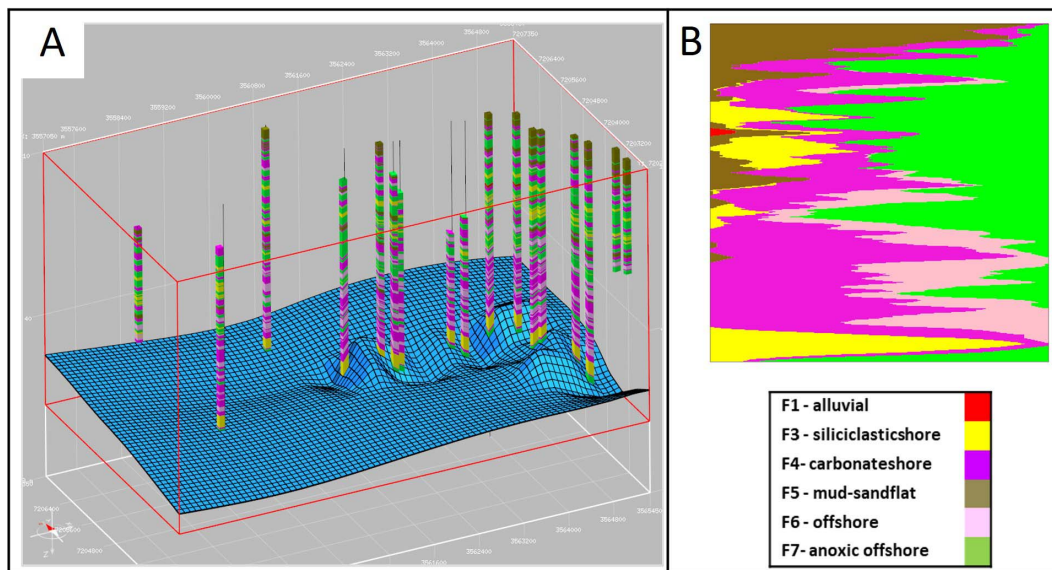
In the case of the bivariate PGS method [Beucher and Renard, 2016, Renard *et al.*, 2008], one categorical variable (usually the geological facies) is simulated by a first complete PGS defined by specific gaussian functions possibly correlated, their variogram model and specific truncation rule defined



**Figure 6.** (A) Basic principles of the Truncated Gaussian Method: transformation of a Gaussian random function into a simulation of three facies  $x$ ,  $y$ , and  $z$  with equal proportions in the domain. (B) Basic principles of the plurigaussian model: transformation of two Gaussian random functions into a simulation of three facies  $x$ ,  $y$  and  $z$ , using the truncation rule illustrated in the middle and the same proportion for the three facies.

according to the sequential and spatial organization of the considered variable. The second categorical variable (for example diagenesis types, intensity or microbialite forms) simulation is defined by a second complete PGS model which parameters might be computed from a completely different dataset (hence the possible heterotopy). The two PGS simulations are linked through the given proportions of the second variable in each of the first variable classes. Finally, the different realizations of the two variables are combined to produce and visualize the joint simulation of both variables through a “combined facies”.





**Figure 7.** (A) 3D view (100 × vertical exaggeration) showing: (i) the bottom surface (SB4) discretized at 100 m resolution upon which is constructed the 3D simulation grid; the top surface is horizontal; (ii) the red 3D box which encompasses the reservoir grid; (iii) the 17 sedimentary sections, represented as vertical wells, with the discretized sedimentary facies association. (B) Vertical proportion curve (VPC) for sedimentary facies association, computed from the 17 wells available in the modelled area.

In our modelling workflow, for each value of the categorical variables, Gaussian distribution of their associated porosity were defined based on the thin-section measurements. The porosity values are drawn at random in these distributions according to the simulated facies.

#### 4.2. Application to the Alemania Sequence 4 in the Cabral Corral area

The 3D reservoir grid has a horizontal regular resolution of 100 m and a vertical one of 10 cm. The top and bottom surfaces were constructed with the Gocad/Skua<sup>®</sup> geomodelling software by Emerson. They are based upon the stratigraphic surface markers [Deschamps *et al.*, 2020] of the sequence boundaries. The top surface is the flattened reference surface and corresponds to the Sequence Boundary SB5 surface of Deschamps *et al.* [2020]. The bottom surface was constructed based on the Sequence Boundary SB4 surface of Deschamps *et al.* [2020] (Figure 1). The grid has the same vertical resolution (0.1 m) than the sections which ensures no loss of information during the numerical conversion. There

is no horizontal spatial trend but a clear vertical organization visible in all wells. Hence only a single representative VPC (Figure 7B) has been computed from the 17 sections with a smoothing over 3 layers.

The Gaussian functions underlying the simulation of both the sedimentary facies associations (GFS1 and GFS2) and microbialite forms (GFM1 and GFM2) were defined according to the lateral continuities observed on the field (Table 1) and the depositional environment interpretation. Carbonate facies (FA5) exhibit a good lateral continuity but with an extent below the kilometer, whereas the other facies show a larger horizontal continuity over 1 km, reaching 3 km for the offshore facies (FA6, FA7). For the microbialite forms, the thrombolite coalescent domal form (M7) is relatively scarce, organized in patches with the smallest lateral extension, whereas the other microbialite forms span laterally over more than half kilometer. Therefore, in addition to different ranges (Table 3), an exponential type Gaussian Function (GFM2) subtends the simulation of the M7 form to enhance the patchy or pixelated structure in opposition to the Gaussian type that leads to smoother and more continuous spatial organization. As no prefer-

ential direction was observed on the field for any facies, variograms are isotropic for both PGS.

The threshold template or truncation rules (Table 3) were designed according to the depositional models (Figures 3 and 4B). In the case of the siliciclastic-dominated margin one, from proximal to distal position, the succession goes from F1 Alluvial, F5 Mudflat, F3 Siliciclastic shoreface to F6 offshore and F7 anoxic offshore. This trend is reproduced horizontally in the truncation rule. Likewise, the carbonate-dominated margin depositional model shows F5 Sandflat, F4 Carbonate shoreface in a proximal position to the offshore facies (F6 and F7) distally. This ordering can be simulated following the second line of the truncation rule. It means also that, following its other dimension, the vertical one, set for the second GFS, F3 siliciclastic shore facies can be in contact to F4 carbonate shore facies as the depositional environment of real shores can change laterally from a siliciclastic-dominated to a carbonate-dominated one farther from clastic sources.

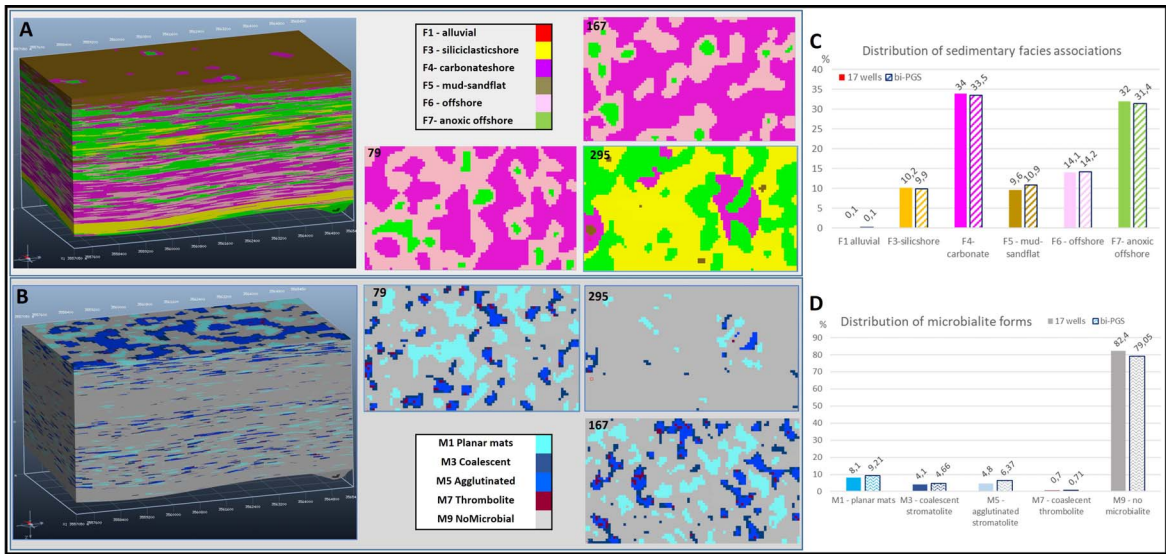
In the truncation rule for the microbialite forms, the planar mats (M1) are separated from the other forms (M3-4, M5 and M7) by the “no microbialite” (M9) facies as it occurs at shallow depths, quite different from the depths of the other microbialite forms. The agglutinated forms (M5) may occur at the same locations than the stromatolite and thrombolite coalescent forms (M3-4 and M7). Hence, the template rule is defined so that the simulated microbialite forms M5 can be in contact with both M3-4 and M7. The stromatolite coalescent forms (M3-4) being more continuous than the M5, the latter was positioned with M7 on the direction defining the succession, for exponential Gaussian function (GFM2). The relative proportions of microbialite forms in each facies association (Table 2) is used to link the two PGS simulations. Due to the available porosity values, the different distributions (Figure 5) were used for their respective sedimentary facies associations and microbialite forms.

## 5. Results

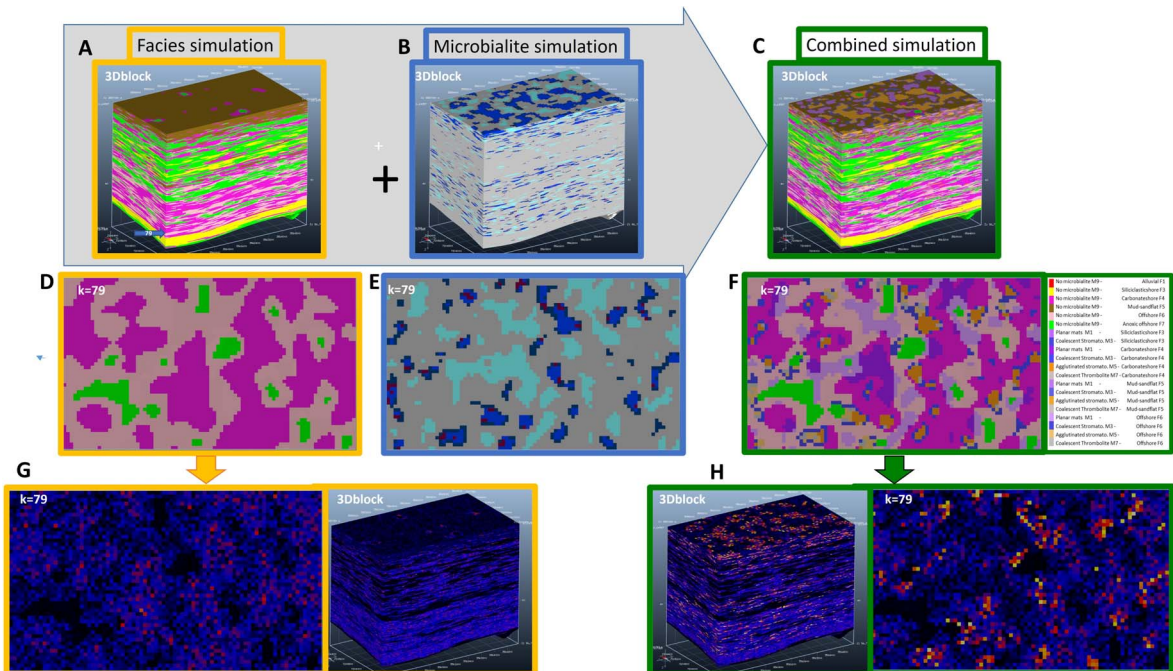
Figure 8 illustrates one realization of the bi-variate simulation in terms of sedimentary facies associations and microbialite forms. The 3D block view (Figure 8A) shows that the vertical trend constrained by the VPC is well represented reproducing the main

stratigraphic architectures visible in Figure 1C. The plan views show the horizontal distribution and lateral succession of the sedimentary facies association following the depositional model (Figure 3) imposed through the truncation template. This depositional model must be considered as a substitution facies diagram rather than a paleogeographic model. It represents the spatial facies distribution and relationships related to bathymetry. Due to the shallow depth of the lake and the palaeo-topography of the lake floor, a patchy facies distribution was expected, and correctly modeled. On palaeo-topographic highs, simulated mud/sand-flat facies associations (F5) pass laterally to carbonate (F4) or siliciclastic (F3) shore facies. In palaeo-topographic lows or depressions, anoxic offshore facies (F7) are simulated in the center and generally surrounded by simulated offshore (F6) or shore facies. Lateral extension of the simulated sedimentary facies reaches a few kilometers to hundreds of meters in accordance with the gaussian function ranges. Vertically, the simulated microbialite forms (Figure 8B) are distributed in agreement with the distribution of the simulated mud/sand-flat (F5) and carbonate shore (F4) facies associations, where they are the most present. They are organized horizontally in different kilometer-scale patches separated by the no-microbialite facies (M9): on one side, the planar mats (M1) are the most extended simulated facies with a maximum continuity of a few kilometers; on the other side, the three other microbialite forms compose the other patches. The simulated coalescent stromatolite forms (M3-4) are located around the agglutinated stromatolite (M5) and coalescent thrombolite (M7) forms. The exponential model of the Gaussian function GFM1 leads to scattered M7 forms which are the scarcest microbialites in the studied area with the shortest lateral extension. For both categorical variables, the simulation reproduces globally the distribution at the observed sections (Figures 8C and D). Except for the F5 facies association, the differences in the proportions of the observed and modelled sedimentary facies (Figure 8C) are below 1.4% with a mean difference below 0.5% (RMSE = 0.63). For the microbialite forms, the maximum difference (Figure 8D) is 3.35% for the majority (82.4%) “no microbialite” category and mean difference of 1.32% (or RMSE 1.75%).

The bi-variate PGS simulation also produces a

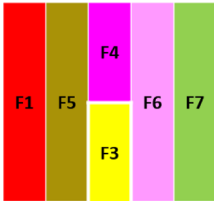
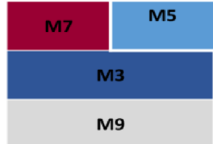



**Figure 8.** Bi-variate simulation of sedimentary facies associations (top) and microbialite forms (bottom). (A, B), 3D views and 3 plan views of the layers 79, 167 and 295. (C, D), comparison of the proportions between simulation (right with colored patterns) and observed sections (left with plain colors).



**Figure 9.** Top, 3D views of the simulated facies associations (A), microbialite forms (B) and combined facies (C), white arrows underline the location of the layer  $k = 79$  shown below for each categorical variable (D–F); Bottom, simulated porosity when considering only sedimentary facies association (G); Simulated porosity when considering the combined facies (sedimentary facies associations + microbialite forms). For these two categorical variables, color legends are defined in Figure 8.

**Table 3.** Bi-variate plurigaussian simulation parameters: truncation rules and gaussian functions variogram

	Truncation rule	Model	Horizontal range	Vertical range	
Sedimentary facies association		Gaussian	800 m	0.4 m	GFS1
		Gaussian	2000 m	0.4 m	GFS2
Microbialite forms		Exponential	200 m	0.4 m	GFM1
		Gaussian	500 m	0.4 m	GFM2

joint simulation of both categorical variables in a combined model and thus providing a visualization of the combined heterogeneity (Figures 9C and F). Planar mats (M1) as well as agglutinated stromatolites (M5), respectively colored in different shades of purple and orange, are simulated across sedimentary facies associations. Furthermore, two simulations of porosity were performed for comparison: the first with porosity distributions defined only according to the sedimentary facies association (Figure 9G), the second with different distributions for each combined facies which also takes into account the simulated microbialite form (Figure 9H). It shows the impact of the additional heterogeneity due to the microbialite forms. For example, mud/sand-flat (F5) is the most abundant sedimentary facies association in the top layers of the 3D block (Figure 9A), characterized by very low porosities (dark blue color on Figure 9G). The addition of microbialite forms such as M3-4, M5 and M7 brings patches of higher porosity in top layers (warm colors in the top layers of the 3D block on Figure 9H). Conversely, planar mats (M1) bring patches of lower porosity when they occur in sedimentary facies association of higher porosity for example carbonate shore facies (F4). On the plan view of layer 79 (Figure 9H), both effects can be identified by comparison with Figure 9G: higher values of porosity are visible where M3-4, M5 and M7 forms are simulated; conversely, additional and diffuse patches of lower

values are simulated with the occurrence of M1.

## 6. Discussion

### 6.1. Data and integration in the modelling process

The defined sedimentary facies associations regroup several sedimentary facies and represent different types of heterogeneity at smaller scale. It implies a loss in details of the description. This can be an important consideration especially when dealing with the petrophysical properties in such varying geological context. As the porosity values were computed on thin sections of characteristic sedimentary facies from 2 cores in the vicinity, we believe we captured the distribution of the porosity for the sedimentary facies associations. The vertical resolution of the grid is exactly the same than the observed sedimentary sections in the field (10 cm). No loss of information occurred during the digitalization of the data as no upscaling nor averaging was needed. Even so, the representativity of these values for the sedimentary facies associations of the area remains questionable and could be further investigated going back in details to the field observations and also by means of uncertainty analysis. It could be also an important consideration when defining the associated permeability.

The alluvial facies association represents only 0.1% of the facies associations observed at the wells. Even if it has an insignificant volumetry, this facies represents the most proximal facies of the “siliciclastic depositional model” and has, as such, an important stratigraphical meaning. This is somehow translated in the simulation through the truncation rule which is defined upon the depositional models and controls the sequence of the simulated facies. So, there is a significant impact for this level to keep this alluvial category. The truncation rule is a way to integrate soft data such as geological interpretation and conceptual models in terms of depositional environments in the stochastic modelling process. With the addition of the ranges defined from the expected continuity of the different sedimentary facies associations, allowing hence the representation of a realistic succession and size of the simulated facies associations.

The modelled microbialite forms were chosen for their potential impact on fluid flow in terms of both lateral and vertical extensions and porosity, and their relevance in relation to the grid cell dimensions. These choices as well as the modelling technique depend on the scale of interest and the model resolution. Such problematics of scale and representativeness were already emphasized by several authors who proposed different approaches. For example, in order to realistically simulate different heterogeneity levels in shoal reservoirs, Petrovic *et al.* [2018] used two nested 3D geo-cellular facies models: (1) a large-scale (30 × 30 km) model based on truncated Gaussian simulation (TGS), forming the basis for (2) a smaller scale (10 × 10 km), more detailed model based on multiple point statistics. Amour *et al.* [2013] also proposed a step-by-step methodology to simulate different orders of stratigraphic sequences and lithofacies. These authors combined a deterministic, surface-based modelling approach to represent the large-scale stratigraphic framework, in which they used stochastic methods (truncated Gaussian and sequential indicator simulations) to distribute respectively the facies association and lithofacies. This stresses out the need to elaborate scale-dependent modelling approaches to take into account a “statistical Representative Elementary Volume (REV)” if one wants to accurately represent the distribution of small-scale heterogeneities. Here the model has a 100 m horizontal resolution and 10 cm vertical reso-

lution which is relevant to model flow barriers. Even with resolutions down to 50 m or 25 m, the choice of the considered microbialite forms in the model would have been the same as the disregarded forms were meter-scale isolated stromatolite domes.

The diagenesis imprints were not considered in this study. This process could have an important impact on the petrophysical properties. One could consider that its impact is accounted for globally in the quantification of the porosity values on the two wells samples used here. Indeed, they are quite low compared to the porosity values measured from plugs and logs from similar facies in the subsurface oil fields in the nearby Lomas de Olmedo basin probably due to different diagenetic processes and pathways [Cesaretti *et al.*, 2000, Durieux and Brown, 2007]. This study of an outcrop is not completely relevant in terms of porosity values for subsurface similar reservoir but interesting as a methodological example performed with a coherent dataset. Furthermore, it would be interesting to study the diagenesis imprints, timing and processes to better understand its spatial distribution which might be controlled by primary sedimentary and microbialite heterogeneity and initial porosity and permeability.

Structural or mechanical constraints may also have an important impact on the final porosity and hydraulic conductivity, through fracturing to be specific. It may be relevant to consider it and requires the definition of relationship between fracturing and the petrophysical properties. In the modelled area, some fractures and micro-fractures, mostly localized, have been observed at outcrop and in thin sections. The distribution of the significant structural elements could be added as an extra-step in the modelling workflow. These fractures would be modelled as DFN (Discrete Fracture Network) based on specific descriptive structural elements such as maps, fracture density, persistence. The simulation of the DFN may also be further controlled by larger structures such as faults or facies as shown by Lapponi *et al.* [2011]. It requires an important specific dataset gathered for this purpose and the identification of existing relationships between larger structures or facies which is not the case in this study.

## 6.2. Simulation results vs observed heterogeneity

The bi-variate method allows the simulation of two types of heterogeneity: the one due to the sedimentary facies variability and the one due to the microbial geobodies. For the first, it is thus possible to consider both depositional models valid for this mixed carbonate and siliciclastic environment. The microbialite heterogeneity can be simulated with its own spatial distribution and correlation dimensions through the parameter definition of the second PGS. The simulated planar mats (M1) have a kilometer scale lateral dimensions and are separated from the other simulated forms in accordance with the conceptual model (Figure 4B). The least extended coalescent thrombolite form (M7) is simulated in a few cells giving place for the agglutinated stromatolite form (M5) which has a larger lateral extension, in good agreement with the associated larger fetch of the depositional environment. The PGS method leads subsequently to the coalescent stromatolites M3-4 simulated generally as a ring or at least in the border of the M5 and M7 simulated forms, which is also coherent with the conceptual model (Figure 4B), where the M3-4 forms is in a shallower position and would appear closer to the lake shore. The correlation with the sedimentary PGS adds consistency to the whole simulation. This advantage is even more blatant when the available data of the categorical data are heterotopic, which is not the case in this study where data are available at the same locations and with the same resolution.

The bi-variate PGS method is well adapted to the case where the microbial geobodies are overprinting the sedimentary heterogeneity. It would be difficult to model this case with a nested approach as the microbial geobodies are not embedded in given sedimentary facies. A single PGS would not be sufficient to embed as many geological concepts and interpretations, spatial continuity and distribution. Some authors suggest the use of growth models [Curtis *et al.*, 2021, Dupraz *et al.*, 2006, Gallois *et al.*, 2016, Johnson and Grotzinger, 2006, Kozłowski *et al.*, 2014, Xi *et al.*, 2022]. However, the controlling factors of these processes are complex and difficult to infer from the simple sedimentary facies or depositional environments. An alternative could be to superimpose to a sedimentary simulation a Boolean method with different objects representing the different types of mi-

crobial geobodies.

## 7. Conclusion

A reservoir-scale geological numerical model of the sequence 4 at the top of the Yacoraite microbialite-bearing carbonate Formation has been generated with the bivariate plurigaussian approach proposed by Renard *et al.* [2008]. This method is able to reproduce both types of heterogeneity: the one due to the sedimentary facies and the microbialite geobodies due to the microbial activity. Simulation of the porosity associated with the combined heterogeneity allows a more realistic description of the volume/storativity of a reservoir and its connectivity. We present how we integrate concepts from the geological interpretation and conceptual models in the definition of the simulation method. We believe this work is a good illustration of how fun and fascinating dealing with the heterogeneity of such complex carbonate rock can be.

## Conflicts of interest

Authors have no conflict of interest to declare.

## Acknowledgments

Part of this work was supported by companies in the frame of the Joint Industrial Project COMPAS led by IFP Energies nouvelles: BHP Billiton, BP, Chevron, Cobalt Inc., Ecopetrol, ENI, ExxonMobil, GalpEnergia, Total, MaerskOil, Petrobras, Repsol, YPF S.A. We also really appreciated the support of José Salfity during several field excursions. Developments of the used prototype embedding the bi-PGS method were supported by Repsol. We would like also to thank the anonymous reviewer whose constructive comments helped in improving this paper.

## References

- Ahr, W. M. (2008). *Geology of Carbonate Reservoirs*. John Wiley & Son, Hoboken.
- Albertão, G. A., Grell, A. P., Badolato, D., and dos Santos, L. R. (2005). 3D geological modeling in a turbidite system with complex stratigraphic-structural framework—an example from Campos Basin, Brazil. In *SPE 95612, SPE Annual Technical Conference & Exhibition, Dallas (TX), USA*.

- Amour, F., Mutti, M., Christ, N., Immenhauser, A., Benson, G. S., Agar, S. M., Tomás, S., and Kabiri, L. (2013). Outcrop analog for an oolitic carbonate ramp reservoir: a scale-dependent geologic modeling approach based on stratigraphic hierarchy. *AAPG Bull.*, 97(5), 845–871.
- Armstrong, M., Galli, A., Beucher, H., Le Loc'h, G., Renard, D., Doligez, B., Eschard, R., and Geffroy, F. (2011). Plurigaussian simulations. In *Geosciences*, page 165. Springer-Verlag, Berlin, Heidelberg.
- Arslan, I., Ribeiro, M. T., Al Neaimi, M., and Hendrawan, I. (2008). Facies modelling using multiple-point statistics: an example from a carbonate reservoir section located in a small part of a large shelf margin of Arabian Gulf, UAE. In *International Petroleum Exhibition and Conference, Abu Dhabi, UAE*. SPE 118089.
- Beucher, H. and Renard, D. (2016). Truncated Gaussian and derived methods. *C. R. Geosci.*, 348(7), 510–519.
- Bosence, D. and Gallois, A. (2022). How do thrombolites form? Multiphase construction of lacustrine microbialites, Purbeck Limestone Group, (Jurassic), Dorset, UK. *Sedimentology*, 69, 914–953.
- Bunevich, R., Borghi, L., Gabaglia, G. P. R., Terra, G. J. S., Bento Freire, E., Lykawka, R., and Fragoso, D. G. C. (2017). Microbialitos da Sequência Balbuena IV (Daniano), Bacia de Salta, Argentina: caracterização de intrabioarquitecturas e de microciclos. *Pesqui. Geoci.*, 44(2), 177–202.
- Carrera, N., Muñoz, J. A., and Roca, E. (2009). 3D reconstruction of geological surfaces by the equivalent dip-domain method: an example from field data of the Cerro Bayo Anticline (Cordillera Oriental, NW Argentine Andes). *J. Struct. Geol.*, 31(12), 1573–1585.
- Carrera, N., Munoz, J. A., Sabat, F., Roca, E., and Mon, R. (2006). The role of inversion tectonics in the structure of the Cordillera Oriental (Argentinean Andes). *J. Struct. Geol.*, 28(11), 1921–1932.
- Cesaretti, N. N., Parnell, J., and Dominguez, E. A. (2000). Pore fluid evolution within a hydrocarbon reservoir: Yacoraite formation (Upper Cretaceous), Northwest Basin, Argentina. *J. Pet. Geol.*, 23(4), 375–398.
- Corbett, P., Hyashi, F. Y., Alves, S. A., Jiang, Z., Wang, H., Demyanov, V., Machado, A., Borghi, L., and Srivastava, N. (2015). Microbial carbonates: a sampling and measurement challenge for petrophysics addressed by capturing the bioarchitectural components. In Bosence, D. W. J., Gibbons, K., Le Heron, D. P., Pritchard, T., and Vining, B., editors, *Microbial Carbonates in Space and Time: Implications for Global Exploration and Production*, volume 418, pages 69–85. Geological Society, London. Special Publications.
- Cristallini, E., Cominguez, A. H., and Ramos, V. A. (1997). Deep structure of the Métan-Guachipas region: tectonic inversion in Northwestern Argentina. *J. South Am. Earth Sci.*, 10(5–6), 403–421.
- Curtis, A., Wood, R., Bowyer, E., Shore, A., Curtis-Walcott, A., and Robertsson, J. (2021). Modelling Ediacaran metazoan–microbial reef growth. *Sedimentology*, 68, 1877–1892.
- de Galard, J.-H., Zoormand, G. H., Ghanizadeh, M., Daltaban, S., and Camus, D. (2005). A case study on redevelopment of a giant highly fractured carbonate reservoir in Iran based on integrated reservoir characterization and 3D modeling studies. In *14th SPE Middle East Oil & Gas Show & Conference, Bahrain, UAE*. SPE 93760.
- de Marsily, G., Delay, F., Gonçalves, J., Renard, P., Teles, V., and Violette, S. (2005). Dealing with spatial heterogeneity. *Hydrogeol. J.*, 13(1), 161–183.
- Deschamps, R., Hamon, Y., Rohais, S., and Gasparini, M. (2020). Dynamic of a lacustrine sedimentary system during late rifting at the Cretaceous–Palaeocene transition: example of the Yacoraite Formation, Salta Basin, Argentina. *Depos. Rec.*, 6, 490–523.
- Deschamps, R., Rohais, S., Hamon, Y., and Gasparini, M. (2017). Dynamic of a lacustrine system in a rift basin: example of the Yacoraite Fm., Salta Basin, Argentina. In *33rd IAS and 16th ASF Joint Meeting, Toulouse, France*.
- Dirner, S. and Steiner, U. (2015). Assessing reservoir uncertainty with stochastic facies modeling of a hydrothermal medium enthalpy reservoir (upper Jurassic carbonates of the southern German Molasse basin). In *Proceedings World Geothermal Congress 2015, Melbourne, Australia*. <https://pangea.stanford.edu/ERE/db/WGC/papers/WGC/2015/32004.pdf>.
- Doligez, B., Beucher, H., Lerat, O., and Souza, O. (2007). Use of a seismic derived constraint: different steps and joined uncertainties in the construction of a realistic geological model. *Oil Gas Sci. Technol.*, 62(2), 237–248.

- Dupraz, C., Pattisina, R., and Verrecchia, E. P. (2006). Translation of energy into morphology: simulation of stromatolite morphospace using a stochastic model. *Sediment. Geol.*, 185(3–4), 185–203.
- Dupraz, C., Reid, R. P., Braissant, O., Decho, A. W., Norman, R. S., and Visscher, P. T. (2009). Processes of carbonate precipitation in modern microbial mats. *Earth Sci. Rev.*, 96, 141–162.
- Durieux, C. G. and Brown, A. C. (2007). Geological context, mineralization, and timing of the Juramento sediment-hosted stratiform copper–silver deposit, Salta district, northwestern Argentina. *Miner. Depos.*, 42, 879–899.
- Emery, X. (2007). Simulation of geological domains using the pluri-Gaussian model: new developments and computer programs. *Comput. Geosci.*, 33, 1189–1201.
- Galli, A. and Beucher, H. (1997). Stochastic models for reservoir characterization: a user-friendly review. In *5th Latin America and Caribbean Petroleum Engineering Conference & Exhibition, Rio de Janeiro, Brazil*.
- Gallois, A., Kozłowski, E. N., Bosence, D. W. J., and Burgess, P. (2016). Multi-scale characterization and modelling of non-marine microbial carbonates: the lower Purbeck Formation (Late Jurassic to Early Cretaceous), Dorset, southern England (UK). In *AAPG and SEG International Conference and Exhibition 2016, Cancun, Mexico*. <http://www.searchanddiscovery.com/abstracts/html/2016/90260ice/abstracts/2475565.htm>.
- Gasparrini, M., López-Cilla, I., Blázquez-Fernández, S., Rosales, I., Lerat, O., Martín-Chivelet, J., and Doligez, B. (2017). A multidisciplinary modeling approach to assess facies-dolomitization-porosity interdependence in a lower Cretaceous platform (northern Spain). In MacNeil, A. J., Lonnee, J., and Wood, R., editors, *Characterization and Modeling of Carbonates—Mountjoy Symposium 1*, SEPM Special Publication 109, pages 130–153. SEPM (Society for Sedimentary Geology), Tulsa, Oklahoma.
- Gomes, J. P. B., Bunevich, R. B., Tonietto, S. N., Alves, D. B., Santos, J. E., and Whitaker, F. F. (2020). Climatic signals in lacustrine deposits of the upper Yacoraite formation, western Argentina: evidence from clay minerals, analcime, dolomite and fibrous calcite. *Sedimentology*, 67, 2282–2309.
- Hamon, Y., Deschamps, R., Joseph, P., Doligez, B., Schmitz, J., and Lerat, O. (2016). Integrated workflow for characterization and modeling of a mixed sedimentary system: the Ilerdian Alveolina limestone formation (Graus-Tremp Basin, Spain). *C. R. Geosci.*, 348(7), 520–530.
- Hamon, Y., Rohais, S., Deschamps, R., and Gasparrini, M. (2012). Outcrop analogue of pre-salt microbial series from south Atlantic: the Yacoraite Fm, Salta rift system (NW Argentina). In *AAPG Hedberg Conference “Microbial Carbonate Reservoir Characterization”, Houston, USA*. <http://www.searchanddiscovery.com/abstracts/html/2011/annual/abstracts/Rohais.html>.
- Hamon, Y., Rohais, S., Deschamps, R., Gasparrini, M., and Barbier, M. (2017). Characterization and distribution of the microbialites within lacustrine series: the example of the Maastrichtian-Danian Yacoraite Fm, Salta Basin, Argentina. In *33rd IAS and 16th ASF Joint Meeting, Toulouse, France*.
- Janson, X. and Madriz, D. D. (2012). Geomodelling of carbonate mounds using two-point and multi-point statistics. In Garland, J., Neilson, J., Laubach, S. E., and Whidden, K., editors, *Advances in Carbonate Exploration and Reservoir Analysis*, volume 370, pages 277–293. Geological Society, London. Special Publications.
- Johnson, J. and Grotzinger, J. P. (2006). Affect of sedimentation on Stromatolite reef growth and morphology, Ediacaran Omkyk Member (Nama Group), Namibia. *South Afr. J. Geol.*, 109(1–2), 87–96.
- Journel, A. G. and Isaaks, E. H. (1984). Conditional indicator simulation: application to a Saskatchewan uranium deposit. *J. Math. Geol.*, 16(7), 685–718.
- Jung, A., Aigner, T., Palermo, D., Nardon, S., and Pontiggia, M. (2012). A new workflow for carbonate reservoir modelling based on MPS: shoal bodies in outcrop analogues (Triassic, SW Germany). In Garland, J., Neilson, J., Laubach, S. E., and Whidden, K., editors, *Advances in Carbonate Exploration and Reservoir Analysis*, volume 370, pages 277–293. Geological Society, London. Special Publications.
- Kirkham, A. and Tucker, M. E. (2018). Thrombolites, spherulites and fibrous crusts (Holkerman, Purbeckian, Aptian): context, fabrics and origins. *Sediment. Geol.*, 374, 69–84.
- Kozłowski, E. N., Burgess, P., Bosence, D. W. J., and Gallois, A. (2014). Stratigraphic forward modelling of Lacustrine carbonates in a tectonically active basin. In *AAPG International*



- Conference & Exhibition, Istanbul, Turkey.* <http://www.searchanddiscovery.com/abstracts/html/2014/90194ice/abstracts/1948616.html>.
- Lapponi, F., Casini, G., Sharp, I., Blendinger, W., Fernández, N., Romaine, I., and Hunt, D. (2011). From outcrop to 3D modelling: a case study of a dolomitized carbonate reservoir, Zagros Mountains, Iran. *Petrol. Geosci.*, 17, 283–307.
- Loucks, R. G. (2018). Domal, thrombolitic, microbialite biostromes and associated lithofacies in the Upper Albian Devils River Trend along the northern, high-energy margin of the Maverick Basin. *Sediment. Geol.*, 371, 75–88.
- Madani, N. and Emery, X. (2015). Plurigaussian modeling of geological domains based on the truncation of non-stationary Gaussian random fields. *Stochast. Environ. Res. Risk Assess.*, 31(4), 893–913.
- Magalhães, A. J. C., Raja Gabaglia, G. P., Fragoso, D. G. C., Bento Freire, E., Lykawka, R., Arregui, C. D., Silveira, M. M. L., Carpio, K. M. T., De Gasperi, A., Pedrinha, S., Artagão, V. M., Terra, G. J. S., Bunevich, R. B., Roemers-Oliveira, E., Gomes, J. P., Hernández, J. I., Hernández, R. M., and Bruhn, C. H. L. (2020). High-resolution sequence stratigraphy applied to reservoir zonation and characterisation, and its impact on production performance - shallow marine, fluvial downstream, and lacustrine carbonate settings. *Earth Sci. Rev.*, 210, article no. 103325.
- Marquillas, R. A. (1985). *Estratigrafía, sedimentología y paleoambientes de la Formación Yacoraite (Cretácico superior) en el tramo austral de la cuenca, Norte Argentino*. PhD thesis, Universidad Nacional de Salta, Salta. 139 pages.
- Marquillas, R. A., del Papa, C. E., and Sabino, I. F. (2005). Sedimentary aspects and paleoenvironmental evolution of a rift basin: Salta group (Cretaceous–Paleogene), northwestern Argentina. *Int. J. Earth Sci.*, 94(1), 94–113.
- Marquillas, R. A., Sabino, I. F., Sial, A. N., Del Papa, C., Ferreira, V., and Matthews, S. (2007). Carbon and oxygen isotopes of Maastrichtian–Danian shallow marine carbonates: Yacoraite Formation, northwestern Argentina. *J. South Am. Earth Sci.*, 23, 304–320.
- Matheron, G., Beucher, H., de Fouquet, C., Galli, A., Guerillot, D., and Ravenne, C. (1987). Conditional simulation of the geometry of fluvio deltaic reservoirs. In *SPE 62nd Annual Conference, Dallas (TX), USA*.
- Michael, H. A., Li, H., Boucher, A., Sun, T., Caers, J., and Gorelick, S. M. (2010). Combining geologic-process models and geostatistics for conditional simulation of 3-D subsurface heterogeneity. *Water Resour. Res.*, 46, 1532–1535.
- Montano, D., Gasparrini, M., Rohais, S., Albert, R., and Gerdes, A. (2022). Depositional age models in lacustrine systems from zircon and carbonate U–Pb geochronology. *Sedimentology*, 69, 2507–2534.
- Moroni, A. M. (1982). Correlación palinológica en las Formaciones Olmedo y Yacoraite. In *Cuenca del Noroeste Argentino. 3° Congreso Geológico Chileno (Concepción, 1982), Proceedings*, pages 340–349.
- Muniz, M. C. and Bosence, D. W. J. (2015). Pre-salt microbialites from the Campos Basin (offshore Brazil): image log facies, facies model and cyclicity in lacustrine carbonates. In Bosence, D. W. J., Gibbons, K. A., Le Heron, D. P., Morgan, W. A., Pritchard, T., and Vining, B. A., editors, *Microbial Carbonates in Space and Time: Implications for Global Exploration and Production*, volume 418. Geological Society, London. Special Publications.
- Petrovic, A., Aigner, T., and Pontiggia, M. (2018). Facies heterogeneities in a ramp carbonate reservoir analogue: a new high-resolution approach for 3D facies modelling. *J. Pet. Geol.*, 41(2), 155–174.
- Pontiggia, M., Ortenzi, A., and Ruvo, L. (2010). New integrated approach for diagenesis characterization and simulation. In *SPE 127236, SPE North Africa Technical Conference & Exhibition, Cairo, Egypt*.
- Renard, D., Beucher, H., and Doligez, B. (2008). Heterotopic bi-categorical variables in PluriGaussian truncated simulations. In *Geostats 2008, Santiago, Chile*.
- Reyes, F. C. (1972). Correlaciones en el Cretácico de la cuenca andina de Bolivia, Perú y Chile. *Rev. Téc. YPF*, 1, 101–144.
- Riding, R. (2000). Microbial carbonates: the geological record of calcified bacterial-algal mats and biofilms. *Sedimentology*, 47(1), 179–201.
- Riding, R. (2002). Structure and composition of organic reefs and carbonate mud mounds: concepts and categories. *Earth Sci. Rev.*, 58(1–2), 163–231.
- Riding, R. (2008). Abiogenic, microbial and hybrid authigenic carbonate crusts: components of Precambrian stromatolites. *Geol. Croat.*, 61(2–3), 73–103.

- Roduit, N. (2007). *JMicroVision: un logiciel d'analyse d'images pétrographiques polyvalent*. Thèse de doctorat, Université de Genève. 3830, 129 pages. Available at <https://doi.org/10.13097/archive-ouverte/unige:468>.
- Rohais, S., Hamon, Y., and Deschamps, R. (2011). Outcrop analogue of pre-salt microbial series from south Atlantic: the Yacoraite Fm, Salta rift system (NW Argentina). In *AAPG Annual Convention & Exhibition, Houston, Texas, USA*. <http://www.searchanddiscovery.com/abstracts/html/2011/annual/abstracts/Rohais.html>.
- Rohais, S., Hamon, Y., Deschamps, R., Beaumont, V., Gasparrini, M., Pillot, D., and Romero-Sarmiento, M.-E. (2019). Patterns of organic carbon enrichment in a lacustrine system across the K-T boundary: insight from a multi-proxy analysis of the Yacoraite Formation, Salta rift basin, Argentina. *Int. J. Coal Geol.*, 210, article no. 103208. ISSN, 0166-5162.
- Rohais, S., Hamon, Y., Deschamps, R., and Gasparrini, M. (2012). Stratigraphic architecture of pre-salt microbial series deposited during the sag phase of the Salta rift system: the Yacoraite Fm (NW Argentina). In *AAPG Annual Convention and Exhibition, Long Beach, USA*. <http://www.searchanddiscovery.com/abstracts/html/2011/ice/abstracts/abstracts367.html>.
- Salfity, J. A. (1982). Evolución paleogeográfica del Grupo Salta (Cretácico-Eogénico), Argentina. In *Congreso Latinoamericano de Geología, 5th, Buenos Aires, Actas, 1*, pages 11–26.
- Suarez-Gonzalez, P., Benito, M. I., Quijada, I. E., Mas, R., and Campos-Soto, S. (2019). 'Trapping and binding': a review of the factors controlling the development of fossil agglutinated microbialites and their distribution in space and time. *Earth Sci. Rev.*, 194, 182–215.
- Viramonte, J. G., Kay, S. M., Becchio, R., Escayola, M., and Novitski, I. (1999). Cretaceous rift related magmatism in central-western South America. *J. South Am. Earth Sci.*, 12(2), 109–121.
- Wright, V. P. (2012). Lacustrine carbonates in rift settings: the interaction of volcanic and microbial processes on carbonate deposition. In Garland, J., Neilson, J., Laubach, S. E., and Whidden, K., editors, *Advances in Carbonate Exploration and Reservoir Analysis*, volume 370, pages 39–47. Geological Society, London. Special Publications.
- Wright, V. P. and Barnett, A. J. (2015). An abiotic model for the development of textures in some south Atlantic early Cretaceous lacustrine carbonates. In Bosence, D. W. J., Gibbons, K., Le Heron, D. P., Pritchard, T., and Vining, B., editors, *Microbial Carbonates in Space and Time: Implications for Global Exploration and Production*, volume 418, pages 209–219. Geological Society, London. Special Publications.
- Xi, H., Burgess, P. M., Kozłowski, E., Hunt, D. W., Jurkiw, A., and Masiero, I. (2022). Spatial self-organization of marine agglutinated microbial carbonate build-ups: insights from stratigraphic forward modelling using Stomatobyte3D. *Sediment. Geol.*, 429, article no. 106081.

# **Geology**





---

Research article

Geo-hydrological Data & Models

# Reflections on the role of chemical osmosis mechanisms on the long-term behavior of a collapsed salt cavity

Emmanuel Ledoux <sup>\*,a</sup>, Emmanuel Hertz <sup>b</sup>, Jean-Charles Robinet <sup>c</sup> and Pierre Combes <sup>d</sup>

<sup>a</sup> MINES Paristech, PSL-Research University, Center of Geosciences, Fontainebleau, 06 07 70 30 68, France

<sup>b</sup> Mining Technical Center, Salins Group, 17 rue Gabriel Péri, 54100 Varangéville, 03 83 18 73 57, France

<sup>c</sup> National Agency for Radioactive Waste Management (Andra), Research and Development Department, 1-7 rue Jean Monnet 92298 Châtenay-Malabry, 01 46 11 81 30, France

<sup>d</sup> Armines, MINES Paristech, PSL-Research University, Center of Geosciences, Fontainebleau, 06 21 60 26 42, France

*E-mails:* emledoux@orange.fr (E. Ledoux), emmanuel.hertz@sfr.fr (E. Hertz), jean-charles.robinet@andra.fr (J.-C. Robinet), pierre.combes34@gmail.com (P. Combes)

**Abstract.** This article focuses on the mechanisms likely to generate salt flux towards the aquifer environment from a cavity created by dissolution within a layer of salt under clay cover, until causing its collapse. This problem is encountered by salt operators, particularly in Lorraine, who exploit a salt deposit located about 200 m deep isolated from an overlying aquifer by an argillaceous layer 130 m thick.

The phenomenon usually invoked to quantify a potential salt emission is the molecular diffusion along the chimney filled by disturbed materials, which are put in place above the cavity following its collapse. Recent work on chemical osmosis in clay media, motivated by research on the geological storage of nuclear waste in a clay layer, raises questions about the role of this mechanism on the long-term behavior of the system.

Based on a bibliographical approach, the article presents a brisk synthesis of the theory of chemical osmosis and discusses its application to the case of clayey rocks which can constitute imperfect membranes capable of generating osmotic pressures.

First, the article presents the equations representing the flow of the fluid and the solute transport according to a simplified formulation adapted to the geometry of the system. It then proposes a set of plausible parameters making it possible to carry out a steady-state modeling of the transfers along the collapsed chimney by means of a numerical resolution in finite differences. Under the conditions representative of the exploitation of the Lorraine salt deposit, the results indicate that, depending on the case, overpressures or downward flows of fresh water may appear. A brine and dissolved salt balance is established and proves to be highly dependent on the osmotic efficiency coefficient, a parameter that is difficult to assess in natural environments. However, the fluxes remain low, indicating that the chemical osmotic effect is unlikely to significantly modify the orders of magnitude

---

\* Corresponding author.

of the salt fluxes evacuated into the environment following intensive exploitation of a salt layer altering the protective clay series.

**Keywords.** Exploitation of salt by dissolution, Salt cavity, Collapse, Chemical osmosis, Salt pollution.

*Manuscript received 31 March 2022, revised 28 July 2022, accepted 6 September 2022.*

## La version française de l'article est disponible à la suite de la version anglaise

### 1. Position of the problem

The exploitation of rock salt in layers or in domes by dissolution *in situ* is the subject of an ancient industry developed since the middle of the 19th century. This is the case of the Lorraine salt deposit, still widely in use today for the production of salt intended for chemicals, food or even snow removal. Among the current operators is the Salins group, which operates near the towns of Dombasle and Varangéville, to the east of Nancy. The layers of salt used are part of a powerful marl-salt ensemble of the Keuper. Two methods of exploitation are practiced, one, by extracting rock salt by mined cavities according to the method of abandoned rooms and pillars for snow removal (mine of Varangéville) and, the other, by pumping the brine obtained by leaching with fresh water of deep cavities, for industry (salines of the Haraucourt plateau).

In order to control the leaching, two questions arise among stakeholders; optimizing the recovery of salt, which is a valuable commodity, and preserving the environment from any pollution by brine and mechanical damage in surface. To this end, with regard to the Salins group, the leaching method currently aims to produce cavities isolated from each other, the long-term stability of which can be demonstrated on the hydraulic and mechanical levels. However, certain sectors of the deposit were utilised by means of communicating cavities, which compromised the control of spontaneous residual dissolutions at the end of their life and led, in certain cases, to collapses reaching the surface, creating craters filled with water.

In order to make the site safe, while optimizing salt production, the operator plans to resume the *in situ* dissolution of a panel on the Gellenoncourt site, on the Drouville concession. The so-called intensive mining method used will result in the enlargement, followed by a collapse, of the already existing cavities, ultimately resulting in a lake extending over approximately 36 ha. With the aim of obtaining authorization to implement the method, the operator submitted an impact file examining, on a scientific basis, the consequences of such a practice, in particular with regard to the risk of pollution of the surrounding aquifers with salt water.

This article discusses the potential mechanisms of salt release from collapsed cavities. It takes advantage of the research work of the French National Agency for Radioactive Waste (Andra), which is interested in understanding the interstitial overpressures observed in the Callovo-Oxfordian argillite layer of Bure in Meuse/Haute-Marne, subject of studies for the creation of a geological repository for high and intermediate long-lived radioactive waste. This work carried out in partnership with the Laboratory of Solid Mechanics of the Ecole Polytechnique and the

UMR Sisyphe at the Pierre and Marie Curie University was the subject, in 2008, of a thesis defended in this university, under the presidency of Professor Ghislain de Marsily. We have structured the theoretical bases of our reasoning from this document, as well as from several resulting publications.

### 1.1. Elements of the geology of the Lorraine salt deposit

The numerous boreholes carried out on the salt concessions make it possible to propose the synthetic geological cross-section which shows from top to bottom the following sequence (Figure 1).

*Black clays with Promicroceras* (Lotharingian). Grey-blue plastic shales, sometimes schistose, with a maximum thickness of 25 m.

*Limestones with Gryphea* (Sinemurian). Decimetric alternation of blue to black marl and beds of rust-coloured clayey limestone, 15–17 m thick.

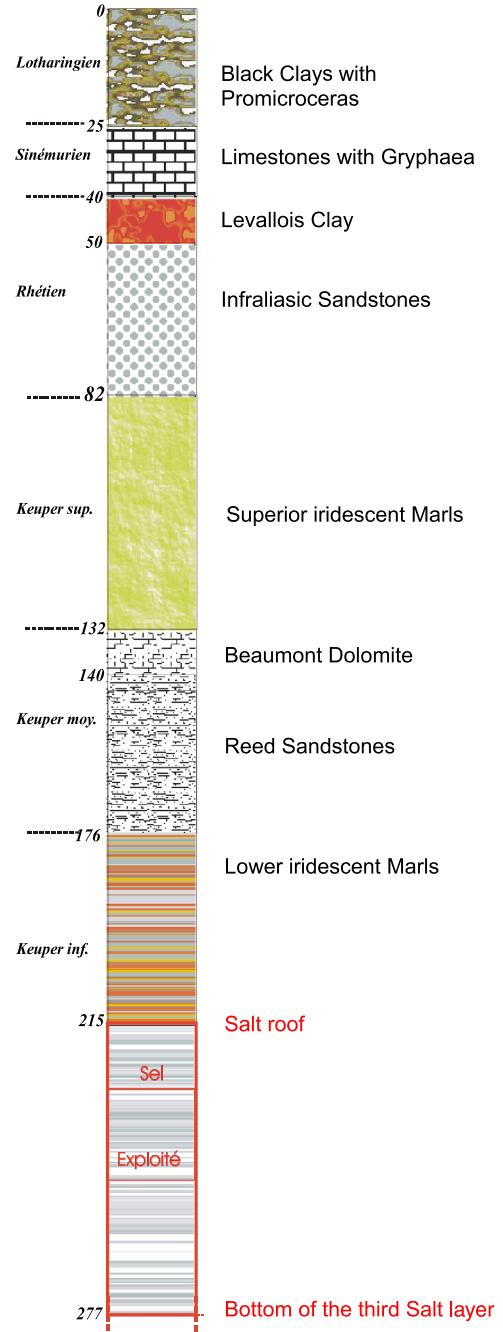
*Levallois clays* (Rhaetian). Salmon pink argillites which constitute a reference level throughout Lorraine. The average thickness is about ten meters.

*Infraliasic sandstone* (Rhaetian). Fine siliceous sandstone with centimetric to decimetric black or green clayey intercalations in the upper part. The average thickness is 32 m.

*Superior iridescent marls* (Upper Keuper). Green and purplish argillites which become brick red towards the base. Centimetric to decimetric dolomitic horizons are present as well as, at the base, decimetric levels of nodular white anhydrite. The thickness reaches 135 m at Gellenoncourt.

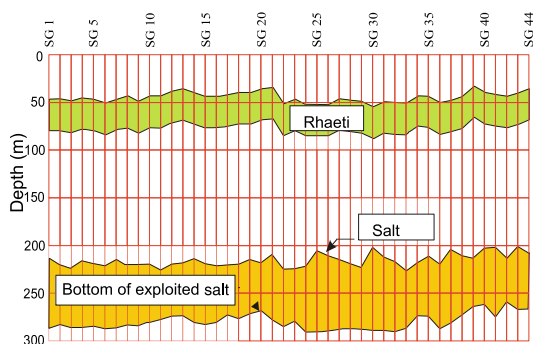
*Beaumont dolomite—Reed sandstone* (Middle Keuper). Beige or gray dolomite in decimetric beds surmounted by a massive anhydrite cap 3 m–4 m thick. The dolomite, 8 m thick, plays an important role in the stability of the land before the collapses. Below there are motley marls then very clayey fine micaceous sandstones containing numerous plant debris.

*Lower iridescent marls* (Lower Keuper). This powerful formation about 200 m thick is composed of argillites alternating with beds of gypsum and anhydrite on the first forty meters. Then comes the salt series 120 m thick but of which only the 3 upper beds, rich in salt and 65 m thick, are mined at Gellenoncourt and contain about 20% of insoluble materials. Thus, the cumulative thickness of leachable pure salt represents on average about forty meters.



**Figure 1.** Synthetic geological log.

The regularity of the layers appears in Figure 2 which shows the position of the reference levels, the Rhaetian and the salt roof, which are present respectively at an average of 50 m and 220 m below ground in the 44 boreholes available on the Gellenoncourt



**Figure 2.** Geological cross-section showing the position of Rhaetian sandstones and utilised salt.

site.

### 1.2. Elements of site hydrogeology

Among the geological formations present in the subsoil, only the Rhaetian sandstones constitute an aquifer likely to be impacted by salt mining.

This thirty meters thick aquifer contributed to the supply of drinking water to neighboring municipalities. It is still weakly utilised for irrigation. Protected by the low permeability and very continuous cover of the Levallois clays, this reservoir is fed on its outcrops at the southern and eastern edges of the plateau. The transmissivity is modest and varies between  $10^{-4}$  m<sup>2</sup>/s and  $10^{-5}$  m<sup>2</sup>/s, which leads to great sensitivity of the piezometric level to pumping.

The aquifer is most often captive under the Levallois clays and is artesian in the lower parts of the topography. The general piezometry indicates a flow towards the northwest to the Roanne river which drains this aquifer.

### 1.3. The consequences of salt mining

The exploitation of the Gellenoncourt panel was carried out for about twenty years from 1967. The method consisted in connecting by hydraulic fracturing at the bottom of several boreholes drilled to the basis of the third salt bed, then introducing fresh water in one or more wells and pumping the resulting brine in remote boreholes called “extractors”, thus causing the salt to dissolve and create cavities in the vicinity of the injectors. A good match between the

flow rate and the arrangement of the device ensured the saturation of the brine. Periodic controls of the geometry of the voids were carried out by ultrasound in order to monitor the evolution of the cavities.

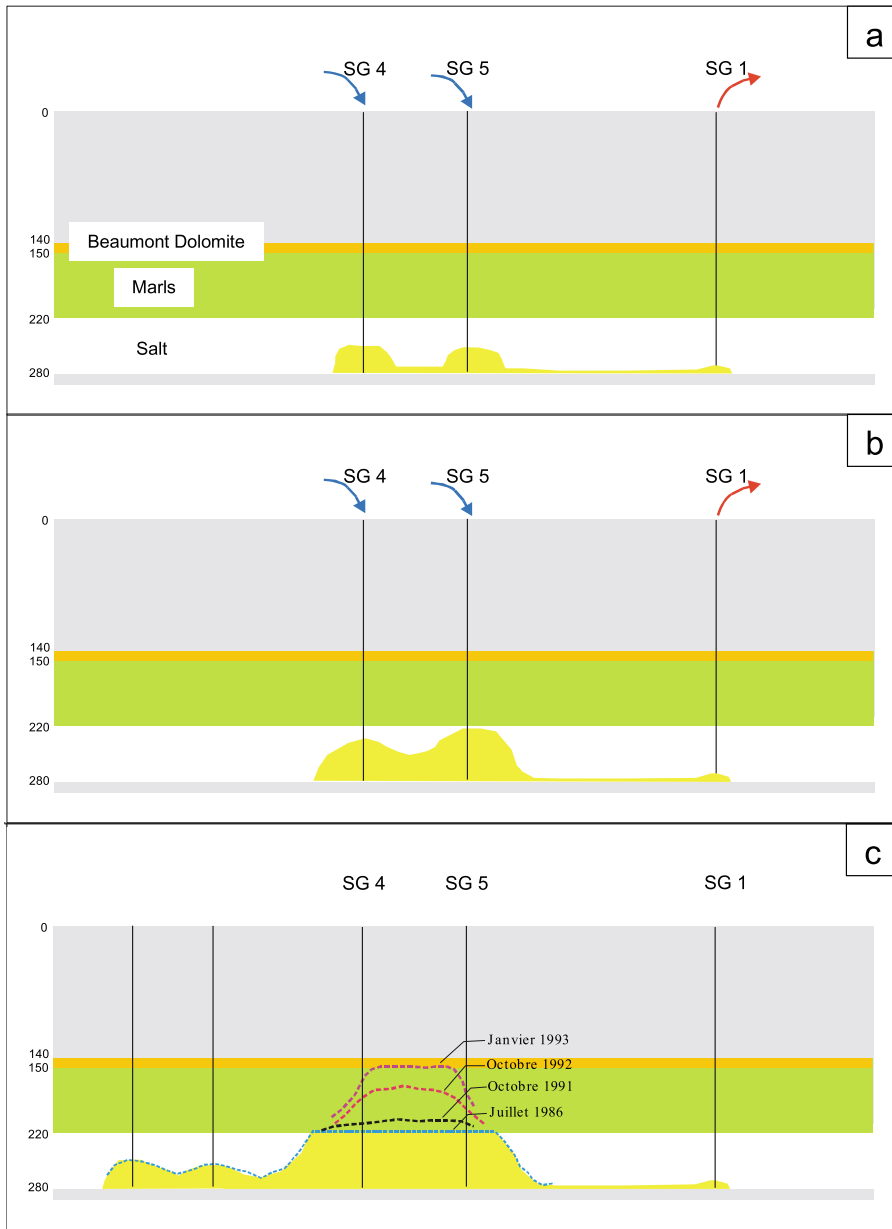
Thus, the extraction of salt has developed voids whose base is around 65 m deep under the roof of the salt and whose morphology evolves in time and space as the production of salt progresses. The precise geometry of the interconnection network of the cavities is poorly known; however, the ultrasonic measurements taken in the boreholes clearly indicate the extension of the cavities created at the level of the injectors.

In normal evolution, all of these voids remain confined within the saliferous formation. However, this evolution was disturbed in the past by the collapse of two central cavities of the Gellenoncourt device, SG4 and SG5 (Figure 3) according to the history described below.

Wells SG4 and SG5 were used as freshwater injectors from the start of operations, SG1 being the brine extractor (Figure 3a). The device was maintained until the salt board at the top of SG5 was found to have reduced to 2 m, (Figure 3b) after which the injection of fresh water was moved to further points. A survey of the cavities subsequently showed that their volume continued to increase and the abandonment of the entire panel was decided. From 1982 to 1993, however, checks revealed a rise in the roof of SG5 in the argillites covering the salt, at first slowly, then rapidly until it came into contact with the Beaumont Dolomite (Figure 3c). The operator then took the decision to cause the collapse, by deliberately lowering the level of brine to maintain an air cushion at the top of the cavity. The event finally occurred in March 1998, leading to the formation of a surface crater about 50 m in diameter and 120 m deep [Buffet, 1998]. The chimney was then partially filled due to sloping of the edges of the crater. The crater was initially dry after the collapse; it then filled with water and ended up as a lake whose level was about 30 m from the surface in June 1998.

This collapse locally disrupted the protective cover of the saline deposits and placed various hydraulic units in potential communication: the crater lake, the Rhaetian aquifer and the mining cavities filled with brine and dissolution residues. Given the lithological diversity of the overburden, the materials that fill the collapse chimney are heterogeneous with





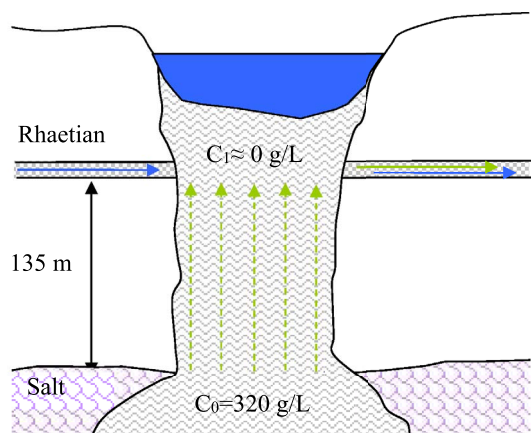
**Figure 3.** Evolution of the SG4 and SG5 cavities up to the Beaumont Dolomite.

a dominant clay content. Their properties regulate the exchanges of fluid and dissolved salt between these units.

This experience motivated the Salins group to consider resuming the leaching of the entire panel until the still stable cavities collapsed, with the dual

aim of increasing salt production but above all of securing the site in the long term.

Currently the lake has an area of about 1.7 ha; it should extend to 36 ha upon completion after resumption of operations. Once all the collapses have been attained, the system will evolve towards a deep



**Figure 4.** Conceptual diagram of the diffusion of salt towards the Rhaetian.

saline aquifer and a lake on the surface hydraulically connected by the low permeability saturated rock materials filling a chimney. The saline aquifer will be confined within the saliferous formation, thanks to the high density of the brine which will prevent any convective circulation despite the presence of a thermal gradient, insofar as the lake will constitute a single body of water not generating hydraulic gradient in the system.

## 2. Long-term behavior review

### 2.1. Diffusion, a potential mechanism of salt emission

We consider here the case of the “post-mining” situation, in the long term, when the configuration of the site will be stabilized. The origin of the salt that can pollute the Rhaetian will be the saturated brine present in the cavities. This dense brine will not be able to circulate by convection but it could be a source of dissolved salt moving by diffusion through the materials of the collapsing chimney towards the fresher waters of the lake. Diffusion flux can be assessed using the following simple model (Figure 4).

The chimney materials are assumed to be homogeneous with an average effective diffusion coefficient  $D$  for the different solutes. A concentration boundary condition at the  $C_0$  value of 320 g/L is imposed by the saturated brine at the base of the chimney where it penetrates the salt layer, while a low concentration of  $C_1$  value is maintained at the Rhaetian

level under the effect of groundwater flow which dilutes the salt. Under these conditions the diffusive flux density in steady state is given by Fick's law:

$$F = -D \text{grad}(C) = -D(C_1 - C_0)/H,$$

where  $H$  is the vertical distance separating the saline layer from the base of the Rhaetian.

The effective diffusion coefficient in the materials filling the chimney is limited by the diffusion coefficient in the free water, the value of which is around  $10^{-9}$ – $2 \times 10^{-9}$  m<sup>2</sup>/s. We have considered a reduction of a factor of 10 due to the presence of solids which hinder diffusion and we have adopted the value of  $2 \times 10^{-10}$  m<sup>2</sup>/s for the calculation of the surface flux of salt:

$$F = 2 \times 10^{-10} \times (320 - 0)/135 \approx 4.7 \times 10^{-10} \text{ kg/m}^2/\text{s}$$

using the numerical values in Figure 4.

For a collapsed chimney of 36 ha, an annual flow of 5–6 tons of salt is obtained. This diffusive flow is discharged into the Rhaetian aquifer under the effect of groundwater flow.

### 2.2. Chemical osmosis, another mechanism

The Regional Mission of the Environmental Authority which assessed the file presented by the Salins group gave an opinion in October 2018, at the request of the Prefect of Meurthe et Moselle and pointed out that osmotic phenomena were not considered and could cause a potential risk to long-term stability.

This article sets out some thoughts on the nature of this phenomenon and on the role it could possibly play in the configuration of the operation envisaged by the Salins.

It is made, initially, by applying the theoretical bases of chemical osmosis. We are then interested in the way in which the theory can be used in the particular context of salt mining, in order to identify orders of magnitude of the consequences that can be expected from the phenomenon. Finally, a confrontation with field observations will provide elements of validation of the approach adopted.

## 3. The theoretical bases of chemical osmosis

Chemical osmosis or chemo-osmosis designates the movement of water occurring through a semi-permeable membrane whose faces are bathed in

ionic solutions of different concentrations. Such a membrane is said to be perfectly semi-permeable when it allows the passage of water molecules but totally opposes the transfer of ions. The water thus crosses the membrane from the zone where the concentration is low towards the zone of high concentration so as to cause a reduction in the concentration gradient by dilution effect.

In practice, membranes are only imperfectly semi-permeable and allow ions to pass at variable speeds depending on their properties; this type of transfer during which the ions of the solution pass from the zone of high concentration to the zone of low concentration is designated by diffusion. Thus, for the two types of mechanisms, osmosis and diffusion, the effect is to tend towards a homogenization of the concentrations on either side of the membrane.

The diagram of Figure 5 [Rousseau-Gueutin, 2008], describes these phenomena in the cases of a perfect and imperfect membrane.

The diagram highlights a pressure difference between the two branches of the U-tube, called osmotic pressure. In the case of a perfect membrane, the pressure difference linked to the chemical activity generates a flow, in accordance with Darcy's law, counter-current to the osmotic flow and the system equilibrates when the flows of water osmotic and darcean are equal for the osmotic pressure corresponding to the difference in residual concentration between the two branches of the tube. In the case of an imperfect membrane, the presence of osmotic pressure is only transitory because the darcean flow and the diffusion tend to homogenize the concentration; the final state then translates into equal pressures and concentrations in the two branches. The duration of this transient phase depends on the permeability of the membrane.

It is therefore noted that the final state of the system is particularly dependent on the properties of the membrane. It also depends on the behavior of the ion source and will be demonstrated when applied to a salt deposit.

### 3.1. *The thermodynamic model*

The thermodynamics of irreversible phenomena provides a theoretical framework for modeling multi-physics coupling phenomena at the macroscopic scale. This theory postulates that, in general,

the flux of a quantity is linked in a tensorial way to a driving force via a matrix of phenomenological coefficients. The different driving forces are expressed, themselves, in the form of the gradient of a scalar field corresponding to the different state variables of the system.

Table 1 brings together the different forces and fluxes taken into account in the problems of fluid mechanics concerned with chemical, thermal and electrical phenomena.

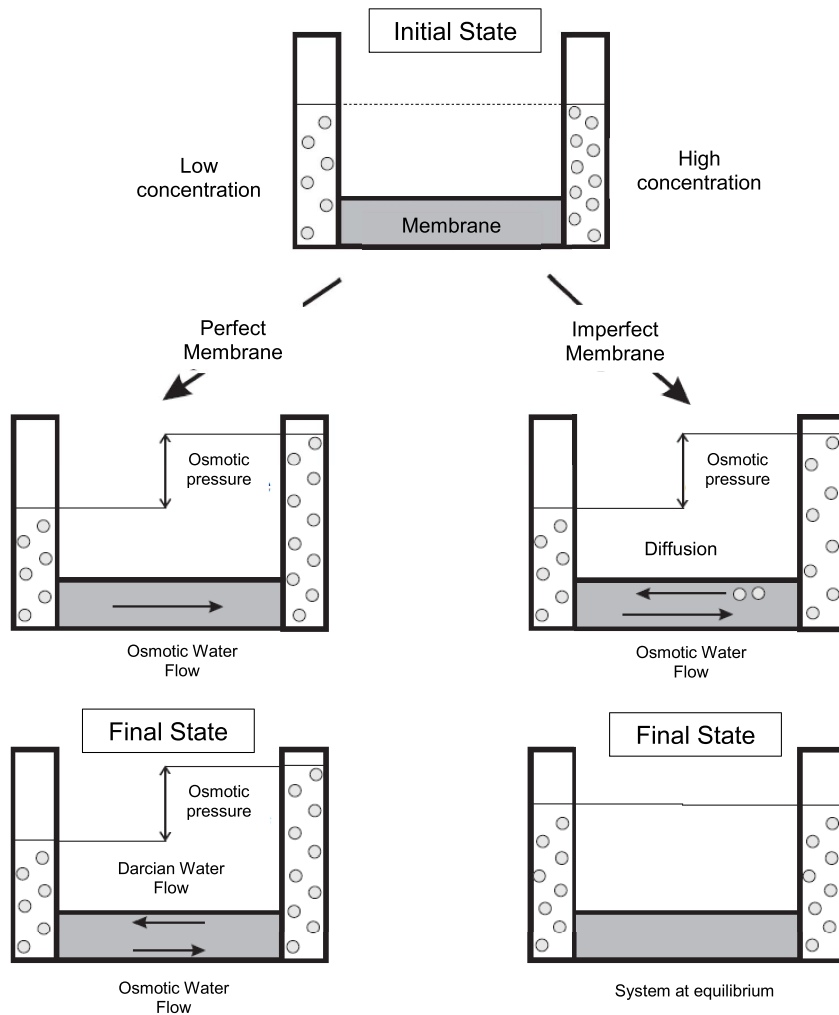
The diagonal coupled mechanisms are said to be dominant and correspond to the phenomena classically taken into account in physics. The non-diagonal mechanisms refer to the non-dominant transfers which only play a significant role in particular configurations; chemical osmosis is one of them.

The phenomenological coefficients linking each flux to each force are not predicted by the macroscopic theory. They can be estimated by changing scale from other models considering the microscopic phenomena or, most often, empirically by means of experiments carried out at the right scale [Gonçalvès and Trémosa, 2010].

### 3.2. *The case of clay materials*

According to Gonçalvès *et al.* [2015], as early as the 1960s, experimenters working on porous media had highlighted deviations from Darcy's law in clay materials, one of the main reasons given was the existence of an osmotic effect in the presence of saline fluids. Knowledge on the subject has developed since the 1990s, in particular, thanks to studies on the storage of radioactive waste in low-permeability clay formations. It should be noted that these studies essentially concerned artificially prepared or reworked materials, on the scale of the laboratory sample; very few involved natural materials and even fewer in *in situ* conditions. On the scale of a geological layer 130 m thick at a depth of approximately 500 m, Andra revealed overpressures in the Callovo-Oxfordian argillites in Meuse/Haute-Marne. The role of chemical osmosis has been assessed and it appears, without being definitively demonstrated, that some of these overpressures could be due to osmotic phenomena [Rousseau-Gueutin, 2008, Gonçalvès *et al.*, 2015, Gueutin *et al.*, 2007].

The origin of the osmotic effect in clays is explained by the fact that the surface of clay minerals,



**Figure 5.** Illustration of diffusion and chemical osmosis in a U-tube [from Rousseau-Gueutin, 2008].

**Table 1.** Coupled mechanisms linking driving forces and fluxes

Fluxes	Driving forces			
	Gradient of head or pressure	Gradient of electric potential	Gradient of concentration	Gradient of temperature
Flow velocity	Darcy's law	Electro-osmosis	Chemical osmosis	Thermal osmosis
Electric current density	Filtration current	Ohm's law	Diffusion current	Seebeck effect
Solute flux	Reverse osmosis	Electrophoresis	Fick's law	Soret effect
Heat flux	Heat filtration	Peltier effect	Dufour effect	Fourier's law

which is negatively charged, imposes electrical restrictions on the ions during their transport, resulting in a concentration of cations near the surface and an exclusion of anions towards the center of the pores.

To respect electrical neutrality, the restriction on the transport of the anions affects the cations, which limits the total ionic transport like what an imperfect membrane would do. Surface charges are more eas-

ily compensated when the ionic strength of the solution is high and the effect is all the more important as the thickness of the electrostatic layers approaches the size of the pores. It is thus understood that the efficiency of the osmotic mechanism is linked to the nature of the clay minerals, to the size of the pores as well as to the concentration of the solutions. Clayey rocks with low permeability, such as Callovo-Oxfordian argillites, are characterized by pore dimensions of a few nanometers comparable to those of electrostatic layers, which explains the occurrence of osmotic phenomena in these systems.

In summary, the existence of the osmotic effect in clay materials is established; the question is to know what its intensity may be in the configuration that concerns us.

### 3.3. Laws and parameters

Goncalvès *et al.* [2015] provided a very detailed synthesis of couplings in membranes and clay materials. It emerges from this synthesis that the phenomenological relations representing the coupling between the flow obeying Darcy's law, the osmotic effect and the transport by convection-diffusion at the macroscopic scale are written as follows, in steady state:

$$\begin{aligned}
 V &= -\frac{k}{\mu}(\text{grad}(p) + \rho g \text{grad}(z)) \\
 V' &= \frac{k}{\mu}(\varepsilon \text{grad}(\pi)) \\
 \text{div}(V + V') &= 0 \\
 J &= -D \text{grad}(C) + VC \\
 \text{div}(J) &= 0 \\
 \text{grad}(\pi) &= \nu RT \text{grad}(C)
 \end{aligned} \tag{1}$$

where  $V$  (m/s) is the Darcy velocity,  $V'$  (m/s) the osmotic velocity of water,  $J$  (mol/s) the diffusive and convective flux density of solute,  $p$  (Pa) the pressure of the fluid,  $z$  (m) the altitude,  $\pi$  (Pa) the osmotic pressure,  $C$  (mol/m<sup>3</sup>) the solute concentration,  $\rho$  (kg/m<sup>3</sup>) and  $\mu$  (kg/m/s) respectively the density and the dynamic viscosity of the fluid,  $D$  (m<sup>2</sup>/s) the effective diffusion coefficient in a porous medium and  $k$  (m<sup>2</sup>) the intrinsic permeability.

The osmotic pressure  $\pi$  is given by Van't Hoff's formula where  $R$  (J/mol/K) is the ideal gas constant,  $T$  the Kelvin temperature and  $\nu$  the number of separated ions (2 for NaCl).

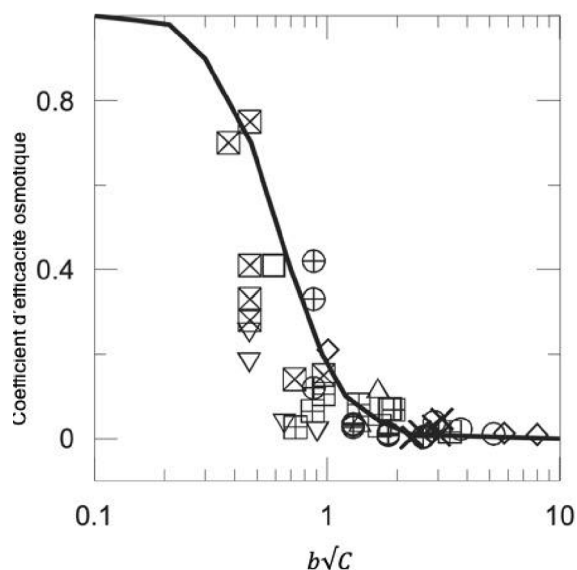
We note the existence of two velocities for water, the classical Darcy velocity  $V$ , generated by the hydraulic head gradient and the osmotic velocity  $V'$  generated by the osmotic pressure, therefore by the concentration gradient. Only the Darcy velocity intervenes on solute transport since the osmotic phenomenon only moves water molecules. The osmotic phenomenon is therefore not likely to cause the displacement of the ions.

An essential parameter is the osmotic efficiency coefficient  $\varepsilon$  (dimensionless) which varies between 0 and 1. In the case of a perfect membrane, this coefficient is equal to 1 and the diffusion coefficient is then zero, reflecting the absence of ionic transport. It is less than 1 for natural environments depending on their more or less semi-permeable character. In such media, the three fundamental parameters are the permeability, the effective diffusion coefficient and the osmotic efficiency coefficient. Their determination is experimental, using systems requiring particularly advanced metrology and their value is influenced by the scale of the problem dealt with.

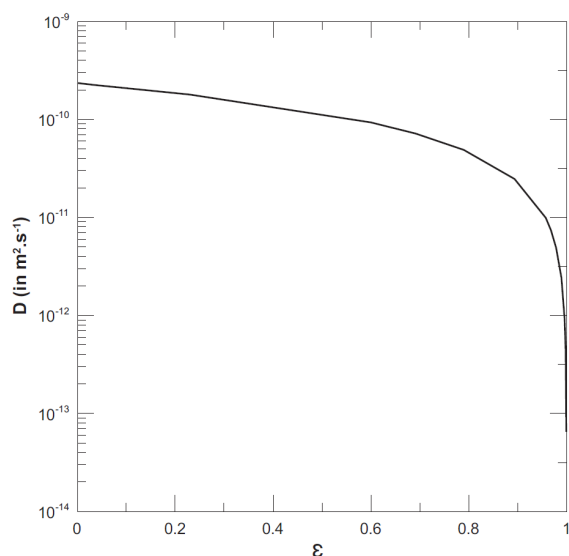
Theoretical considerations [Bresler, 1973] establish that the osmotic efficiency coefficient is a decreasing function of the product of the pore size of the medium times the square root of the concentration. Goncalvès *et al.* [2015] have brought together in a single figure (Figure 6) the measurements available in the literature according to these parameters compared to the theoretical Bresler curve.

These data show that the measured osmotic efficiency coefficient is generally low with values in fairly good agreement with the theoretical curve. Note that the points relating to the very compact Bure argillites are at low values.

The theory also accounts for the dependence of the diffusion coefficient with the osmotic efficiency coefficient. Figure 7 taken from Rousseau-Gueutin [2008], based on the work of Revil *et al.* [2005] and Gonçalvès *et al.* [2007], indicates the shape of this dependence, established for a material composed of a smectite-type clay. We note that it is weak for low values of osmotic efficiency, which suggests that osmosis has little effect on the diffusive transport of ionic solute in media that behave like imperfect membranes.



**Figure 6.** Osmotic efficiency coefficient as a function of  $b\sqrt{C}$ .  $b$  is the average pore size in nm and  $C$  the concentration in mol/L.  $\boxtimes$  represent the experimental values for the Bure argillite [from Goncalvès *et al.*, 2015].



**Figure 7.** Variation of the effective diffusion coefficient as a function of the osmotic efficiency coefficient for a smectite in the presence of an NaCl solution [Rousseau-Gueutin, 2008].

### 3.4. Application to a saline cavity

In order to apply the above to the situation of a collapsed salt cavity, we will retain the following elements:

- it has been proven that chemical osmosis phenomena can occur in clay materials. This is linked to the surface electrical properties of clay minerals and to the small pore size which give to clay media the behavior of an imperfect semi-permeable membrane;
- at the macroscopic scale, the phenomenon is governed by an osmotic efficiency coefficient varying between 0 and 1 which weights the influence of osmotic pressure in the phenomenological relationship which describes the flow of water;
- the osmotic efficiency coefficient is highly variable, of complex and uncertain measurement, and has only rarely been assessed in natural environments;
- the efficiency coefficient depends on the size of the pores and on the concentration, which makes the non-linear phenomenon therefore difficult to quantify by a simple approach because the size of the pores is not a directly accessible parameter;
- the materials filling the collapse chimney of the saline cavities are essentially disturbed clayey materials whose porosity is increased compared to that of an argillite in place, which leads to a relative increase in the average size of the pores. Moreover, the salt concentrations are very high, at least in the lower part of the system. These elements are in favor of a very low value of the coefficient of osmotic efficiency if they are compared with the knowledge acquired on the argillites of the Callovo-Oxfordian of Bure which are considered as well characterized and having pores of very small dimensions. The hypothesis considering that the filling of the chimney is made up of homogeneous clay, leads to increasing the role of osmosis mechanisms, which goes in the direction of safety in the assessment of the danger;
- the osmotic phenomenon does not modify the transport of ionic solute within the

medium, which remains dependent on the convection–diffusion mechanism.

#### 4. Exploratory modeling of the behavior of a collapsed cavity

In order to quantify the potential effect of osmosis on the behavior of a collapsed salt cavity, we attempted a simple modeling based on the conceptual diagram in Figure 4.

##### 4.1. Principles of the model

The assumptions are as follows:

- the transfers are considered to take place in a one-dimensional vertical fashion between the deep collapsed cavity and the Rhaetian aquifer system, through the chimney over a height of 135 m;
- the depth concentration is maintained by dissolving the salt at 320 g/L, i.e. 5480 mol/m<sup>3</sup>. It is imposed at 0 g/L at the Rhaetian level. This fixes the concentration gradient at the value 0.04 mol/L/m;
- the system is represented in steady state, in two distinct configurations of hydraulic boundary conditions illustrated in Figure 8;
  - *configuration 1*: the cavity is in hydraulic relation with other cavities whose brine level is kept constant; this is currently the case at Gellenoncourt, where the collapse of the operated panel is only partial. We will consider a hydrostatic equilibrium between the Rhaetian and the cavity;
  - *configuration 2*: the collapsed cavity is confined, with no possible hydraulic exchange with the external environment; this configuration is meant to represent the final state when the whole panel is collapsed on approximately 36 ha;
- in both configurations the concentration profile at equilibrium in the chimney is considered to be diffusive with a constant gradient between 320 and 0 g/L.

Under these assumptions the system of equations (1) simplifies to (2)

$$\frac{\partial}{\partial z} K \left( -\frac{\partial h}{\partial z} + \frac{\varepsilon \nu RT}{\rho g} \frac{\partial C}{\partial z} \right) = 0 \quad (2)$$

$$D \frac{\partial^2 C}{\partial z^2} = 0$$

with the notations of (1) supplemented by  $h$  (m), hydraulic head of the fluid and  $K$  (m/s), Darcy permeability of the materials filling the chimney.

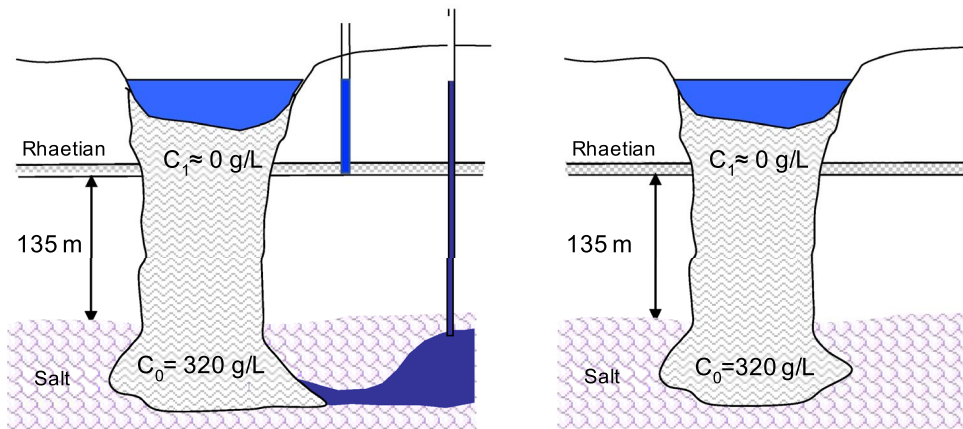
The system of equations (2) is non-linear because the coefficient of osmotic efficiency depends on the concentration. The numerical resolution by finite differences provides the profile of the head in the chimney and allows the calculation of the flow of water and salt flux exchanged between the cavity and the Rhaetian aquifer.

##### 4.2. Choice of parameters

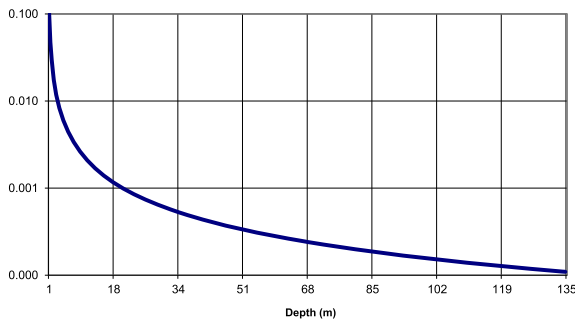
The Darcy permeability  $K$  and the effective diffusion coefficient  $D$  have been fixed at the values 10<sup>-10</sup> m/s and 2 × 10<sup>-10</sup> m<sup>2</sup>/s respectively. Note that if these parameters are constant on the vertical, they have no influence on the calculated profile. They intervene on the other hand in a linear way on the fluxes of water and salt. We consider that the values chosen are in the high range and that the fluxes will thus be assessed on an upward basis.

As indicated previously, the essential parameter of the osmotic phenomenon is the coefficient of osmotic efficiency  $\varepsilon$ . We estimated it from the Bressler curve (cf. Figure 6) considering an average pore size of 10 nm and a linear concentration profile in the chimney between 320 and 0 g/L. This pore size was set arbitrarily by reference to the 4 nm value presented by Rousseau-Gueutin [2008], for the Callovo-Oxfordian argillites of Bure. It is very likely that it is largely underestimated, if one considers that the porosity of the unstructured clay materials of the chimney must be notably increased compared to that of compact argillites. By comparison, surface clays from the Albien in Aube show pore sizes between 50 and 80 nm.

With these hypotheses, deliberately chosen to increase the osmotic effects, the values of the osmotic efficiency coefficient range between 10% in the upper part where the concentration is low and 0.01% in the lower part where the brine reaches saturation



**Figure 8.** Simulation configurations, (on the left connected cavity, on the right isolated cavity).



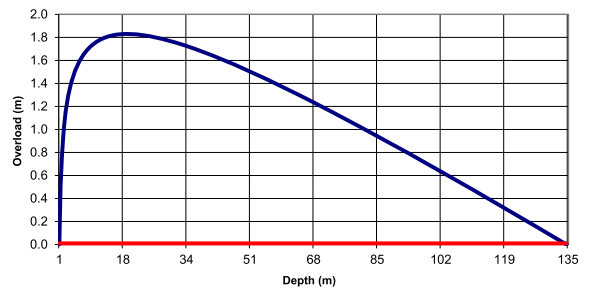
**Figure 9.** Profile of osmotic efficiency coefficient as a function of depth for a pore size of 10 nm.

(Figure 9). In order to obtain an appreciation of the sensitivity to this parameter, a simulation was also carried out by adopting a homogeneous value of 1% along the profile.

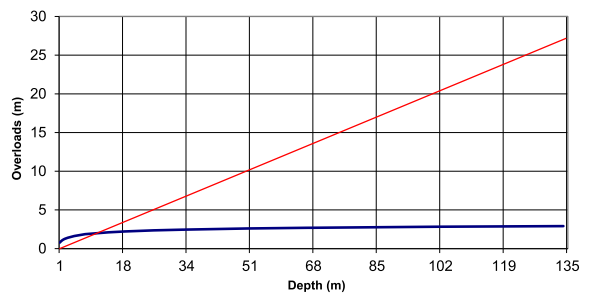
4.3. *Simulation results*

Figures 10 and 11 show the calculated hydraulic over-head profiles for the two model boundary condition configurations. The over-head represents the hydraulic head difference along the profile compared to the Rhaetian head considered as a reference.

*In the case of configuration 1* where the cavity is connected to the aquifer system, a hydraulic head of 0 m is arbitrarily imposed at the top and bottom of the chimney. The model therefore calculates an over-head compared to the hydrostatic head. It



**Figure 10.** Configuration 1, hydraulic over-head profiles as a function of depth calculated with a variable (in blue) and constant (in red) osmotic efficiency coefficient.



**Figure 11.** Configuration 2, hydraulic over-head profiles as a function of depth calculated with a variable (in blue) and constant (in red) osmotic efficiency coefficient.

can be seen that the osmotic effect only deforms the head profile if the osmotic efficiency coefficient is



variable as a function of depth. We thus find the results of Rousseau-Gueutin [2008], and Gonçalves *et al.* [2007]. In both cases, the chimney is traversed by a downward flow of water, the speeds of which are respectively  $2.2 \times 10^{-12}$  m/s and  $2.02 \times 10^{-11}$  m/s. This flow is solely the consequence of the osmotic effect; it is evacuated to the outside by the boundary condition of constant level imposed at the base of the chimney. It should be noted that an interpretation in purely “Darcean” terms of the deformed profile would lead to an erroneous evaluation of the intensities and orientations of the fluxes.

*In the case of configuration 2* where the cavity is totally confined, no water flow can travel through the chimney in steady state, so that the head rises with depth to compensate for the osmotic pressure. The over-head thus reaches 2.9 m or 27 m at the level of the cavity, depending on the osmotic efficiency coefficient profile adopted. Such over-heads would not endanger the mechanical stability of the cavity given the thickness of the covering. Let us recall that the values of osmotic efficiency coefficients are certainly chosen by excess.

#### 4.4. Cavity salt balance

The preceding simulations make it possible to quantify the flux of salt evacuated from the cavity in the different configurations.

This flux has two components:

- a component by diffusion generating a flux towards the top of the chimney. This part is independent of the osmotic effect insofar as the diffusion coefficient is only slightly influenced by this effect for the low values of osmotic efficiency coefficient considered (cf. Figure 7);
- a so-called “convective” component linked to the downward flow of water which reaches the cavity under the effect of osmosis in the case of configuration 1 where the cavity is hydraulically connected to the external environment. Note that this flow is made up of fresh water but that it becomes saturated with salt when it reaches the salt formation. It will be considered that the mass of NaCl dissolved by 1 L of pure water is 355–360 g for a brine saturated at 320 g/L with a density of  $1220 \text{ kg/m}^3$ .

Table 2 indicates these different fluxes for a 36 ha collapse chimney, as should be the case when the Gellenoncourt panel operation ends.

Let us recall the values adopted for the Darcy permeability and the diffusion coefficient, which are respectively  $10^{-10}$  m/s and  $2 \times 10^{-10} \text{ m}^2/\text{s}$ . Salt fluxes are directly proportional to these values.

It will be noted that the diffusive flux of salt is independent of the osmotic effect but that the convective flux is directly linked to this effect with a high sensitivity to the profile of the osmotic efficiency coefficient. The line of the table entitled “volume increase of the cavity” corresponds to the volume of salt dissolved at depth under the effect of the various flows. The brine “overflow volume” represents the quantity of brine that must theoretically be evacuated to cancel the over-head generated in the cavity by the osmotic effect.

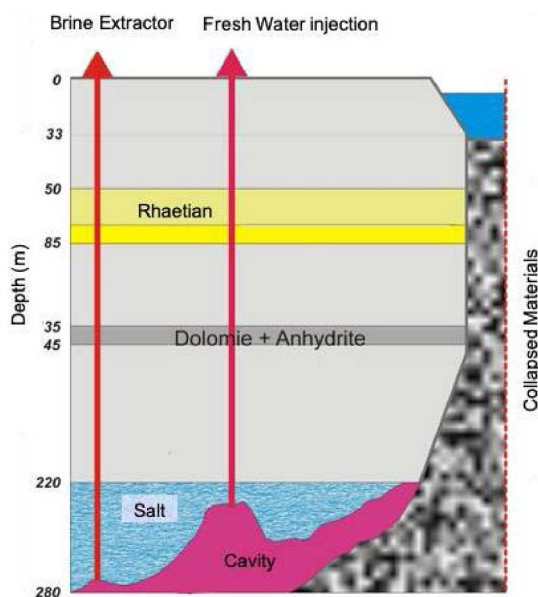
In addition to the diffusion and osmosis processes mentioned above, experience shows that deep saline cavities are subject to the creep of the surrounding saline mass which generates pressure rises or brine overflows depending on whether the cavity is kept closed or open [Brouard *et al.*, 2017]. This mechanism does not appear to play an important role in the case that concerns us because the cavities are shallow and are self-filled by the materials from the collapse chimney. Ground movements are however likely to exist to compensate for the increase in volume generated by the dissolution at depth under the effect of the convective flow of fresh water caused by osmosis. However, these movements would be very slow given the low volumes involved and would primarily concern the depth of the lake.

## 5. On-site observations and conclusion

Figure 12 illustrates the current configuration of the Gellenoncourt panel where the collapse of two coalescing cavities (SG4 and SG5) in 1998 created a chimney opening on the surface into a crater, occupied by a lake whose initial diameter was 50 m, i.e. an area of 0.2 ha. This area then evolved over time to currently reach 1.7 ha under the effect of the natural sloping. The figure shows that the collapsed part of the cavities is in direct hydraulic relationship with other non-collapsed cavities whose brine level is controlled by operating boreholes. This diagram corresponds to configuration 1 of the model developed above. The

**Table 2.** Annual salt balance at the end of a collapsed panel on 36 ha

	Setup 1		Setup 2	
	Variable $\varepsilon$ profile 10% $\rightarrow$ 0.01%	Constant $\varepsilon$ profile 1%	Variable $\varepsilon$ profile 10% $\rightarrow$ 0.01%	Constant $\varepsilon$ profile 1%
Diffusive salt flux (tons/year)	5.4	5.4	5.4	5.4
“Convective” salt flux (tons/year)	6.6	69	0	0
Total salt flux (tons/year)	12	74	5.4	5.4
Volume increase of the cavity (m <sup>3</sup> /year)	6.5	40	2.5	2.5
Overflow volume of brine (m <sup>3</sup> /year)	21	214	0	0

**Figure 12.** Conceptual diagram of the current operation of Gellenoncourt.

simulation indicates that the annual volume of brine that would have to be evacuated to prevent overflow in the boreholes would amount to between 0.1 and 1 m<sup>3</sup> according to the assumption on the osmotic efficiency coefficient, considering the value of 0.2 ha for the area of the collapse chimney. Such volumes would go unnoticed by comparison to the various pumping or purging operations carried out for mining purposes. It should be noted that the boreholes of the Gellenoncourt field remained open between 1998, the date of the collapse, and 2002, the date of the resumption of leaching operations, and that no

variation in level was observed during this period.

In conclusion, the preceding quantitative evaluations come from a theoretical approach to the problem. They highlight the sensitivity of chemo-osmosis to the macroscopic coefficient of osmotic efficiency, the value of which remains poorly understood for natural environments, at the relevant scales.

An application of the theory with values of this coefficient chosen on an upper bound basis with reference to a well-documented compact argillite such as that investigated by Andra on the site of the geological disposal project for radioactive waste, shows that the chemical osmotic effect is not likely to significantly modify the orders of magnitude of the salt fluxes evacuated into the environment by intensive exploitation of a salt layer, as envisaged by the Salins group in Gellenoncourt.

In general, field observations on the sites that exploit the layered salt deposit of the Lorraine plateau indicate that no unexplained overpressure has ever been observed in the salt aquifer created by dissolution within the salt formation while collapsing chimneys, filled with clayey materials, bring the deep saturated brine in contact with the fresh waters of the superficial aquifer system. This is consistent with the theoretical approach presented which shows that osmotic phenomena are not likely to generate disturbances of an exceptional nature compared to the various disturbances inherent in the practices of exploitation of the salt deposit by dissolution.

### Conflicts of interest

Authors have no conflict of interest to declare.

## Version française

### 1. Position du problème

L'exploitation du sel gemme en couches ou en dômes par dissolution *in situ* fait l'objet d'une industrie ancienne développée depuis le milieu du 19<sup>e</sup> siècle. C'est le cas du gisement salifère lorrain encore largement exploité de nos jours pour la production de sel destiné à la chimie, à l'alimentation ou encore au déneigement. Parmi les exploitants actuels, figure le groupe Salins qui opère à proximité des villes de Dombasle et Varangéville, dans l'est de Nancy. Les couches de sel sollicitées font partie d'un puissant ensemble marno-salifère du Keuper. Deux méthodes d'exploitation sont pratiquées, d'une part, en extrayant du sel gemme par cavités minées selon la méthode des chambres et piliers abandonnés pour le déneigement (mine de Varangéville) et, d'autre part, en prélevant la saumure obtenue par lixiviation à l'eau douce de cavités profondes, pour l'industrie (salines du plateau d'Haraucourt).

Dans le respect d'une exploitation par lixiviation bien contrôlée, deux questions se posent parmi d'autres, optimiser la récupération du sel qui est une denrée de valeur et préserver l'environnement de toute pollution par de la saumure et de mouvements de terrain en surface. A cette fin, s'agissant du groupe Salins, la méthode par lixiviation vise actuellement à réaliser des cavités isolées les unes des autres, dont on peut démontrer la stabilité à long terme sur les plans hydrauliques et mécaniques. Cependant, certains secteurs anciens du gisement ont été exploités au moyen de cavités communicantes, ce qui a compromis la maîtrise des dissolutions résiduelles spontanées en fin de vie et entraîné, dans certains cas, des effondrements parvenant en surface, créant des cratères remplis d'eaux.

Afin de mettre en sécurité les lieux, tout en optimisant la production de sel, l'exploitant envisage de reprendre la dissolution *in situ* d'un panneau du site de Gellenoncourt, sur la concession de Drouville. La méthode d'exploitation employée, dite intensive, aura pour conséquence l'élargissement, suivi d'un effondrement, des cavités déjà existantes, aboutissant au final à un lac s'étendant sur l'emprise du panneau sur environ 36 ha. Dans le but d'obtenir l'autorisation de mettre en œuvre la méthode, l'exploitant a constitué un dossier d'impact examinant, sur

des bases scientifiques, les conséquences d'une telle pratique, notamment pour ce qui concerne le risque de pollution des aquifères environnants par des eaux salées.

Le présent article traite des mécanismes potentiels de relâchement de sel en provenance de cavités effondrées. Il met à profit des travaux de recherche de l'Agence nationale pour les déchets radioactifs (Andra) qui s'intéresse à la compréhension des surpressions interstitielles observées dans la couche d'argilite du Callovo-Oxfordien de Bure en Meuse/Haute-Marne, objet d'études pour la création d'un stockage géologique des déchets radioactifs de haute et moyenne activité à vie longue. Ces travaux réalisés en partenariat avec le Laboratoire de Mécanique des Solides de l'Ecole Polytechnique et l'UMR Sisyphé à l'Université Pierre et Marie Curie ont fait l'objet, en 2008, d'une thèse soutenue dans cette université, sous la présidence du Professeur Ghislain de Marsily. Nous avons emprunté les bases théoriques de notre raisonnement à cet ouvrage, ainsi qu'à plusieurs publications qui en ont résulté.

#### 1.1. *Éléments de géologie du gisement salifère de Lorraine*

Les nombreux sondages réalisés sur les concessions salines permettent de proposer la coupe géologique synthétique qui montre du haut vers le bas la succession suivante (Figure 1).

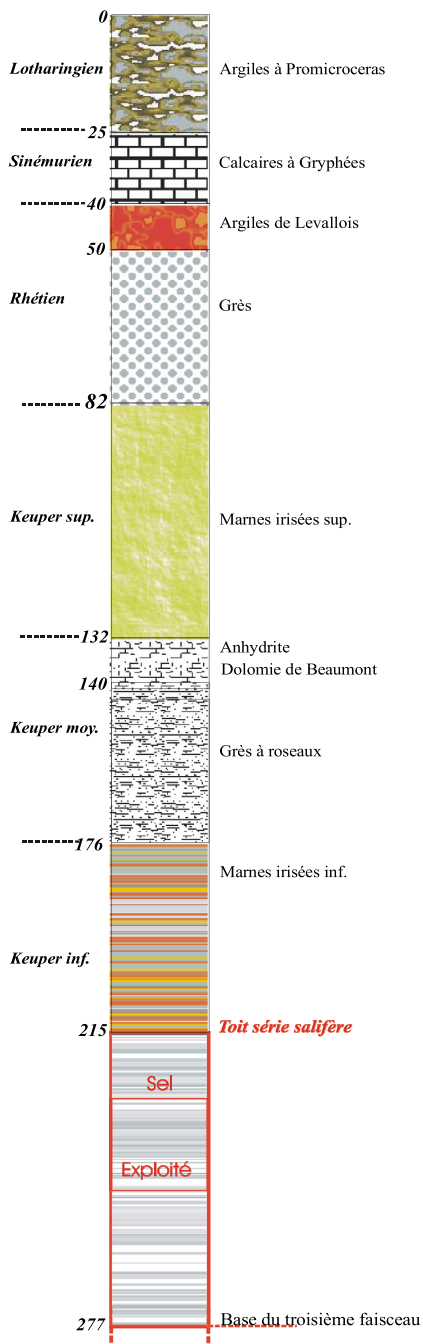
*Argiles noires à Promicroceras* (Lotharingien). Argilites gris-bleu plastiques, parfois schisteuses, d'épaisseur maximale 25 m.

*Calcaires à Gryphées* (Sinémurien). Alternance décimétrique de marnes bleues à noires et de bancs de calcaires argileux de couleur rouille, de 15 à 17 m d'épaisseur.

*Argiles de Levallois* (Rhétien). Argilites rose-saumon qui constituent un niveau repère dans toute la Lorraine. L'épaisseur moyenne est d'une dizaine de mètres.

*Grès infraliasiques* (Rhétien). Grès siliceux fins avec intercalations argileuses centimétriques à décimétriques noires ou vertes dans la partie sommitale. L'épaisseur moyenne est de 32 m.

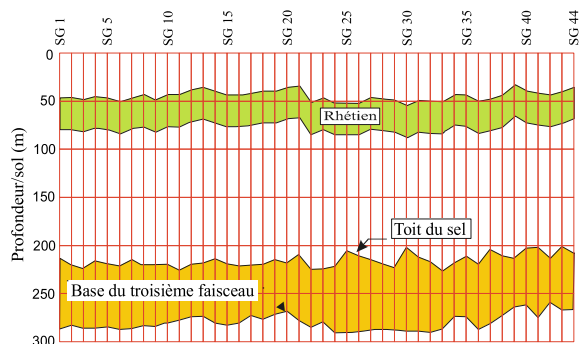
*Marnes irisées supérieures* (Keuper supérieur). Argilites vertes et violacées qui deviennent rouge



seur atteint 135 m à Gellenoncourt.

**FIGURE 1.** Log géologique synthétique.

brique vers la base. Des horizons dolomitiques centimétriques à décimétriques sont présents ainsi que, à la base, des passées décimétriques d’anhydrite blanche nodulaire ou en filets anastomosés. L’épais-



**FIGURE 2.** Coupe géologique illustrant la position du sel exploité par rapport à l'aquifère rhétien.

*Dolomie de Beaumont — Grès à roseaux* (Keuper moyen). Dolomie beige ou grise en bancs décimétriques surmontée d'un chapeau d'anhydrite massive de 3 m à 4 m d'épaisseur. La dolomie, épaisse de 8 m, joue un rôle important dans la stabilité des terrains avant les effondrements. Au-dessous on trouve des marnes bariolées puis des grès fins micacés très argileux contenant de nombreux débris végétaux.

*Marnes irisées inférieures* (Keuper inférieur). Cette puissante formation de l'ordre de 200 m d'épaisseur est composée d'argilites alternant avec des lits de gypse et d'anhydrite sur les quarante premiers mètres. On rencontre ensuite la série salifère épaisse de 120 m mais dont seuls les 3 faisceaux supérieurs, riches en sel et épais de 65 m, sont exploités à Gellenoncourt et renferment environ 20% d'insolubles. Ainsi, l'épaisseur cumulée de sel pur lixiviable représente en moyenne une quarantaine de mètres.

La régularité des couches apparaît sur la Figure 2 qui montre la position des niveaux repères que sont le Rhétien et le toit du sel, qui se placent respectivement à une moyenne de 50 m et 220 m sous le sol dans les 44 sondages présents sur l'ensemble du site de Gellenoncourt.

### 1.2. Éléments d'hydrogéologie du site

Parmi les formations géologiques présentes dans le sous-sol, seuls les grès du Rhétien constituent un aquifère susceptible d'être impacté par l'exploitation saline.

Épais d'une trentaine de mètres, cet aquifère a contribué à l'alimentation en eau potable des communes voisines. Il est encore faiblement exploité pour l'agriculture. Protégé par la couverture peu perméable et bien continue des argiles de Levallois, ce réservoir est alimenté par ses affleurements en bordure sud et est du plateau. La transmissivité est modeste et varie entre  $10^{-4}$  m<sup>2</sup>/s et  $10^{-5}$  m<sup>2</sup>/s, ce qui conduit à une grande sensibilité du niveau piézométrique aux prélèvements.

La nappe est le plus souvent captive sous les argiles de Levallois et est artésienne dans les parties basses de la topographie. La piézométrie générale indique un écoulement vers le nord-ouest vers la rivière Roanne qui draine cette nappe.

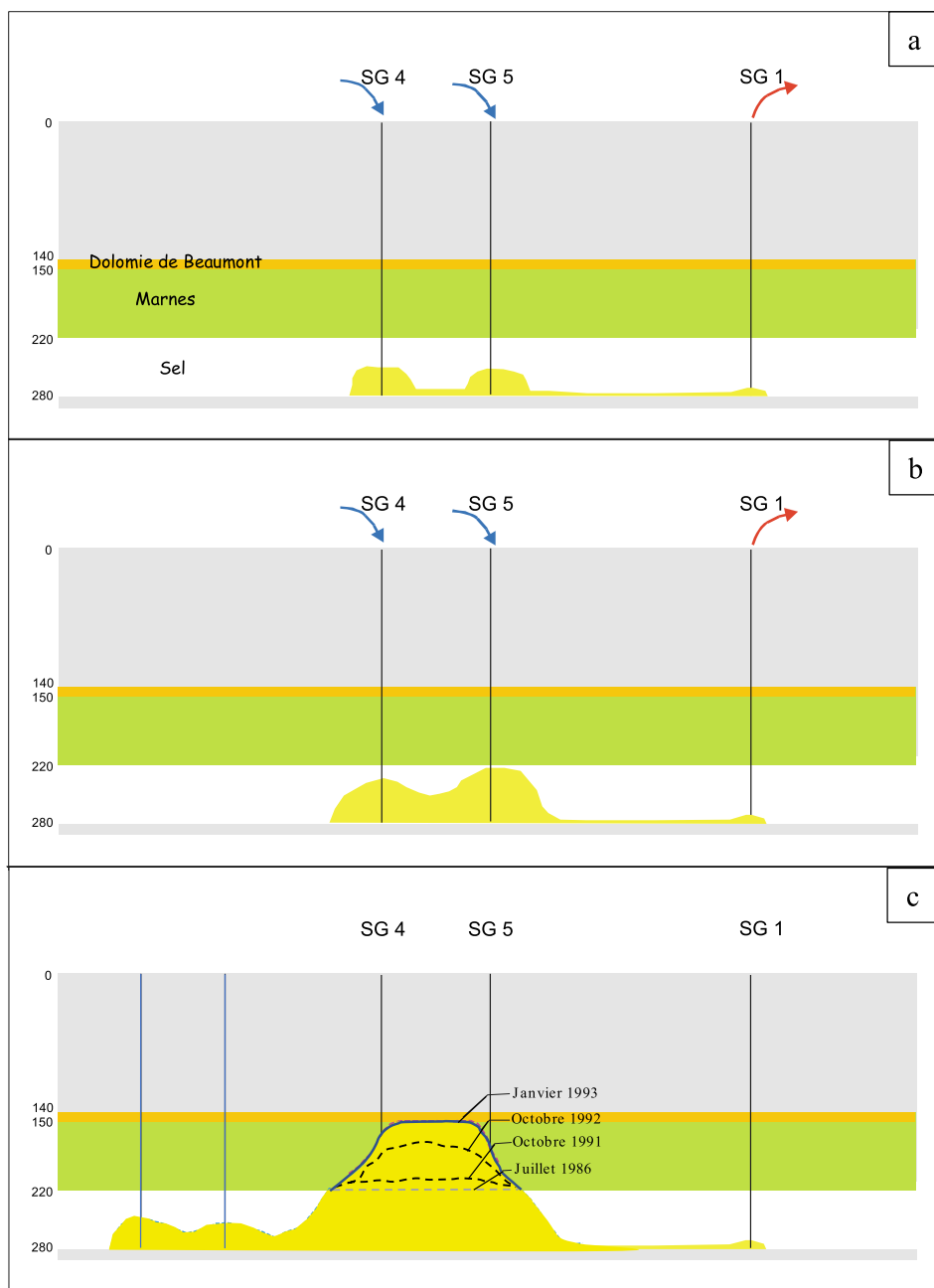
### 1.3. Les conséquences de l'exploitation du sel

L'exploitation du panneau de Gellenoncourt a été menée pendant une vingtaine d'années à partir de 1967. La méthode a consisté à relier par fracturation hydraulique la base de plusieurs sondages creusés jusqu'au mur du troisième faisceau salifère, puis à introduire de l'eau douce dans un ou plusieurs ouvrages et à pomper la saumure dans des sondages éloignés dits « extracteurs », entraînant ainsi par dissolution du sel la création de cavités au voisinage des injecteurs. Une bonne adéquation entre le débit et l'agencement du dispositif assurait la saturation de la saumure. Des contrôles périodiques de la géométrie des vides étaient réalisés par ultra-sons afin de surveiller l'évolution des cavités.

Ainsi l'exploitation du sel a développé des vides dont la base se trouve vers 65 m de profondeur sous le toit du sel et dont la morphologie évolue dans le temps et dans l'espace au fur et à mesure de la production de sel. La géométrie précise du réseau d'interconnexion des cavités est mal connue ; les mesures ultra-soniques réalisées dans les forages, indiquent par contre clairement l'extension des cavités créées au niveau des injecteurs.

En évolution normale l'ensemble de ces vides reste confiné au sein de la formation salifère. Cependant cette évolution a été perturbée dans le passé par l'effondrement de deux cavités centrales du dispositif de Gellenoncourt, SG4 et SG5 (Figure 3) selon l'historique décrit ci-dessous.

Les puits SG4 et SG5 ont été utilisés comme injecteurs d'eau douce dès le début de l'exploitation, SG1



**FIGURE 3.** Evolution des cavités SG4–SG5 jusqu’à la Dolomie de Beaumont.

étant l’extracteur de saumure (Figure 3a). Le dispositif a été maintenu jusqu’à ce qu’il ait été constaté que la planche de sel au sommet de SG5 était réduite à 2 m, (Figure 3b) après quoi l’injection d’eau douce a été reportée en d’autres points. Un suivi des cavi-

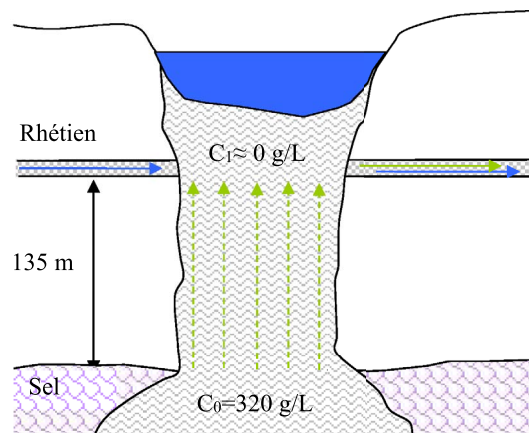
tés a par la suite montré que leur volume continuait à s’accroître et l’arrêt d’exploitation de l’ensemble du panneau a été décidé. De 1982 à 1993 les contrôles ont cependant mis en évidence une remontée du toit de SG5 dans les argilites de couverture du sel,

d'abord lente, puis rapide jusqu'à parvenir au contact de la Dolomie de Beaumont (Figure 3c). L'exploitant a alors pris la décision de provoquer l'effondrement en abaissant volontairement le niveau de saumure pour maintenir un matelas d'air au sommet de la cavité. L'événement s'est finalement produit en mars 1998, entraînant la formation d'un cratère en surface d'une cinquantaine de mètres de diamètre et d'une profondeur de 120 m [Buffet, 1998]. La cheminée s'est ensuite partiellement comblée par talutage des bords du cratère. Le cratère était initialement sec après l'effondrement; il s'est ensuite rempli d'eau et était occupé par un lac dont le niveau se situait à environ 30 m de la surface en juin 1998.

Cet effondrement a déstructuré localement la couverture protectrice du gisement salin et mis en communication potentielle diverses unités hydrauliques : le lac de cratère, l'aquifère du Rhétien et les cavités d'exploitation remplies de saumure et de résidus de dissolution. Compte tenu de la diversité lithologique des terrains de couverture, les matériaux qui remplissent la cheminée d'effondrement sont hétérogènes avec une dominante argileuse. Leurs propriétés règlent les échanges de fluide et de sel dissous entre ces unités.

Cette expérience vécue a motivé le groupe Salins à envisager de reprendre la lixiviation du panneau jusqu'à effondrement des cavités encore stables dans le double but d'accroître la production de sel mais surtout de mettre le site en sécurité sur le long terme.

Actuellement le lac présente une superficie d'environ 1,7 ha; il devrait s'étendre jusqu'à 36 ha à terminaison après reprise de l'exploitation. Une fois la totalité des effondrements obtenus, le système évoluera vers une nappe salée en profondeur et un lac en surface reliés hydrauliquement par les terrains saturés peu perméables remplissant une cheminée d'effondrement. La nappe salée sera confinée au sein de la formation salifère grâce à la forte densité de la saumure qui empêchera toute circulation convective malgré la présence d'un gradient thermique, dans la mesure où le lac constituera un plan d'eau unique n'engendrant pas de gradient hydraulique dans le système.



**FIGURE 4.** Schéma conceptuel de la diffusion du sel dans la cheminée d'effondrement vers le Rhétien.

## 2. Examen du comportement à long terme

### 2.1. La diffusion, un mécanisme potentiel de l'émission de sel

On considère ici le cas de la situation « après-mine », à long terme, lorsque la configuration du site sera stabilisée. L'origine du sel pouvant polluer le Rhétien sera la saumure saturée présente dans les cavités. Cette saumure dense ne pourra pas circuler par convection mais elle pourrait être une source de sel dissous se déplaçant par diffusion à travers les matériaux de la cheminée d'effondrement vers les eaux plus douces du lac. On peut évaluer le flux de diffusion au moyen du modèle simple suivant (Figure 4).

Les matériaux de la cheminée sont supposés homogènes avec un coefficient de diffusion effectif  $D$  moyen pour les différents solutés. Une condition à la limite de concentration à saturation à la valeur  $C_0$  de 320 g/L est imposée par la nappe salée à la base de la cheminée qui pénètre dans la couche de sel, tandis qu'une faible concentration de valeur  $C_1$  est maintenue au niveau du Rhétien sous l'effet de l'écoulement de la nappe qui dilue le sel. Dans ces conditions la densité de flux diffusif en régime stationnaire est donnée par la loi de Fick :

$$F = -D \text{grad}(C) = -D(C_1 - C_0)/H$$

où  $H$  est la distance verticale séparant la nappe salée de la base du Rhétien.

Le coefficient de diffusion effectif dans les matériaux remplissant la cheminée est majoré par le coefficient de diffusion dans l'eau libre dont la valeur est de l'ordre de  $10^{-9}$  à  $2 \times 10^{-9}$  m<sup>2</sup>/s. Nous avons considéré une réduction d'un facteur 10 du fait de la présence de solides qui gênent la diffusion et avons adopté la valeur de  $2 \times 10^{-10}$  m<sup>2</sup>/s pour le calcul du flux surfacique de sel :

$$F = 2 \times 10^{-10} \cdot (320 - 0) / 135 \approx 4,7 \times 10^{-10} \text{ kg/m}^2/\text{s}$$

en utilisant les valeurs numériques de la Figure 4.

Pour une cheminée d'effondrement de 36 ha, on obtient un flux annuel de 5 à 6 tonnes de sel. Ce flux diffusif est évacué dans l'aquifère du Rhétien sous l'effet de l'écoulement de la nappe.

## 2.2. L'osmose chimique, un autre mécanisme

La Mission Régionale de l'Autorité environnementale qui a évalué le dossier présenté par le groupe Salins a remis un avis en octobre 2018, à la requête du Préfet de Meurthe-et-Moselle et a fait remarquer que les phénomènes osmotiques n'étaient pas considérés et pouvaient présenter un risque potentiel vis-à-vis de la stabilité à long terme.

La présente note expose quelques réflexions sur la nature de ce phénomène et sur le rôle qu'il pourrait éventuellement jouer dans la configuration de l'exploitation envisagée par les Salins.

Il est fait, dans un premier temps, appel aux bases théoriques de l'osmose chimique. On s'intéresse dans un second temps à la manière dont la théorie peut être utilisée dans le contexte particulier de l'exploitation saline, afin de dégager des ordres de grandeur des conséquences que l'on peut attendre du phénomène. Enfin, une confrontation avec des observations de terrain fournira des éléments de validation de l'approche retenue.

## 3. Les bases théoriques de l'osmose chimique

L'osmose chimique ou chemo-osmose désigne le mouvement de l'eau se produisant au travers d'une membrane semi-perméable dont les faces sont baignées par des solutions ioniques de concentrations différentes. Une telle membrane est dite parfaitement semi-perméable lorsqu'elle autorise le passage des molécules d'eau mais s'oppose totalement au transfert des ions. L'eau traverse ainsi la membrane

depuis la zone où la concentration est faible vers la zone à forte concentration de façon à provoquer une diminution du gradient de concentration par effet de dilution.

En pratique, les membranes ne sont qu'imparfaitement semi-perméables et laissent passer les ions à des vitesses variables selon leurs propriétés; on désigne par diffusion ce type de transfert au cours duquel les ions de la solution passent de la zone de concentration élevée à la zone de concentration faible. Ainsi, pour les deux types de mécanismes, osmose et diffusion, l'effet est de tendre vers une homogénéisation des concentrations de part et d'autre de la membrane.

Le schéma de la Figure 5 [Rousseau-Gueutin, 2008], décrit ces phénomènes dans les cas d'une membrane parfaite et imparfaite.

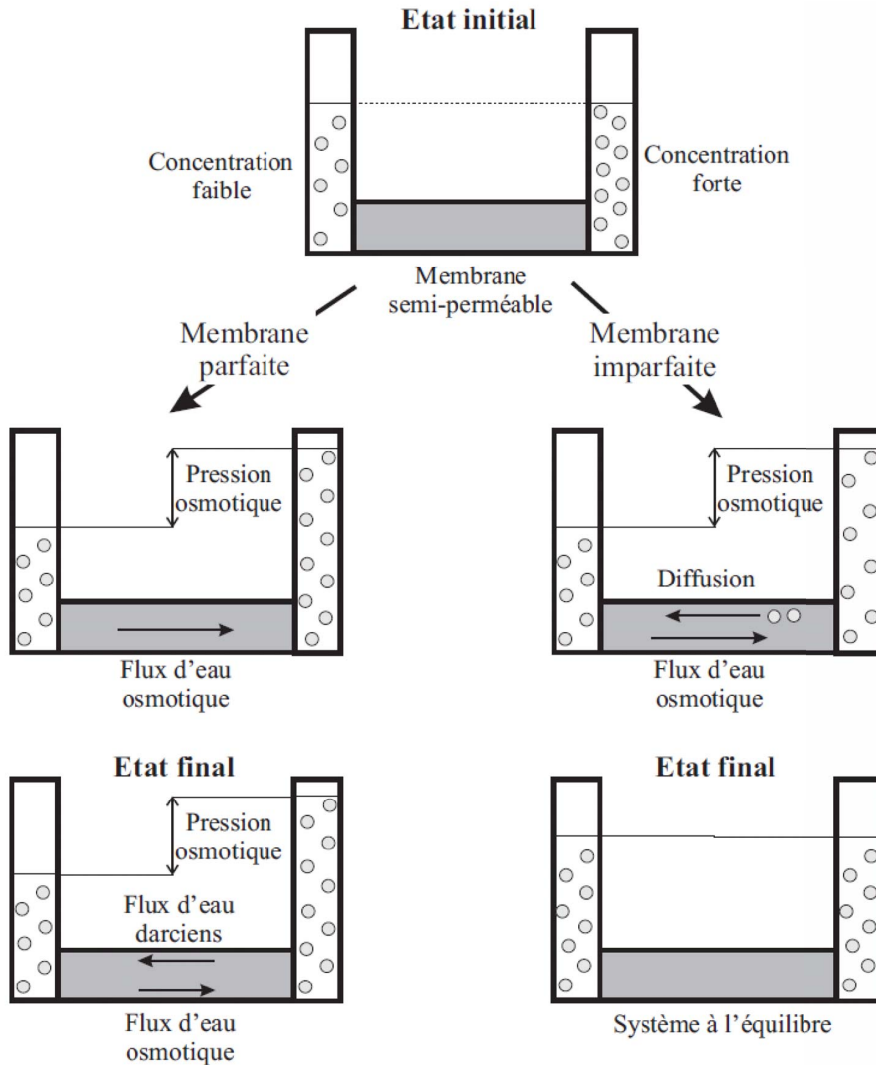
Le schéma met en évidence une différence de pression entre les deux branches du tube en U, dite pression osmotique. Dans le cas d'une membrane parfaite, la différence de pression liée à l'activité chimique engendre un écoulement, conforme à la loi de Darcy, à contre-courant de l'écoulement osmotique et le système s'équilibre lorsque les flux d'eau osmotique et darcéen sont égaux pour la pression osmotique correspondant à la différence de concentration résiduelle entre les deux branches du tube. Dans le cas d'une membrane imparfaite, la présence de la pression osmotique n'est que transitoire car l'écoulement darcéen et la diffusion tendent à homogénéiser la concentration; l'état final se traduit alors par l'égalité des pressions et des concentrations dans les deux branches. La durée de cette phase transitoire dépend de la perméabilité de la membrane.

On note donc que l'état final du système est particulièrement dépendant des propriétés de la membrane. Il dépend également du comportement de la source d'ions comme cela sera mis en évidence lors de l'application à un gisement salifère.

### 3.1. Le modèle thermodynamique

La thermodynamique des phénomènes irréversibles fournit un cadre théorique pour modéliser les phénomènes de couplage multi-physiques à l'échelle macroscopique. Cette théorie postule que, de manière générale, le flux d'une grandeur est relié de façon tensorielle à une force motrice par l'intermédiaire d'une





**FIGURE 5.** Illustration de la diffusion et de l'osmose chimique dans un tube en U [d'après Rousseau-Gueutin, 2008].

matrice de coefficients phénoménologiques. Les différentes forces motrices s'expriment, elles-mêmes, sous forme du gradient d'un champ scalaire correspondant aux différentes variables d'état du système.

Le tableau 1 suivant rassemble les différentes forces et flux entrant en ligne de compte dans les problèmes de mécanique des fluides concernés par les phénomènes chimiques, thermiques et électriques.

Les mécanismes couplés diagonaux sont dits dominants et correspondent aux phénomènes classiquement pris en compte en physique. Les mécanismes non-diagonaux traduisent les trans-

ferts non-dominants qui ne jouent un rôle notable que dans des configurations particulières; l'osmose chimique en fait partie.

Les coefficients phénoménologiques reliant chaque flux à chaque force ne sont pas prédits par la théorie macroscopique. Ils peuvent être estimés par changement d'échelle à partir d'autres modèles considérant l'aspect microscopique des milieux ou, le plus souvent, empiriquement au moyen d'expériences réalisées à la bonne échelle [Gonçalves and Trémosa, 2010].

**TABLEAU 1.** Mécanismes couplés reliant forces et flux

Flux	Forces			
	Gradient de charge ou de pression	Gradient de potentiel électrique	Gradient de concentration	Gradient de température
Vitesse d'écoulement	Loi de Darcy	Electro-osmose	Osmose chimique	Osmose thermique
Densité de courant électrique	Courant de filtration	Loi d'Ohm	Courant de diffusion	Effet Seebeck
Flux de soluté	Osmose inverse	Electrophorèse	Loi de Fick	Effet Soret
Flux de chaleur	Filtration de la chaleur	Effet Peltier	Effet Dufour	Loi de Fourier

### 3.2. Le cas des matériaux argileux

Selon Goncalvès *et al.* [2015], dès les années 60, les expérimentateurs travaillant sur les milieux poreux avaient mis en évidence des écarts à la loi de Darcy dans les matériaux argileux, une des raisons principales évoquées étant l'existence d'un effet osmotique en présence de fluides salins. Les connaissances sur le sujet se sont développées à partir des années 90, notamment à la faveur des études sur le stockage des déchets radioactifs dans les formations argileuses peu perméables. Notons que ces études concernaient essentiellement des matériaux préparés artificiellement ou remaniés, à l'échelle de l'échantillon de laboratoire; très peu portaient sur des matériaux naturels et encore moins en conditions *in situ*. A l'échelle d'une couche géologique de 130 m d'épaisseur à 500 m environ de profondeur, l'Andra a mis en évidence des surpressions dans les argilites du Callovo-Oxfordien en Meuse/Haute-Marne. Le rôle de l'osmose chimique a été évalué et il apparaît, sans que cela soit démontré de manière définitive, qu'une part de ces surpressions pourraient être dues aux phénomènes osmotiques [Rousseau-Gueutin, 2008, Goncalvès *et al.*, 2015, Gueutin *et al.*, 2007].

L'origine de l'effet osmotique dans les argiles est expliquée par le fait que la surface des minéraux argileux qui est chargée négativement impose des restrictions électriques sur les ions lors de leur transport, entraînant une concentration des cations près de la surface et une exclusion des anions vers le centre des pores. Pour respecter la neutralité électrique la restriction sur le transport des anions se répercute sur les cations, ce qui limite le transport ionique total à l'instar de ce que ferait une membrane imparfaite. Les charges de surface sont plus facilement compensées lorsque la force ionique de

la solution est élevée et l'effet est d'autant plus important que l'épaisseur des couches électrostatiques se rapproche de la taille des pores. On conçoit ainsi que l'efficacité du mécanisme osmotique soit liée à la nature des minéraux argileux, à la taille des pores ainsi qu'à la concentration des solutions. Les roches argileuses peu perméables, telles que les argilites du Callovo-Oxfordien, sont caractérisées par des dimensions porales de quelques nanomètres comparables à celles des couches électrostatiques, ce qui explique l'occurrence des phénomènes osmotiques dans ces systèmes.

En résumé, l'existence de l'effet osmotique dans les matériaux argileux est établie; la question est de savoir quelle peut-être son intensité dans la configuration qui nous occupe.

### 3.3. Les lois et les paramètres

Goncalvès *et al.* [2015] ont dressé une synthèse très détaillée des couplages dans les membranes et les matériaux argileux. Il ressort de cette synthèse que les relations phénoménologiques représentant le couplage entre l'écoulement obéissant à la loi de Darcy, l'effet osmotique et le transport par convection-diffusion à l'échelle macroscopique s'écrivent de la façon suivante, en régime stationnaire :

$$\begin{aligned}
 V &= -\frac{k}{\mu}(\text{grad}(p) + \rho g \text{grad}(z)) \\
 V' &= \frac{k}{\mu}(\varepsilon \text{grad}(\pi)) \\
 \text{div}(V + V') &= 0 \\
 J &= -D \text{grad}(C) + VC \\
 \text{div}(J) &= 0 \\
 \text{grad}(\pi) &= \nu RT \text{grad}(C)
 \end{aligned} \tag{1}$$

où  $V$  (m/s) est la vitesse de Darcy,  $V'$  (m/s) la vitesse osmotique de l'eau,  $J$  (mol/s) la densité de flux diffusif et convectif de soluté,  $p$  (Pa) la pression du fluide,  $z$  (m) l'altitude,  $\pi$  (Pa) la pression osmotique,  $C$  (mol/m<sup>3</sup>) la concentration en soluté,  $\rho$  (kg/m<sup>3</sup>) et  $\mu$  (kg/m/s) respectivement la masse volumique et la viscosité dynamique du fluide,  $D$  (m<sup>2</sup>/s) le coefficient de diffusion effectif en milieu poreux et  $k$  (m<sup>2</sup>) la perméabilité intrinsèque.

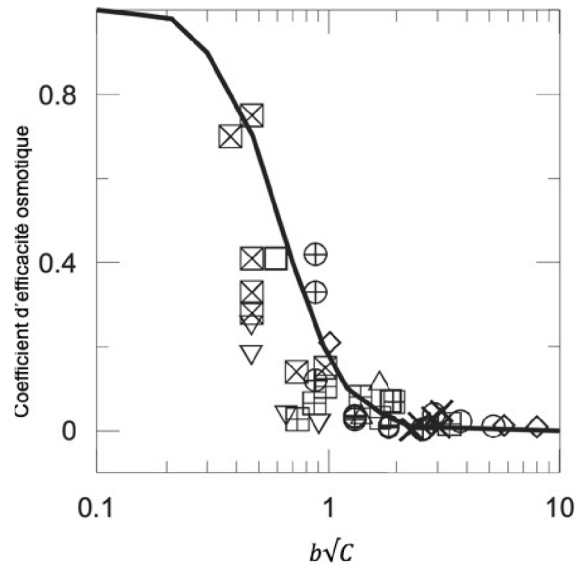
La pression osmotique  $\pi$  est donnée par la formule de Van't Hoff où  $R$  (J/mol/K) est la constante des gaz parfaits,  $T$  la température Kelvin et  $\nu$  le nombre d'ions séparés (2 pour NaCl).

On remarque l'existence de deux vitesses pour l'eau, la vitesse de Darcy  $V$  classique, engendrée par le gradient de charge hydraulique et la vitesse osmotique  $V'$  engendrée par la pression osmotique, donc par le gradient de concentration. Seule la vitesse de Darcy intervient sur le transport de soluté puisque le phénomène osmotique ne déplace que les molécules d'eau. Le phénomène osmotique n'est donc pas susceptible de provoquer le déplacement des ions.

Un paramètre essentiel est le coefficient d'efficacité osmotique  $\varepsilon$  (sans dimension) qui varie entre 0 et 1. Dans le cas d'une membrane parfaite, ce coefficient vaut 1 et le coefficient de diffusion est alors nul, traduisant l'absence de transport ionique. Il est inférieur à 1 pour les milieux naturels en fonction de leur caractère plus ou moins semi-perméable. Dans de tels milieux les trois paramètres fondamentaux sont la perméabilité, le coefficient de diffusion effectif et le coefficient d'efficacité osmotique. Leur détermination est expérimentale, à l'aide de systèmes nécessitant une métrologie particulièrement pointue et, s'agissant de coefficients macroscopiques, leur valeur est influencée par l'échelle du problème traité.

Des considérations théoriques [Bresler, 1973] établissent que le coefficient d'efficacité osmotique est une fonction décroissante du produit de la taille des pores du milieu par la racine carrée de la concentration. Goncalvès et al. [2015] ont rassemblé sur une même figure (Figure 6) les mesures disponibles dans la littérature en fonction de ces paramètres en regard de la courbe théorique de Bresler.

Ces données montrent que le coefficient d'efficacité osmotique mesuré est en général faible avec des valeurs en assez bon accord avec la courbe théorique. Notons que les points relatifs aux argilites de Bure, très compactes, se situent dans les faibles valeurs.



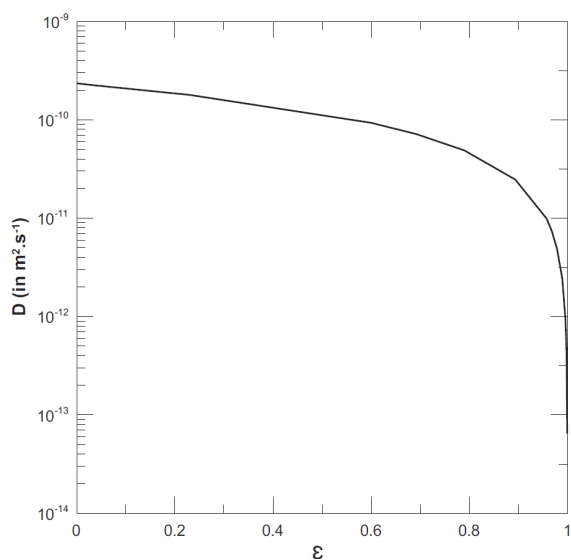
**FIGURE 6.** Coefficient d'efficacité osmotique en fonction de  $b\sqrt{C}$ .  $b$  est la taille moyenne des pores en nm et  $C$  la concentration en mol/L.  $\boxtimes$  représentent les valeurs expérimentales pour l'argilite de Bure [d'après Goncalvès et al., 2015].

La théorie rend également compte de la dépendance du coefficient de diffusion avec le coefficient d'efficacité osmotique. La Figure 7 tirée de Rousseau-Gueutin [2008], d'après les travaux de Revil et al. [2005] et Gonçalvès et al. [2007], indique l'allure de cette dépendance, établie pour un matériau composé d'une argile de type smectite. On remarque qu'elle est peu marquée pour les faibles valeurs de l'efficacité osmotique, ce qui suggère que l'osmose a peu d'effet sur le transport diffusif de soluté ionique dans les milieux qui se comportent comme des membranes imparfaites.

### 3.4. Application à une cavité saline

Dans le but d'appliquer ce qui précède à la situation d'une cavité saline effondrée, nous retiendrons les éléments suivants :

- il est prouvé que des phénomènes d'osmose chimique peuvent se manifester dans les matériaux argileux. Ceci est lié aux propriétés électriques de surface des minéraux argileux et à la faible dimension des pores qui



**FIGURE 7.** Variation du coefficient de diffusion effectif en fonction du coefficient d'efficacité osmotique pour une smectite en présence d'une solution NaCl [Rousseau-Gueutin, 2008].

confèrent aux milieux argileux un comportement de membrane semi-perméable imparfaite;

- à l'échelle macroscopique, le phénomène est gouverné par un coefficient d'efficacité osmotique variant entre 0 et 1 qui pondère l'influence de la pression osmotique dans la relation phénoménologique qui décrit le flux d'eau;
- le coefficient d'efficacité osmotique est très variable, de mesure complexe et incertaine, et n'a été que rarement évalué dans les milieux naturels;
- le coefficient d'efficacité dépend de la taille des pores et de la concentration, ce qui rend le phénomène non-linéaire donc difficile à quantifier par une approche simple car la taille des pores n'est pas un paramètre directement accessible;
- les matériaux remplissant la cheminée d'effondrement des cavités salines sont pour l'essentiel des matériaux argileux déstructurés dont la porosité est accrue par rapport à celle d'une argilite en place, ce qui

conduit à une augmentation relative de la taille moyenne des pores. De plus les concentrations en sel sont très élevées, tout au moins en partie basse du système. Ces éléments sont en faveur d'une valeur très faible du coefficient d'efficacité osmotique si on les compare aux connaissances acquises sur les argilites du Callovo-Oxfordien de Bure qui sont considérées comme bien caractérisées et ayant des pores de très faibles dimensions. L'hypothèse qui consiste à considérer que le remplissage de la cheminée est constitué d'argile homogène conduit ainsi à majorer le rôle des mécanismes d'osmose, ce qui va dans le sens de la sécurité dans l'appréciation du danger;

- le phénomène osmotique ne modifie pas le transport de soluté ionique au sein du milieu, qui reste dépendant du mécanisme de convection-diffusion.

#### 4. Modélisation exploratoire du comportement d'une cavité effondrée

Dans le but de quantifier l'effet potentiel de l'osmose sur le comportement d'une cavité saline effondrée, nous avons tenté une modélisation simple sur la base du schéma conceptuel de la Figure 4.

##### 4.1. Principes du modèle

Les hypothèses sont les suivantes :

- les transferts sont considérés s'effectuer de façon monodimensionnelle verticale entre la cavité effondrée profonde et le système aquifère du Rhétien, au travers de la cheminée sur 135 m de hauteur;
- la concentration en profondeur est maintenue par la dissolution du sel à 320 g/L, soit 5480 mol/m<sup>3</sup>. Elle est imposée à 0 g/L au niveau du Rhétien. Ceci fixe le gradient de concentration à la valeur 0,04 mol/L/m;
- le système est représenté en régime stationnaire, dans deux configurations distinctes de conditions aux limites hydrauliques illustrées sur la Figure 8

- *configuration 1* : la cavité est en relation hydraulique avec d'autres cavités dont le niveau de saumure est maintenu constant; c'est actuellement le cas à Gellenoncourt, où l'effondrement du panneau exploité n'est que partiel. On considèrera un équilibre hydrostatique entre le Rhétien et la cavité;
- *configuration 2* : la cavité effondrée est confinée, sans échange hydraulique possible avec le milieu extérieur; cette configuration est censée représenter l'état final lorsque l'ensemble du panneau sera effondré sur environ 36 ha;
- dans les deux configurations le profil de concentration à l'équilibre dans la cheminée est considéré comme diffusif avec un gradient constant entre 320 et 0 g/L.

Sous ces hypothèses le système d'équations (1) se simplifie en (2)

$$\begin{aligned} \frac{\partial}{\partial z} K \left( -\frac{\partial h}{\partial z} + \frac{\varepsilon \nu RT}{\rho g} \frac{\partial C}{\partial z} \right) &= 0 \\ D \frac{\partial^2 C}{\partial z^2} &= 0 \end{aligned} \quad (2)$$

avec les notations de (1) complétées par  $h$  (m), charge hydraulique de la saumure et  $K$  (m/s), perméabilité de Darcy des matériaux remplissant la cheminée.

Le système d'équations (2) est non linéaire car le coefficient d'efficacité osmotique dépend de la concentration. La résolution numérique par différences finies fournit le profil de charge dans la cheminée et permet le calcul des flux d'eau et de sel échangés entre la cavité et l'aquifère rhétien.

#### 4.2. Choix des paramètres

La perméabilité de Darcy  $K$  et le coefficient de diffusion effectif  $D$  ont été fixés respectivement aux valeurs  $10^{-10}$  m/s et  $2 \times 10^{-10}$  m<sup>2</sup>/s. Remarquons que dans le cas où ces paramètres sont constants sur la verticale, ils n'ont pas d'influence sur le profil de charge calculé. Ils interviennent par contre de manière linéaire sur les flux d'eau et de sel. Nous considérons que les valeurs choisies sont dans la gamme haute et que les flux seront ainsi évalués de façon majorante.

Comme indiqué précédemment, le paramètre essentiel du phénomène osmotique est le coefficient

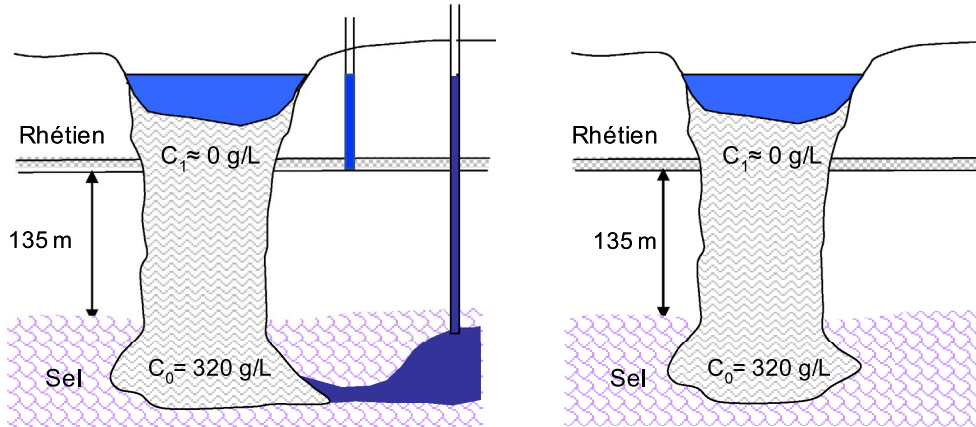
d'efficacité osmotique  $\varepsilon$ . Nous l'avons estimé à partir de la courbe de Bressler (cf. Figure 6) en considérant une taille de pores moyenne de 10 nm et un profil de concentration linéaire dans la cheminée entre 320 et 0 g/L. Cette taille de pores a été fixée arbitrairement par référence à la valeur 4 nm présentée par Rousseau-Gueutin [2008], pour les argilites du Callovo-Oxfordien de Bure. Il est très probable qu'elle soit largement sous-estimée si l'on considère que la porosité des matériaux argileux déstructurés de la cheminée doit être notablement accrue par rapport à celle d'argilites compactes. Par comparaison, les argiles de surface de l'Albien dans l'Aube présentent des tailles de pores entre 50 et 80 nm.

Avec ces hypothèses, choisies délibérément pour majorer les effets osmotiques, les valeurs du coefficient d'efficacité osmotique s'échelonnent entre 10% en partie supérieure où la concentration est faible et 0,01% en partie inférieure où la saumure atteint la saturation (Figure 9). Afin d'obtenir une appréciation de la sensibilité à ce paramètre, une simulation a également été réalisée en adoptant une valeur homogène de 1% le long du profil.

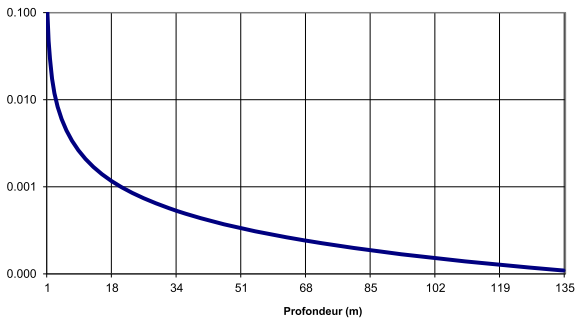
#### 4.3. Résultats des simulations

Les Figures 10 et 11 montrent les profils de surcharge hydraulique calculés pour les deux configurations de conditions aux limites du modèle. La surcharge représente la différence de charge hydraulique le long du profil par rapport à la charge du Rhétien considérée comme référence.

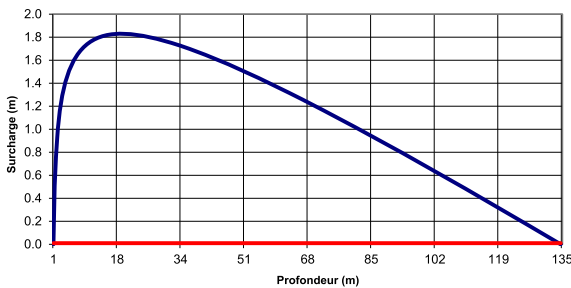
Dans le cas de la configuration 1 où la cavité est connectée au système aquifère, une charge hydraulique de 0 m est arbitrairement imposée en haut et en bas de la cheminée. Le modèle calcule donc une surcharge par rapport à la charge hydrostatique. On constate que l'effet osmotique ne déforme le profil de charge que si le coefficient d'efficacité osmotique est variable en fonction de la profondeur. On retrouve ainsi les résultats de Rousseau-Gueutin [2008], et Gonçalves et al. [2007]. Dans les deux cas, la cheminée est parcourue par un flux d'eau descendant, dont les vitesses sont respectivement de  $2,2 \times 10^{-12}$  m/s et  $2,02 \times 10^{-11}$  m/s. Ce flux est uniquement la conséquence de l'effet osmotique; il est évacué vers l'extérieur par la condition aux limites de niveau constant



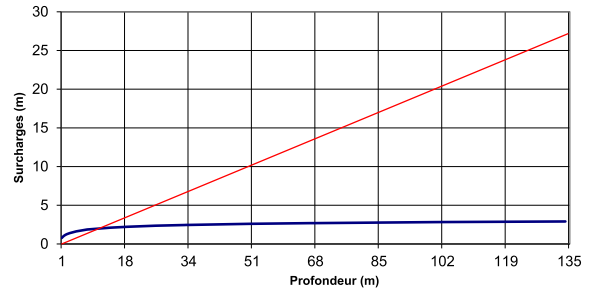
**FIGURE 8.** Configurations de simulation (à gauche cavité connectée, à droite cavité isolée).



**FIGURE 9.** Profil du coefficient d'efficacité osmotique en fonction de la profondeur pour une taille de pores de 10 nm.



**FIGURE 10.** Configuration 1, profils de surcharge hydraulique en fonction de la profondeur calculés avec un coefficient d'efficacité osmotique variable (en bleu) et constant (en rouge).



**FIGURE 11.** Configuration 2, profils de surcharge hydraulique en fonction de la profondeur calculés avec un coefficient d'efficacité osmotique variable (en bleu) et constant (en rouge).

imposée à la base de la cheminée. Notons qu'une interprétation en termes purement « darcéens » du profil déformé conduirait à une évaluation erronée des intensités et des orientations des flux.

Dans le cas de la configuration 2 où la cavité est totalement confinée, aucun flux d'eau ne peut parcourir la cheminée en régime stationnaire, de telle sorte que la charge s'élève avec la profondeur pour compenser la pression osmotique. La surcharge atteint ainsi 2,9 m ou 27 m au niveau de la cavité, selon le profil de coefficient d'efficacité osmotique adopté. De telles surcharges ne mettraient pas en danger la stabilité mécanique de la cavité compte tenu de la profondeur du recouvrement. Rappelons que les valeurs de coefficients d'efficacité osmotique sont certainement choisies par excès.

#### 4.4. Bilan en sel de la cavité

Les simulations précédentes permettent de quantifier le flux de sel évacué à partir de la cavité dans les différentes configurations.

Ce flux comporte deux composantes :

- une composante par diffusion engendrant un flux vers le haut de la cheminée. Cette part de flux est indépendante de l'effet osmotique dans la mesure où le coefficient de diffusion n'est que très peu influencé par cet effet pour les faibles valeurs de coefficient d'efficacité osmotique considérées (cf. Figure 7) ;
- une composante, dite « convective », liée au flux d'eau descendant qui atteint la cavité sous l'effet de l'osmose dans le cas de la configuration 1 où celle-ci est connectée hydrauliquement au milieu extérieur. Remarquons que ce flux est constitué d'eau douce mais qu'il se charge en sel à saturation lorsqu'il atteint la formation salifère. On considèrera que la masse de NaCl dissoute par 1 L d'eau pure est de 355 à 360 g pour une saumure saturée à 320 g/L de masse volumique 1220 kg/m<sup>3</sup>.

Le Tableau 2 indique ces différents flux pour une cheminée d'effondrement de 36 ha, telle que cela devrait être le cas à terminaison de l'exploitation du panneau de Gellenoncourt.

Rappelons les valeurs adoptées pour la perméabilité de Darcy et le coefficient de diffusion, qui sont respectivement de  $10^{-10}$  m/s et  $2 \times 10^{-10}$  m<sup>2</sup>/s. Les flux de sel sont directement proportionnels à ces valeurs.

On notera que le flux diffusif de sel est indépendant de l'effet osmotique mais que le flux convectif est directement lié à cet effet avec une forte sensibilité au profil du coefficient d'efficacité osmotique. La ligne du tableau intitulée « augmentation de volume de la cavité » correspond au volume de sel dissous en profondeur sous l'effet des différents flux. Le volume de débordement de saumure représente la quantité de saumure qu'il faut théoriquement évacuer pour annuler la surcharge engendrée dans la cavité par l'effet osmotique.

En complément des processus de diffusion et d'osmose invoqués ci-dessus, l'expérience montre que les cavités salines profondes sont soumises au fluage du massif salifère environnant qui engendre

des remontées de pression ou des débordements de saumure selon que la cavité est maintenue fermée ou ouverte [Brouard *et al.*, 2017]. Ce mécanisme n'apparaît pas devoir jouer un rôle important dans le cas qui nous occupe car les cavités sont peu profondes et se trouvent auto-remblayées par les matériaux issus de la cheminée d'effondrement. Des mouvements de terrain sont cependant susceptibles d'exister pour compenser l'augmentation de volume engendrée par la dissolution en profondeur sous l'effet du flux convectif d'eau douce provoqué par l'osmose. Ces mouvements seraient toutefois très lents étant donné les faibles volumes en jeu et concerneraient en priorité la profondeur du lac.

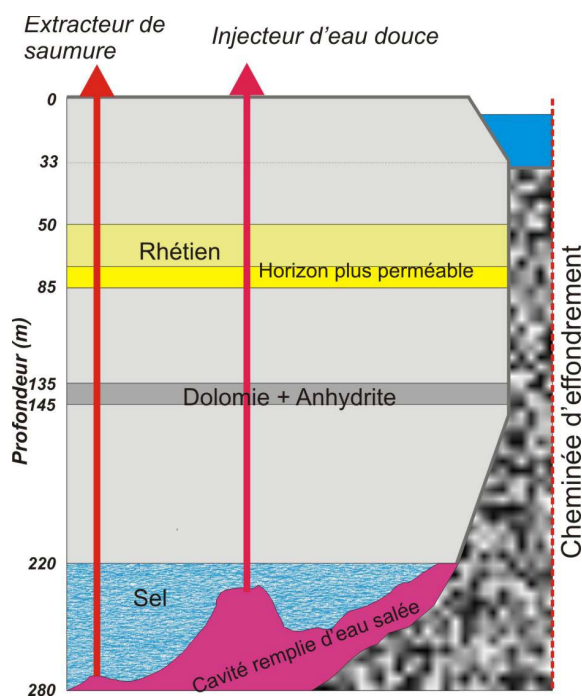
## 5. Les observations sur site et conclusion

La Figure 12 illustre la configuration actuelle du panneau de Gellenoncourt où l'effondrement de deux cavités coalescentes (SG4 et SG5) en 1998 a créé une cheminée débouchant en surface sur un cratère, occupé par un lac dont le diamètre initial était de 50 m, soit une superficie de 0,2 ha. Cette superficie a ensuite évolué avec le temps pour atteindre actuellement 1,7 ha sous l'effet du talutage naturel des terrains. La figure montre que la partie effondrée des cavités se trouve en relation hydraulique directe avec d'autres cavités non effondrées dont le niveau de saumure est contrôlé par des forages d'exploitation. Ce schéma correspond à la configuration no 1 du modèle développé ci-dessus. La simulation indique que le volume annuel de saumure qu'il faudrait évacuer pour empêcher le débordement dans les forages s'élèverait entre 0,1 et 1 m<sup>3</sup> selon l'hypothèse sur le coefficient d'efficacité osmotique, en considérant la valeur de 0,2 ha pour la superficie de la cheminée d'effondrement. De tels volumes passeraient inaperçus face aux différentes opérations de pompage ou de purge réalisées dans le gisement pour les besoins de l'exploitation. On peut noter que les forages du champ de Gellenoncourt sont restés ouverts entre 1998, date de l'effondrement et 2002, date de reprise de l'exploitation et qu'aucune variation de niveau n'y a été constatée.

En conclusion, les évaluations quantitatives qui précèdent sont issues d'une approche théorique du problème. Elles mettent en évidence la sensibilité de

**TABLEAU 2.** Bilan annuel en sel à terminaison d'un panneau d'exploitation effondré sur 36 ha

	Configuration 1		Configuration 2	
	Profil $\varepsilon$ variable 10% → 0,01%	Profil $\varepsilon$ constant 1%	Profil $\varepsilon$ variable 10% → 0,01%	Profil $\varepsilon$ constant 1%
Flux sel diffusif (tonnes/an)	5,4	5,4	5,4	5,4
Flux sel « convectif » (tonnes/an)	6,6	69	0	0
Flux sel total (tonnes/an)	12	74	5,4	5,4
Augmentation de volume de la cavité (m <sup>3</sup> /an)	6,5	40	2,5	2,5
Volume de débordement de saumure (m <sup>3</sup> /an)	21	214	0	0

**FIGURE 12.** Schéma conceptuel de l'exploitation actuelle de Gellenoncourt.

la chemo-osmose au coefficient macroscopique d'efficacité osmotique dont la valeur reste mal appréhendée pour les milieux naturels, aux échelles pertinentes.

Une application de la théorie avec des valeurs de ce coefficient choisies de façon majorante par référence à une argilite compacte bien documentée telle que celle investiguée par l'Andra sur le site du projet de stockage géologique de déchets radioactifs, montre que l'effet osmotique chimique n'est pas susceptible de modifier notablement les ordres de grandeur des flux de sel évacués dans l'environnement

par une exploitation intensive d'une couche salifère, telle qu'envisagé par le groupe Salins à Gellenoncourt.

D'une manière générale, les observations de terrain sur les sites qui exploitent le gisement de sel en couches du plateau lorrain indiquent qu'il n'a jamais été constaté de surpression inexpliquée dans la nappe salée créée par dissolution au sein de la formation salifère alors que des cheminées d'effondrement, comblées de matériaux argileux, mettent en contact la saumure saturée profonde avec les eaux douces du système aquifère superficiel. Ceci est cohérent avec l'approche théorique présentée qui montre que les phénomènes osmotiques ne sont pas susceptibles d'engendrer des perturbations de caractère exceptionnel par rapport aux différentes perturbations inhérentes aux pratiques d'exploitation du gisement de sel par dissolution.

### Conflit d'intérêt

Les auteurs n'ont aucun conflit d'intérêt à déclarer.

### References

- Bresler, E. (1973). Anion exclusion and coupling effects in nonsteady transport unsaturated soils: I. Theory. *Soil Sci. Soc. Am. Proc.*, 37, 663–669.
- Brouard, B., Bérest, P., Hertz, E., and Lheur, C. (2017). A cavern abandonment test at Gellenoncourt, Lorraine—an update. In *Proceedings SMRI Fall Technical Conference, Münster, Germany*, page 14. Solution Mining Research Institute, New York, NY.
- Buffet, A. (1998). The collapse of Compagnie des Salins SG4 and SG5 drillings. In *Proceedings SMRI Fall Meeting, Roma, Italy*, pages 79–105. Solution Mining Research Institute, New York, NY.



- Gonçalvès, J., Adler, P. M., Cosenza, P., Pazdniakou, A., and De Marsily, G. (2015). Semipermeable membrane properties and chemomechanical coupling in clay carriers. In *Developments in Clay Science*, volume 6, pages 269–327. Elsevier, Amsterdam.
- Gonçalvès, J., Rousseau-Gueutin, P., and Revil, A. (2007). Introducing interacting diffuse layers in TLM calculations. A reappraisal of the influence of the pore size on the swelling pressure and the osmotic efficiency of compacted bentonites. *J. Colloid Interface Sci.*, 316, 92–99.
- Gonçalvès, J. and Trémosa, J. (2010). Estimating thermo-osmotic coefficients in clay-rocks: I. Theoretical insights. *J. Colloid Interface Sci.*, 342(1), 166–174.
- Gueutin, P., Altmann, S., Gonçalvès, J., Cozenza, P., and Violette, S. (2007). Osmotic interpretation of overpressures from monovalent based triple layer model, in the Callovo-Oxfordian at the Bure site. *Phys. Chem. Earth*, 32, 434–440.
- Revil, A., Leroy, P., and Titov, K. (2005). Characterization of transport properties of argillaceous sediments: application to the Callovian-Oxfordian argillite. *J. Geophys. Res.*, 110, article no. B06202.
- Rousseau-Gueutin, P. (2008). *Les processus couplés dans les argilites du Callovien-Oxfordien sur le site de Bure : implications pour les mouvements de fluide et de solutés*. PhD thesis, Université Pierre et Marie Curie.





Research article

Geo-hydrological Data & Models

# Scaling of fractured rock flow. Proposition of indicators for selection of DFN based flow models

Philippe Davy<sup>Ⓢ,\*<sup>a</sup></sup>, Romain Le Goc<sup>Ⓢ,<sup>b</sup></sup>, Caroline Darcel<sup>Ⓢ,<sup>b</sup></sup> and Jan-Olof Selroos<sup>Ⓢ,<sup>c,d</sup></sup>

<sup>a</sup> Univ Rennes, CNRS, Geosciences Rennes, UMR 6118, 263 Avenue General Leclerc, Rennes 35042, France

<sup>b</sup> Itasca Consultants S.A.S., 29 Avenue Joannes Masset, Lyon 69009, France

<sup>c</sup> Swedish Nuclear Fuel and Waste Management Company (SKB), Evenemangsgatan 13, Box 3091, Solna SE-169 03, Sweden

<sup>d</sup> Department of Sustainable Development, Environmental Science and Engineering, Royal Institute of Technology (KTH), Sweden

*E-mails:* [davy@univ-rennes1.fr](mailto:davy@univ-rennes1.fr) (P. Davy), [r.legoc@itasca.fr](mailto:r.legoc@itasca.fr) (R. Le Goc), [c.darcel@itasca.fr](mailto:c.darcel@itasca.fr) (C. Darcel), [jan-olof.selroos@skb.se](mailto:jan-olof.selroos@skb.se) (J.-O. Selroos)

**Abstract.** The objective of the paper is to better understand and quantify the flow structure in fractured rocks from flow logs, and to propose relevant indicators for validating, calibrating or even rejecting hydrogeological models. We first studied what the inflow distribution tells us about the permeability structure from a series of analyses: distribution of transmissivities as a function of depth, proportion of flowing sections as a function of section scale, and scaling of the arithmetically-averaged and geometrically-averaged permeability. We then define three indicators that describe few fundamental characteristics of the flow/permeability, whatever the scale: a percolation scale  $l_s$ , the way permeability increases with scale above  $l_s$ , and the variability of permeability. A 4th indicator on the representative elemental volume could in principle be defined but the data show that this volume/scale is beyond the 300 m investigated. We tested a series of numerical models built in three steps: the geo-DFN based on the observed fracture network, the open-DFN which is the part of the geo-DFN where fractures are open, and a transmissivity model applying on each fracture of the open-DFN (Discrete Fracture Network). The analysis of the models showed that the percolation scale is controlled by the open-DFN structure and that the percolation scale can be predicted from a scale analysis of the percolation parameter (basically, the third moment of the fracture size distribution that provides a measure of the network connectivity). The way permeability increases with scale above the percolation threshold is controlled by the transmissivity model and in particular by the dependence of the fracture transmissivity on either the orientation of the fractures via a stress-controlled transmissivity or their size or both. The comparison with data on the first two indicators shows that a model that matches the characteristics of the geo-DFN with an open fraction of 15% as measured adequately fits the data provided that the large fractures remain open and that the fracture transmissivity model is well selected. Most of the other models show unacceptable differences with data but other models or model combinations has still to be explored before rejecting them.

\* Corresponding author.

The third indicator on model variability is still problematic since the natural data show a higher variability than the models but the open fraction is also much more variable in the data than in the models.

**Keywords.** Fracture network, Crystalline rocks, Permeability, Scaling, Indicator, Percolation.

*Manuscript received 27 April 2022, revised 12 October 2022, accepted 17 October 2022.*

## 1. Introduction

In naturally fractured geological reservoirs, *fluid-flow structure* is a major question for hydrogeological modeling [Neuman, 2005]. Especially for the crystalline rocks targeted for energy recovery exploitation, or geological storage of industrial wastes like CO<sub>2</sub> or spent nuclear fuel, fluid flow only takes place in a tiny part of the whole rock, through the connected network of fractures between hydraulically active boundaries, while intact rock matrix permeability is considered as negligible. Fluid flow structure is hence a subpart of the fractured system.

While *fractures are ubiquitous* in crystalline rocks, the complexity of their spatial organization precludes the use of classical methods of homogenization to derive hydraulic properties for equivalent continuous media. Fracture systems result from long-standing geological histories and fracturing processes, leading to a complex multiscale structure. Fracture-size scales commonly cover several orders of magnitudes from fractures smaller than micrometers to millimeters to tectonic faults larger than hundreds of meters to kilometers with fracture size distributions adequately modeled by power-law distributions [Bonnet *et al.*, 2001, Davy *et al.*, 2013]. Their capacity to be hydraulically active is dependent on their topological structure (connectivity, intersections), their openness and transmissivity. These factors together contribute to present a very large variability—in space and intensity—of the inflows that can be measured at depth from flow logging [Follin, 2008]. In parallel, usual data acquisition capacities are far from what would be needed to easily reduce modeling uncertainties for typical site conditions. These aspects contribute to the difficulty to define relevant models of the fluid-flow structure for naturally fractured rocks and emphasizes the need to model the discrete nature of the flow before deriving any homogenized properties.

*Discrete Fracture Network (DFN) approaches* emerged in the late eighties [Cacas *et al.*, 1990a,b, Long and Witherspoon, 1985, Long and Billaux,

1987]. Since then, they are widely used and developed, most often for applications that involve hydrogeological and hydro-mechanical modeling, e.g. [Davy *et al.*, 2006, 2018, De Dreuzy *et al.*, 2013, Lei *et al.*, 2017, Park *et al.*, 2002, Selroos *et al.*, 2022]. As they realistically reproduce the discrete nature of the flows—typically spacing between inflows ranging from a few meters to tens of meters along wellbore—they are the most suitable for understanding flow structures in-situ. They also permit aggregation of various types of data (geological, mechanical, hydraulic, etc.) in a multidisciplinary approach where the DFN is the unifying element, as part of a strategy to minimize modeling uncertainties [Selroos *et al.*, 2022].

The major theoretical questions for hydrogeological modeling of fractured media, which directly drive and impact the flow structure modeling, are scale dependence and upscaling or homogenization. The issue of the *scale dependence* of the hydraulic properties is widely debated. One of the first compilations on that subject is the seminal paper of Clauser [1992] which compiled permeability measurements from lab to regional scales, emphasizing a permeability increase with scale increase up to a hundred meters. This conclusion has been questioned by Hunt [2003b], who posited that although experiments often indicate an increase in the hydraulic conductivity with increasing scale, this effect may be explained by sampling issues rather than intrinsic scale effect. Other studies, based on field observation [Illman, 2006, Neuman, 2003, Neuman and Di Federico, 2003, Ren *et al.*, 2021] or modeling [de Dreuzy *et al.*, 2001b, 2002] are in line with the observations from [Clauser, 1992] and a scale effect emerging from the fracture network connectivity structure and transmissivity distribution. Illman [2006] also observed a directional permeability scale effect in cross-hole test results and hypothesizes that the effect is controlled by the connectivity of fluid conducting fractures. A scale increase of effective permeability with scale is also shown in Martinez-Landa and Carrera [2005] and Ren *et al.* [2021] from single hole and cross-hole

pumping tests data and identify a scale dependence likely site-dependent too. de Dreuzy *et al.* [2001b, 2002] use a DFN model (in 2D) for which the fracture length and fracture aperture distributions are power laws and study the resulting hydraulic properties as the flow structure and equivalent permeability. They determine the equivalent permeability scale effects, characterize the flow structure by a channeling indicator, and establish the relationship with the fracture distributions parameters. *Upscaling* permeability is an active research area in hydrogeology both for fractured and unfractured media. Upscaling consists of deriving the equivalent permeability from the smaller scale permeability distribution [de Dreuzy *et al.*, 2010]. Wen and Gómez-Hernández [1996] and Renard and Marsily [1997] reviewed the state of the art on upscaling conductivities in heterogeneous media. Oda [1985] developed an analytical approach to define equivalent permeability from a DFN description based on geometric characteristics of fractures with arbitrary orientations. de Dreuzy *et al.* [2001b, 2002] develop the relationship between equivalent permeability and the parameters of multiscale DFN models in 2D and hence the scale dependency. Chen *et al.* [2015, 2018] develop numerical approaches to compute the equivalent permeability from a DFN description.

The question investigated in this paper is how to best use available hydro data for capturing the multiscale nature of in-situ flow structures and to derive manageable and relevant metrics for site modeling purposes. At the Forsmark area in Sweden, site investigations have been performed by Svensk Kärnbränslehantering AB (SKB), the Swedish Nuclear Fuel and Waste Management Company, for more than two decades in view of the future deep repository for spent nuclear fuel [Follin, 2008, Follin *et al.*, 2007, Selroos *et al.*, 2022]. Follin *et al.* [2014] use flow logs and metrics based on the relative proportion of sealed, open, and flowing fracture frequencies together with equivalent transmissivity distributions to calibrate the parameters of DFN models for flow simulations. Maillot *et al.* [2016] compare the performance of different DFN models from the comparison of channeling and equivalent permeability indicators. Zou and Cvetkovic [2020, 2021] also use DFN models and varying hydraulic boundary conditions to simulate steady-state pumping tests and derive indicators of cumulative and of complementary cumulative

distributions to evaluate the impact of fracture in-plane heterogeneities on the flow structure at the DFN scale.

The first objective of this study is to develop flow-based metrics adapted to in-situ flow testing capacities and suited to emphasize the multiscale nature of the flow structure. To address these issues, the database of fractures and discrete inflows at Forsmark, which includes flow logging and steady-state pumping tests, was analyzed with metrics revisited from the former work of Maillot *et al.* [2016]. The second objective is to test the capacity of the latter to discriminate between DFN models otherwise calibrated to similar geological and geometrical observations. A modeling environment is set up in the software “DFN.lab” to generate DFN realizations from a predefined range of DFN models and perform numerical flow simulations in steady-state pumping configuration simulations.

The paper is organized as follows: the data and conditions of the Forsmark site are recalled in Section 2. The framework to define the metrics is presented in Section 3. Data analyses are presented in Section 4. In Section 5 the DFN models are defined and finally, several a priori equiprobable DFN models are compared with each other and with the field data and the interpretations and comparisons between data and models are presented. Results and outcomes of the study are finally discussed in Section 6 and summarized in the conclusion. To facilitate ease of reading this study, specific parts have been moved to appendices.

Note that, while the analytical tools and methodology are generic, the results are specific to the studied area (mostly the Nordic granites). It is not presumed that thresholds and trends are generalizable to all crystalline rocks [e.g., Dewandel *et al.*, 2006].

## 2. Site and available data overview

The Forsmark site in Sweden, located in crystalline bedrock in south-eastern Sweden and located roughly 130 km north of Stockholm, is selected by the Swedish regulatory authorities to build a nuclear waste repository for spent nuclear fuel at 470 m depth [SKB, 2011]. Svensk Kärnbränslehantering AB (SKB) is in charge of site investigations, future construction, and long-term safety assessment of the final repository. The database built over the course

of the site investigations encompasses data from geological, mechanical, hydrogeological, geochemical, etc., aspects. Several tens of boreholes each reaching lengths up to about 1 km have been drilled and exhaustively logged.

The presented analyses relies on a detailed description of the fracture transmissivity structure in the different boreholes that have been extensively carried out at the Forsmark site by SKB with the Posiva Flow Logs (PFL). PFL was originally developed by Posiva Oy to meet the demand on flow measuring techniques adapted to sparsely fractured rocks and low-permeability environments [Öhberg and Rouhiainen, 2000]. The method couples a very low threshold for flow detection (down to 0.1 l/min) and a precise positioning of the device that allows them to detect single fractures inflow with a depth resolution of a couple of centimeters. The hydraulic properties of every cored borehole at the Forsmark site are hence tested [Rouhiainen and Pöllänen, 2003, Rouhiainen *et al.*, 2004] by performing a standard hydraulic test where the applied pressure head and resulting flow are recorded. The derivation of a transmissivity is based on the Thiem (or Dupuits) solution for steady-state radial flow without skin effect, so that the flow rate  $Q_1$  at a borehole test section of size  $L$  is equal to:

$$Q_1 = K \frac{2\pi L}{\ln\left(\frac{R}{r_0}\right)} (h_0 - h_1) \quad (1)$$

with

$Q_1$ : measured flow rate in the test section ( $\text{m}^3/\text{s}$ )

$K$ : hydraulic conductivity of the test section ( $\text{m}/\text{s}$ )

$L$ : length of the test section limited by packers ( $\text{m}$ )

$R$ : radius of influence (500 m) ( $\text{m}$ )

$r_0$ : radius of the borehole ( $\text{m}$ )

$h_0$ : undisturbed hydraulic head far from test section ( $\text{m}$ )

$h_1$ : pump-induced or natural borehole hydraulic head ( $\text{m}$ ).

The two unknowns  $K$  and  $h_0$  can be deduced from two pairs of hydraulic heads ( $h_1$  and  $h_2$ ) and resulting flow rates ( $Q_1$  and  $Q_2$ ) provided that  $K$  is not dependent on the induced pressure (no hydromechanical effects).

A combination of sequential and overlapping flow logging allows identifying water conductive fractures down to a spatial resolution of 0.1 m and flow rates down to  $30 \text{ ml}\cdot\text{h}^{-1}$ . This leads to a measurement limit

of approximately  $T = 10^{-10} \text{ m}^2/\text{s}$  when expressed in equivalent transmissivity as defined below.

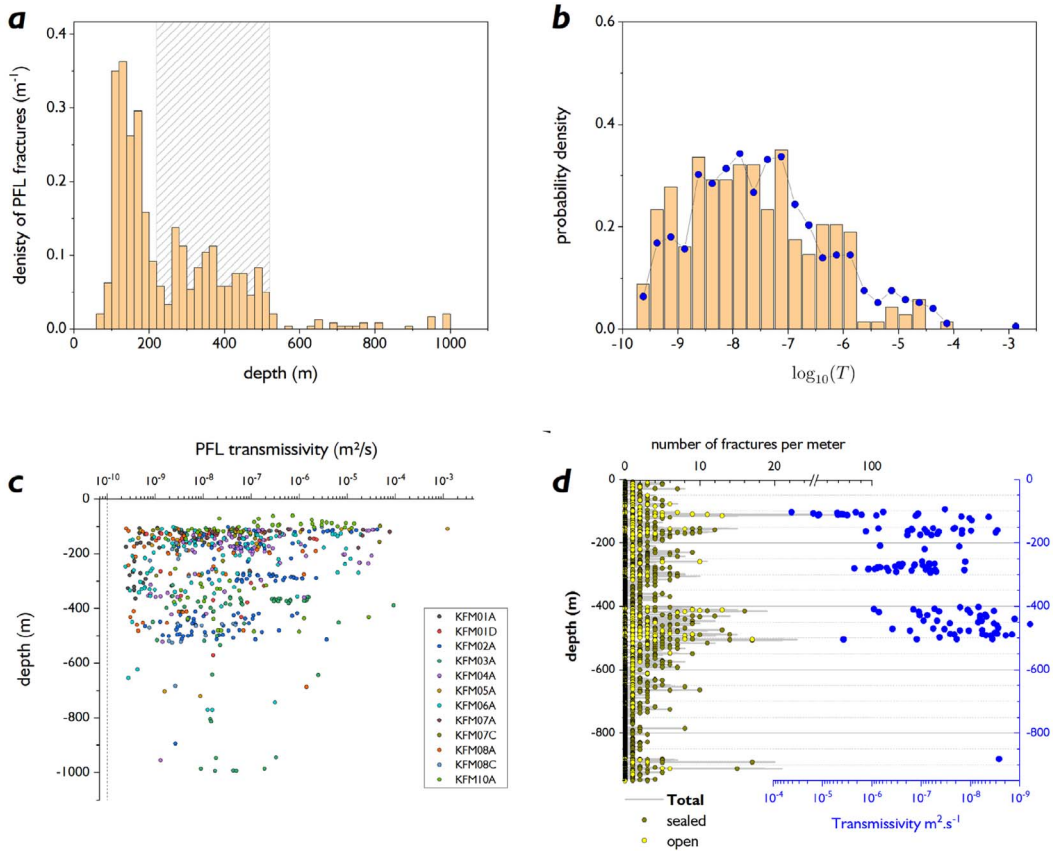
The transmissivity of a test section, or the fracture identified by the sequential mode, is defined as:

$$T = KL = \frac{\ln\left(\frac{R}{r_0}\right)}{2\pi} \cdot (Q_2 - Q_1)/(h_2 - h_1) \quad (2)$$

with  $R \sim 500 \text{ m}$  it comes that  $T \sim \Delta Q/\Delta h$ . The ratio  $\Delta Q/\Delta h$  is also called a specific capacity (in  $\text{m}^2/\text{s}$ ). Martin and Follin [2011] rightly note that: "The PFL transmissivity value ( $T$  in (2)) reported to SICADA (the SKB database) will be close to the local value at the borehole (i.e., the transmissivity of the intersecting flowing fracture), if the local value is less than the overall transmissivity of the network to which it is connected. If the local value at the borehole is greater than the overall transmissivity of the network to which it is connected, the PFL transmissivity value represents a "hydraulic choke" (bottleneck) phenomenon, which means that it is not the transmissivity of the intersecting flowing fracture that is determined". The PFL transmissivity  $T$  (Equation (2)) is not the actual fracture transmissivity, but rather the transmissivity of the closely connected structure. Other limitations of (1) and (2) are that  $R$  in (1) and (2) varies with time [Perrochet, 2005] and, in case of several fractures,  $T$  or  $K$  tends towards the geometric mean [Meier *et al.*, 1999]. Thus  $T$  must be taken as an estimate rather than an exact value of the fracture transmissivity, which is correct considering the high variability of transmissivities (i.e., the distribution of the estimate gives a faithful measure of the transmissivity distribution).

In practice, the same set of flow rate measurements is done twice all along a tested borehole: the first pass is for natural flow rate conditions (no pumping), and the second one is for a typically imposed drawdown of  $\Delta h = 10 \text{ m}$ .

The distribution of PFL transmissivities for all wells is given in Figure 1b. The log average is about  $3 \times 10^{-8} \text{ m}^2\cdot\text{s}^{-1}$  with values ranging from  $2.5 \times 10^{-10}$  to  $1.3 \times 10^{-3} \text{ m}^2\cdot\text{s}^{-1}$ . The average intensity of PFL log-transmissivities does not vary significantly with depth (Figure 1c) in contrast to the density of PFL fractures, i.e., the number of fractures where a transmissivity has been detected, and to the transmissivity variability (Figure 1a). From these observations, we identify three main domains: the near surface (0–200 m), the domain of interest for the repository



**Figure 1.** (a) (top left) Number of PFL fractures (i.e., fractures whose transmissivity is detectable with the PFL flow log) per 20 m depth increment. (b) (top left), Distribution of the decimal logarithm of PFL transmissivities in  $\text{m}^2\cdot\text{s}^{-1}$ ; the bars are for the whole domain (100–1000 m) and the solid blue dots for the target domain between 220 and 520 m. No value smaller than  $10^{-9.5}$  are observed; the impact of this threshold, although very low, is discussed in Section 4.3 and in Figure 6. (c) (bottom left), PFL transmissivity as a function of depth for all the boreholes. (d) (bottom right), Example of a borehole logging (KFM02A) showing the measured density of fully intersecting fractures (number per borehole length with sealed and open densities indicated by the solid and yellow dot, respectively) on the left and the measured PFL transmissivity on the right.

(200–500 m), and a deeper domain (>500 m). The near surface is highly fractured and highly permeable; the density of PFL fractures is 1 PFL every 5 m on average (Figure 1a). It is characterized by horizontal sheet joints, which makes them distinct from the deeper parts. The intermediate domain is rather homogeneous with a density of PFL fractures about half the near surface. The deep domain is very dry with only a few PFL fractures, on average 1 every 2T50 m.

The spatial distribution of transmissive fractures varies from one borehole to another but with

common features as clustering in flow zones, presence of large domains without detectable flowing structures, and positive correlation between zones of high fracture density and zones of large PFL transmissivities (Figure 1d). Note that the correlation between fracture density and PFL transmissivity is far from perfect and many counterexamples exist where PFL fractures are observed corresponding to zones of normal fracture density.

In the following, we focus on the domain of interest for the waste repository. We arbitrarily set the upper and lower depths to 220 m and 520 m,

respectively, in accordance with the density of the PFL fractures observed in Figure 1a. The distribution of PFL transmissivity between 220 and 520 m is not different from that observed on the whole domain (100–1000 m), despite the variability in the density of PFL fractures (Figure 1b).

### 3. Scaling metrics for flow log interpretation

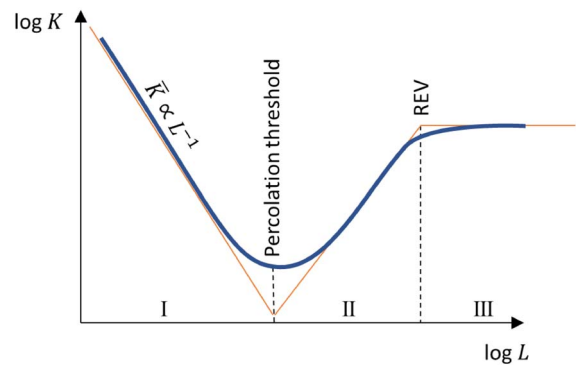
In what follows, an interpretation framework based on the PFL data, including the transmissivity values and their spatial distribution is introduced. This is done through the prism of scale analysis, where the in-situ scale-dependent permeability is defined and evaluated for domains of increasing size.

The primary component of the analysis is the total inflow through a given core log section:

$$K_i = \frac{1}{L_i} \sum_{j=1}^{n_i} T_{i,j} \quad (3)$$

where  $K_i$  is the equivalent hydraulic conductivity (in (m/s)) of the core section  $i$  of length  $L_i$ , into which there are  $n_i$  positive inflows interpreted as transmissivities noted  $T_{i,j}$ , from (2). The core section length  $L_i$  will be used to analyze the scale dependency of  $K_i$  in order to highlight fundamental properties of the site permeability, and in particular, the scaling laws that govern it. The scale analysis of hydraulic conductivity relies on the continuous records of all the discrete inflows (the PFL transmissivities as recalled in Section 2) over an entire borehole. The maximum investigated scale of the analysis is the length of the borehole in the investigated domain. For a given scale  $L$  (i.e., all section length  $L_i = L$  in (3)) and a given borehole, the arithmetic average of  $K_i(L)$  is exactly the hydraulic conductivity of the borehole if there is no overlap between the different sections, or a very close value, if there is some. In Appendices A and B, we argue from numerical simulations that this borehole conductivity is representative of the permeability of a 3D system with a typical size equal to the section length. Hereafter, we use the term permeability to refer to  $K_i$ .

To reveal the flow network organization, we focus on the distribution of “wet” sections (i.e., sections containing at least one inflow), as it is done when determining the fractal properties of a network [Tél et al., 1989]. Whether a section is dry or wet depends on the resolution of the inflow measurement. Even if



**Figure 2.** Schematic evolution of the geometric average of the permeability  $K$  as a function of the measurement scale  $L$  (solid blue line). The thin red segments indicate the three main regimes: below the percolation threshold scale  $p_c$  (I), between  $p_c$  and the representative elementary volume (REV) (II), and above the REV (III) [Charlaix et al., 1987, de Dreuzy et al., 2001b, 2002]. See text for an explanation of the three regimes.

the PFL instrument is extremely sensitive, it is clear that the transmissivity distribution is not complete and truncated at the PFL resolution (Figure 1b). This issue will be discussed further down in the text.

By taking the geometric average of the section permeability, as in de Dreuzy et al. [2002], we expect to identify some important characteristics of the flow structure (Figure 2): the percolation threshold scale  $p_c$  (transition between regime I and II in Figure 2), the permeability increase above  $p_c$  (regime II), and possibly a Representative Elementary Volume (REV) that indicates the scale above which the permeability no longer varies. The correspondence between the percolation threshold and a scale was established for fracture networks with a power-law size distribution [Berkowitz et al., 2000, Bour and Davy, 1997, 1998, Darcel et al., 2003b, de Dreuzy et al., 2000]. Compared to networks of constant size, the percolation threshold reflects the connectivity of small fractures and the probability of encountering fractures larger than the size of the systems that connect the networks on their own. It is this dual connectivity that makes connectivity scale-dependent and gives a correspondence between the percolation threshold and a given scale. Below the percolation threshold, the wet sections are likely containing one main frac-



ture/channel with a permeability that decreases as  $\bar{K} = T_o/L$  where  $T_o$  is the channel transmissivity and  $L$  the section length. The permeability increase above the threshold has been demonstrated by [Charlaix et al., 1987, de Dreuzy et al., 2002]. It reflects the distribution of channel transmissivities. Above  $p_c$ , the scale increase results in a merging of different channels with different transmissivities, and the permeability increase with scale reflects the broadness of the transmissivity distribution [de Dreuzy et al., 2001b, 2002]. Depending on the transmissivity distribution, it may exist a REV over which systems can be handled as homogeneous media. The REV is a correlation length of the transmissivity distribution but its existence is questioned if the distribution of fracture transmissivities and lengths is too broad [de Dreuzy et al., 2001b].

To calculate  $K(L)$ , we subsample each borehole with section lengths smaller than the total borehole length obtaining a set of equivalent hydraulic conductivities  $K_j$  (Equation (3)). We note  $N_{>0}(L)$  the number of core sub-sections containing at least one positive inflow and hence a positive  $K_j$ . The first proposed metric, the function  $K_a(L)$ , is the arithmetic average of the above-defined equivalent permeabilities derived from the ensemble of flowing core sections (non-flowing core sections are discarded) for a size scale  $L$ , as summarized in the equation below:

$$K_a(L) = \frac{1}{N_{>0}(L)} \sum_{i=1}^{N_{>0}(L)} K_i(L). \quad (4)$$

Since the permeability arithmetic average including zero-flow sections  $\bar{K}_a$  is independent of  $L$  (see section following (3)),  $K_a(L)$  varies as the inverse of the percentage of flowing sections:

$$K_a(L) = \frac{N(L)}{N_{>0}(L)} \bar{K}_a. \quad (5)$$

The second proposed metric, the function  $K_g(L)$ , is the geometric average of the same set of elementary bricks:

$$K_g(L) = \exp\left(\frac{1}{N_{>0}} \sum_{i=1}^{N_{>0}} \log(K_{(i)}(L))\right). \quad (6)$$

This metrics has been proposed by de Dreuzy et al. [2002] as the most representative mean for a permeability distribution that tends to be log-normally distributed (Figure 1).

Both  $K_a(L)$  and  $K_g(L)$  are likely to be dependent on the inflow detection limit estimated at

$10^{-10} \text{ m}^2 \cdot \text{s}^{-1}$  for PFL measurements. Even if this value is very low, the transmissivity distribution in Figure 1b suggests that smaller values exist. By artificially increasing the detection limit, we test in the following section that the metrics do not depend too much on its value as long as it remains much in the lower end of the transmissivity distribution.

To characterize the flow channeling, we also calculate the average number of inflows in all flowing sections of length  $L$ ,  $n(L)$ , and a channeling metric  $n_Q(L)$  derived from [Maillot et al., 2016]:

$$n_{Q,i} = \frac{\left(\sum_{j=1}^{n_i} T_{ij}\right)^2}{\sum_{j=1}^{n_i} T_{ij}^2} \quad (7)$$

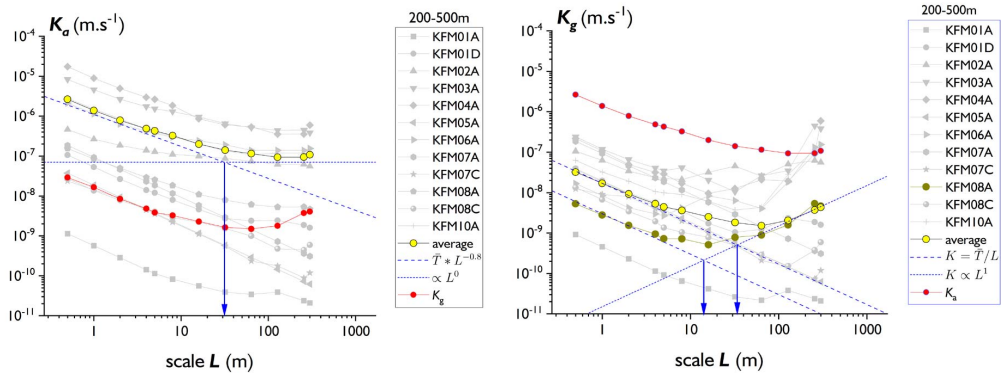
where  $i$  refers to the core section  $i$  that contains  $n_i$  inflows noted  $T_{ij}$ , over the section size  $L_i$ . The quantity  $n_Q(L)$  is then the arithmetic average over all flowing core sections of size  $L$ , and it can be viewed as the number of efficient channels in a flowing core section. If all inflows have the same intensity,  $n_Q$  is equal to the number of inflow  $n$ , however, in the general case with a distribution of inflow intensity,  $n_Q$  is smaller than  $n$ . The ratios  $n/L$  and  $n_Q/L$  have the same meaning as the number of fracture intersections per unit core length, classically referenced as  $p_{10}$ , but only for flowing ( $n$ )—or significantly flowing ( $n_Q$ )—fractures.

## 4. Hydraulic data interpretations

In this section, we compute the different metrics and indicators defined in the previous section for the datasets listed in Table 1 in the depth range 220–520 m (see Section 2 for an explanation of this depth range).

### 4.1. Permeability averages

The scale evolution of equivalent arithmetic and geometric permeabilities are plotted in Figure 3. From borehole to borehole and comparable size scale, the calculated values are spread over 4 orders of magnitude. Despite this large dispersion, all the plots display a similar shape with a  $L^{-1}$  decrease over about 1 to 2 orders of magnitude when increasing scales until a stable (for  $K_a$ ) or increasing (for  $K_g$ ) regime is reached with large-scale permeabilities that range between  $2 \times 10^{-10}$  and  $4 \times 10^{-7} \text{ m} \cdot \text{s}^{-1}$ . The transition



**Figure 3.**  $K_a(L)$  (left) and  $K_g(L)$  (right) computed for each borehole dataset listed in Table 1 (grey symbols) as a function of scale. For both, the yellow dots indicate the average of all boreholes and the red dots, the permeability average (i.e.,  $K_a$  for the  $K_g$  graph and vice versa). The blue lines represent power-law fits at small and large scales, respectively. The dark yellow dots in the  $K_g$  plot are for KFM08A, a borehole that will be discussed in the modeling section.

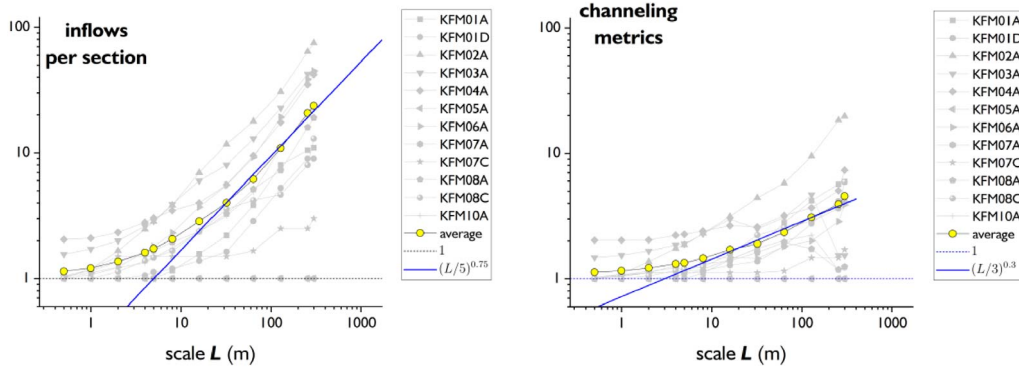
**Table 1.** List of boreholes (IDCODES in the SICADA database) with PFL data including the total number of inflows, the depth of the uppermost and lowermost inflow, the smallest and the largest PFL transmissivity

IDCODE	Number of inflows ( $T > 0$ )	Uppermost inflow (m)	Lowest inflow (m)	Min( $T$ ) ( $\text{m}^2/\text{s}$ )	Max( $T$ ) ( $\text{m}^2/\text{s}$ )	Total logged length (m)
KFM01A	34	105.3	363.4	$2.45 \times 10^{-10}$	$5.31 \times 10^{-8}$	891
KFM01D	34	106	571.17	$6.59 \times 10^{-10}$	$2.30 \times 10^{-6}$	708
KFM02A	125	101.8	894	$6.16 \times 10^{-10}$	$4.21 \times 10^{-5}$	899
KFM03A	71	106.4	994	$8.90 \times 10^{-10}$	$9.21 \times 10^{-5}$	896
KFM04A	142	109.6	954.8	$7.16 \times 10^{-10}$	$3.54 \times 10^{-5}$	876
KFM05A	27	108.9	720	$4.45 \times 10^{-10}$	$1.23 \times 10^{-3}$	897
KFM06A	99	102.4	770.8	$2.40 \times 10^{-10}$	$1.92 \times 10^{-5}$	895
KFM07A	23	110.8	261.4	$3.57 \times 10^{-9}$	$7.67 \times 10^{-5}$	892
KFM07C	15	98.4	279.8	$8.99 \times 10^{-10}$	$4.81 \times 10^{-5}$	413
KFM08A	41	107.6	687	$2.48 \times 10^{-10}$	$2.20 \times 10^{-6}$	846
KFM08C	21	102.4	683.6	$7.51 \times 10^{-10}$	$2.95 \times 10^{-6}$	848
KFM10A	56	60.3	484.4	$1.64 \times 10^{-9}$	$2.79 \times 10^{-5}$	437

between both regimes is at about 30 m for both permeability averages.

The evolution of  $K_a(L)$  and  $K_g(L)$  reflects the organization of flows along the boreholes. In the first regime, as long as the length of the core sections is smaller than the minimum distance between two inflows, every element contains only one inflow and hence the equivalent permeability is inversely proportional to the core section length:

$$K_j = \frac{T_j}{L}. \quad (8)$$

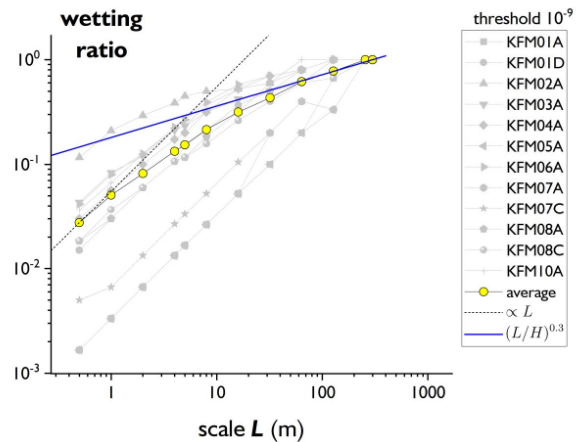


**Figure 4.** Plot of the average number of inflows per section (left) and of the channeling metrics defined in (7) (right) as a function of section scale.

As expected, the arithmetic average of  $K_j$  is larger than the geometric average but the scaling trends are similar for this small-scale regime. We note that, for  $K_a$ , the average of all boreholes (yellow dots) seems to decrease slightly less than  $L^{-1}$ ,  $\propto L^{-0.8}$ . This reflects a difference between the least and most permeable boreholes; the latter (KFM02A, KFM03A, KFM04A, KFM10A) having a less steep decrease with scale than the former. A basic explanation is that a high permeability is also associated with a higher density of inflows. This is not observed for the borehole-averaged  $K_g$  because the geometric average gives a more distributed weight to all boreholes whatever their permeability.

When the section length is large enough to gather several inflows, the permeability stabilizes with scale for the arithmetic average  $K_a$ , or grows as  $\sim L^{0.7}$  for the geometric average  $K_g$  as the result of the transmissivity distribution. The transition between both regimes spans over one order of scale magnitude between  $\sim 6$  and  $\sim 60$  m with a critical value at 30 m determined from the intersection of both end-member regimes (Figure 3). There is no limit to the increase of  $K_g$  with scale in this regime for the observed scale range (30–300 m), which means that a REV such as anticipated in Figure 2 is beyond 300 m, if it exists.

The average number of inflows per section confirms the trends observed on the permeability averages (Figure 4, left). For sections smaller than 4–5 m, only one inflow is observed on average per wet section, and then the number increases as  $L^{0.75}$ . The maximum number of inflows for the largest section of 300 m varies between 10 and 80 with an average



**Figure 5.** Evolution of the wetting ratio, i.e., the proportion of “wet” sections, i.e., containing at least one inflow, with section length scale  $L$  for all the boreholes (grey symbols) and for the average over all the boreholes (yellow dots). The scaling is  $L$  (dashed line) and  $L^{0.3}$  (solid blue line).

of 20, except for the two boreholes KFM01A and KFM07C where only 1 and 2 inflows are recorded, respectively. We also calculate the channeling metrics  $n_Q$  (Equation (7)), which quantifies the number of main channels in each section.  $n_Q$  shows the same two-regime scaling trend as the number of inflows per section but with smaller values, as expected. It increases above 4–5 m approximately as the square root of the number of inflows. For the largest section (300 m), only 3 main channels on average are

detected.

and it can be used to identify a potential fractal nature of the flowing structure as it is classically done in box-counting methods [Mandelbrot, 1982, Pavón-Domínguez and Moreno-Pulido, 2022]. In a fractal of dimension  $D_{1D}$  in 1D, the number of wet sections decreases as  $(L/H)^{D_{1D}}$  with  $H$  the investigated borehole length (here 300 m), and the wetting ratio increases as  $r_w \sim (L/H)^{1-D_{1D}}$ .  $D_{1D} = 0$  is a point-like structure in 1D corresponding to a plane in 3D. Figure 5 shows the two regimes already described: for scales smaller than  $\sim 30$  m, the wetting ratio increases as  $L$  corresponding to individual flow planes. Above 30 m,  $r_w = (L/H)^{0.3}$  indicating a fractal structure with a 1D dimension of 0.7 (2.7 in 3D). This complex structure and the distribution of transmissivity are likely to be responsible for the increase in the geometric average of permeability.

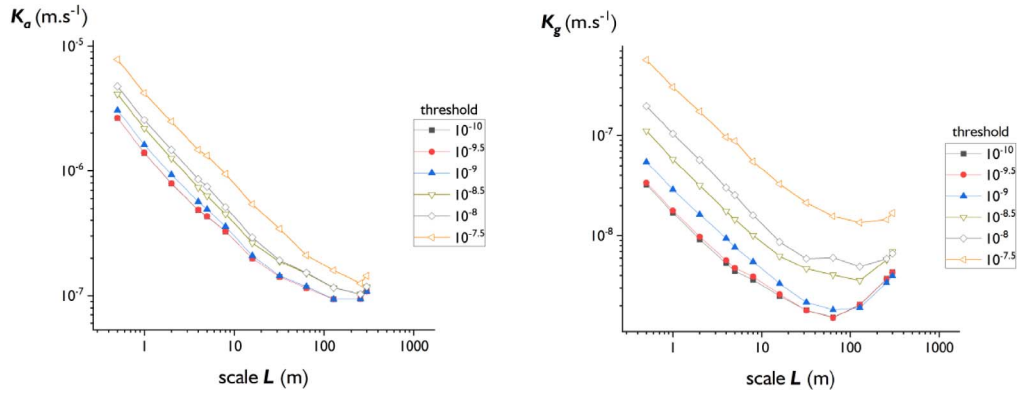
#### 4.2. Flow structure organization

We plot in Figure 5 the evolution with  $L$  of the ratio of the proportion of “wet” sections—i.e., number of sections that have at least one inflow divided by the total number of sections—with scale. This ratio, hereafter called the *wetting ratio*  $r_w$ , controls the evolution of  $K_a(L)$  (see discussion above and (5))

#### 4.3. Dependency on the detection threshold of transmissivity

The previous analysis is potentially dependent on the flow detection resolution to identify wet and dry (no-flow detected) sections. Both  $K_a$  and  $K_g$  rely on the permeability of “wet” sections defined as the dual of no-flow sections. If inflows smaller than the detection limit of transmissivity exist, the number of wet sections will increase and the average permeability decreases. Even if the detection limit is very small, it does not seem to represent the lower bound of the transmissivity distribution (Figure 1b).

Considering that we cannot decrease the measurement resolution, we first apprise the reader that the previous scaling analysis is performed with a transmissivity threshold of  $10^{-10} \text{ m}^2 \cdot \text{s}^{-1}$  and that the results may depend on this threshold. Then, we performed the same analysis as above but with different detection thresholds of transmissivity (Figure 6). The curves are identical for detection threshold ( $T_d$ ) up to  $10^{-10} \text{ m}^2 \cdot \text{s}^{-1}$  and similar, i.e., presenting the same log-shape but with a higher permeability for  $T_d$  up to  $10^{-8} \text{ m}^2 \cdot \text{s}^{-1}$ , which is the mode of the probability distribution of log-transmissivity. We thus conclude that the detection threshold is likely to be small enough to allow for a relevant scaling analysis of the flow structure, but we will be certain of this conclusion only if analyses were carried out with an even lower detection limit.



**Figure 6.** Evolution of the permeability averages  $K_a$  (left) and  $K_g$  (right) as a function of section scales for different transmissivity thresholds (see text).

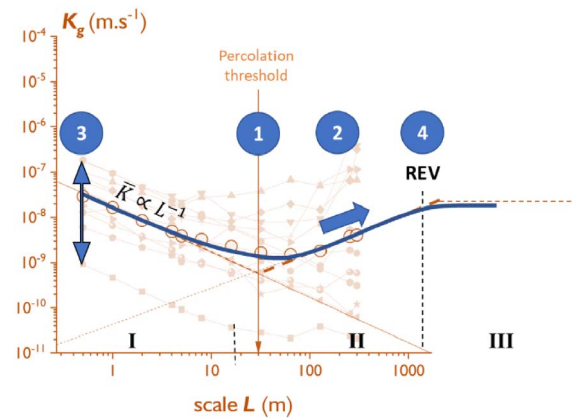
#### 4.4. Flow structure indicators

From the previous analysis, we identify a series of flow structure indicators which could serve as a basis for testing, and calibration of validating hydrogeological models. The indicators are best chosen to be characteristics of the fractured media that are likely to be controlling the flow properties, that can be measured, and whose measure is not too much affected by the conditions of the experiment. The 4 indicators listed below and in Figure 7 are chosen to be the measures of transmissivity distribution and flow structure.

For each borehole, the equivalent geometric permeability average (6) with scale  $L$  can be characterized by a V shape (Figure 7) with a scattering from one borehole to another. These trends are basic to the definition of our 4 indicators.

The first indicator is the scale at which network percolates, which is about 30 m in the studied area. This scale also represents the average distance between inflows or inflow clusters. The evolution of  $K_g(L)$  below the percolation is representative of sections in which the number of inflows does not increase with scale, which explains the  $L^{-1}$  scaling.

The second indicator is the increase of the geometric permeability average with scale above the percolation threshold. It reflects the spatial distribution of inflows and the probability to encounter a flow path of large intensity when increasing the section scale. The slope of this increase is characteristic of the flow organization, a steeper slope indicating a higher heterogeneity of flows. Note that the increase of permeability with scale is long debated in the hydrogeological literature [Hunt, 2003a,b, Illman, 2006, Neuman, 2005, 1994, 2003, Neuman and Di Federico, 2003].



**Figure 7.** Representation of the 4 indicators (number in circles) described in the text with a graph which superposes the data of Figure 3 and the schematic scaling evolution of Figure 2. The first is the percolation threshold scale, the second is the permeability increase above the percolation threshold, the third is the distribution of transmissivities, and the fourth is the REV scale.

drogeological literature [Hunt, 2003a,b, Illman, 2006, Neuman, 2005, 1994, 2003, Neuman and Di Federico, 2003].

The third indicator is the dispersion of the borehole curves around the  $L^{-1}$  decrease. From borehole to borehole and comparable size scale, the calculated values are spread over 3 orders of magnitude (Figure 7). At sizes  $L$  smaller than the minimum distance between two inflows  $d_{\min}$ ,  $K_g(L)$  is simply equal to

the geometric average of the set of individual transmissivities divided by  $L$ :

$$K_g(L)_{L < d_{\min}} = \frac{1}{L} \times \left( \prod_{i=1}^N T_i \right)^{\frac{1}{N}} = \frac{\overline{T}_g}{L}. \quad (9)$$

A borehole of total investigated length  $H$  is thus characterized by a set of  $N$  transmissivities  $T_i$  drawn from the full transmissivity distribution of the whole site with a geometric average  $\overline{T}_g$ . Both  $N$  and the distribution of  $T_i$ —and thus  $\overline{T}_g$ —are indicators of the full transmissivity distribution. Note that the distribution of  $\overline{T}_g$  must vary with the investigated borehole length  $H$  around the geometric average of  $\overline{T}_g$  (yellow dots in Figure 3).

In theory, it exists a fourth indicator which is the representative elementary volume (REV), above which the permeability becomes scale independent. Figure 3 shows that the REV scale, if it exists, is larger than the investigated section length of 300 m.

## 5. Calibration/validation of DFN models

This section aims at testing the relevance of the indicators to calibrate, validate or reject DFN models. It should be taken as an illustration of the potential of the indicators and not as an in-depth analysis of the modeling.

The DFN modeling process presented here is in line with the DFN methodology defined in Selroos *et al.* [2022], where the definition of a DFN model for a hydrogeological application involves several steps. In the first, geological and geometrical data are used to define the DFN model of all the fractures, whatever their internal properties may be (by convenience hereafter referred to as geo-DFN). The second one is an in-between between geometric and hydraulic modeling steps. It consists in defining the DFN model of all open and partly open fractures, as a subpart of the geo-DFN model (by convenience hereafter referred to as open-DFN). Before even assigning transmissivity properties to the fractures of the open-DFN model, and hence defining the hydro-DFN model, the structure of the open-DFN model is critical for the flow: having a connected path of open fractures between hydraulically active boundaries is the prerequisite to define hydraulic properties. In each step of the modeling process, data are integrated with prior models and hypotheses and eval-

uated in a cycle of Sensitivity Analyses, Calibration, and Rejection tests Selroos *et al.* [2022].

In this section, we present only a small part of this modeling process which focuses on the integration of hydrological data analyzed in Section 4. The other parts are only briefly described. The analysis focuses on the fracture domain FFM01 [Follin *et al.*, 2014, Olofsson *et al.*, 2007], which is the target geological formation for the spent nuclear fuel repository [SKB, 2011]. Models are compared with the borehole KFM08A (Figure 3b, dark yellow dots) that is representative of FFM01.

### 5.1. Description of the selected models

The geo-DFNs rely on site investigations and data interpretations at the Forsmark site [Darcel *et al.*, 2009, Fox *et al.*, 2007, Olofsson *et al.*, 2007]. The fracture density and fracture orientation distributions are measured from borehole logging and core mapping. The number of fractures fully intersecting the boreholes per unit core length  $P_{10}$  is converted into total surface of fracture per unit volume  $P_{32}$  by means of stereological rules [Dershowitz and Herda, 1992, Darcel *et al.*, 2003a, Davy *et al.*, 2006, Piggott, 1997]. The values of  $P_{32}$  for the different orientation poles recorded in Forsmark are given in Table 2 with references therein.

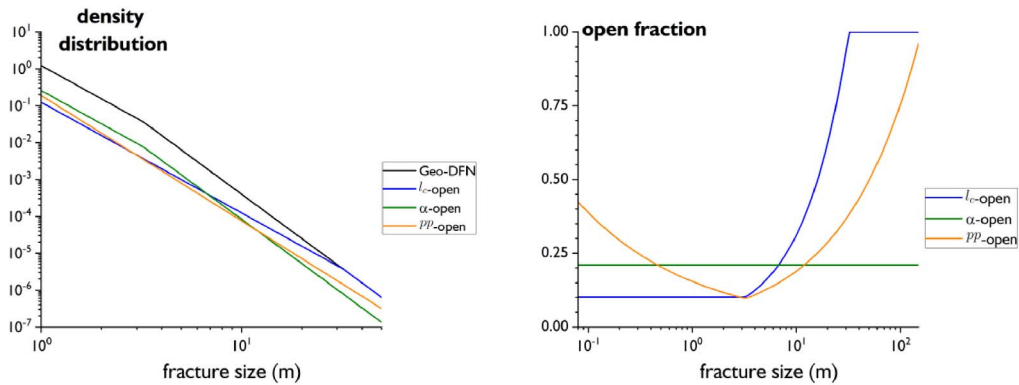
The multiscale organization of fracture networks is a critical attribute for fracture connectivity and flow [Bonnet *et al.*, 2001, Bour and Davy, 1998, Davy *et al.*, 2006]. It results in fractal properties and power law fracture size distributions, which are confirmed by the analysis of outcrop and lineament trace maps [Darcel *et al.*, 2009]. Davy *et al.* [2010] have shown that the outcrop data in Forsmark are consistent with a kinematic model of fracture nucleation, growth and arrest (hereafter called the UFM model). The resulting fracture-size distribution is a double power law with a scaling exponent close to  $-4$  above a critical fracture size  $l_c$  and  $-3$  below. The geo-DFN relies on the genetic UFM model with rules given in Davy *et al.* [2013]; it is characterized by the critical fracture size  $l_c$ , which fixes the fracture density  $P_{32}$ ; the larger is  $l_c$ , the smaller is  $P_{32}$ .

The open-DFN is a subset of the geo-DFN where only open fractures—i.e., fractures with a non-nil open aperture—are represented. The only data available is the surface ratio of open fractures (hereafter

**Table 2.** List of the different fracture sets for the FFM01 fracture domain at depths between  $-200$  and  $-400$  m with their intensity  $P_{32}$  and the Fisher parameters that define their orientation distributions

Set name	$P_{32}$ ( $\text{m}^2/\text{m}^3$ )	Fisher distribution parameters		
		Trend ( $^\circ$ )	Plunge ( $^\circ$ )	Kappa (-)
NS	1.292	292	1	17.8
NE	1.733	326	2	14.3
NW	0.948	60	6	12.9
EW	0.169	15	2	14.0
HZ	0.624	5	86	15.2

$P_{32}$  values are taken from Tables 6–33 in Glamheden et al. [2010] and Fisher parameters from Table C-1 in Follin [2008].



**Figure 8.** Density distribution (left) and open fraction (right) as a function of the fracture size for the Geo-DFN,  $l_c$ -open,  $\alpha$ -open and  $pp$ -open models, respectively.

**Table 3.** Parameters of the open-DFN models

Model	Open fraction	$pP_{32}$ borehole scale	Fracture size distribution
Geo-DFN	0	4.6	UFM size model, $l_c = 3.2$ m
$l_c$ -open 25%	0.25	1.2	UFM size model, $l_c = 17.6$ m
$l_c$ -open 21%	0.21	1	UFM size model, $l_c = 21.7$ m
$\alpha$ -open	0.25	1.2	UFM size model, $l_c = 3.2$ m
$pp$ -open	0.21	1.02	Power-law, exponent $-3.4$

The open fraction and  $P_{32}$ , which are scale-dependent parameters, are calculated for a fracture size of  $0.076$  m equivalent to the borehole diameter. Both parameters are thus comparable to field measurements. For all the models, the smallest fracture for flow calculation is  $2$  m, and the system is a cube of  $150$  m side. The number of fractures is between  $37,000$  and  $70,000$  depending on the open-DFN model.

referred as the open fraction  $f_{op}$ ), which is varying between  $0.15$  and  $0.25$  in the geological formation [Follin et al., 2014, Olofsson et al., 2007] with a value around  $0.15$  for the target fracture domain FFM01 [Doolaeghe Wehowsky, 2021].  $f_{op}$  can be measured from the core logging by identifying open

and partially-open fractures and measuring the ratio of fractures presenting open voids, but this is a bulk ratio that does not indicate how the fractures are opened individually or depending on their characteristics. We suspect that the relationship between the open fraction and the fracture attributes (size or ori-

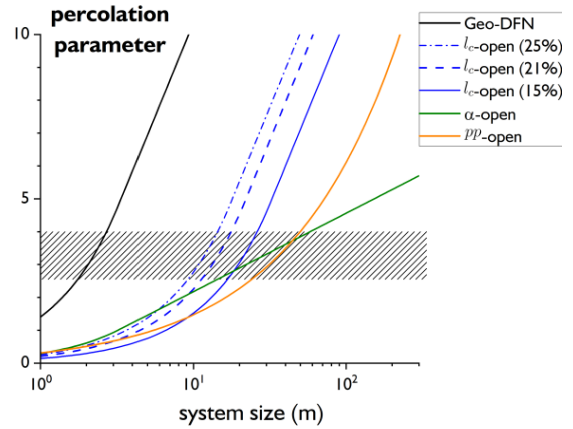
entation) is a key parameter to define the hydraulic properties of the open-DFN.

We analyze different models for the distribution of the open fraction in the network, whose parameters are given in Table 3. In all the models, we assume that fractures are either completely open or sealed. The first two models—further referred to as “ $l_c$ -open”—are built as genetic UFM models [Davy et al., 2013] with a larger critical fracture size  $l_c$  than the geo-DFN to account for the decrease of fracture density  $P_{32}$ . The first and second models differ by the proportion of open fraction. The second model—further referred to as  $\alpha$ -open, is a geo-DFN for which a number of fractures have been removed to account for the open fraction independently of their size (Figure 8b). The last model, originally presented in Follin et al. [2014], relies on a different approach where the geo-DFN step was skipped and a open/hydro DFN was directly generated with only a single power-law fracture size distribution and no genetic base but only Poisson-point processes (no spatial correlations between fractures), with the orientation of Table 2 [Follin et al., 2014]. For simplicity, this last model is named  $pp$ -open ( $pp$  for poisson-point). The size distribution and the open fraction of the Geo-DFN,  $l_c$ -open,  $\alpha$ -open and  $pp$ -open models are given in Figure 8 and the other model characteristics summarized in the legend of Table 3.

Although the 4 DFN models are consistent with the same data, they are very different in terms of connectivity, as highlighted by the analysis of the percolation parameter  $p$  as a function of system size  $L$  (Figure 9). This parameter measures the degree of connectivity of the networks and is expressed as [Charlaix et al., 1984, de Dreuzy et al., 2000]:

$$p(L) = \int_{l_{\min}}^L \frac{\pi^2}{8} l^3 n(l) dl \quad (10)$$

where  $n(l)$  is the density distribution of fracture size  $l$  per unit volume, and  $l_{\min}$  is the smallest value of the distribution  $n(l)$ . Bour and Davy [1998] demonstrate that the percolation parameter increases with scale and that the percolation threshold marks the limit scale above which statistically networks are connected by several fractures. The larger the percolation threshold, the larger the number of connected paths in a system. Figure 9 shows that the connection scale ranges from 10 to 50 m for the studied models,



**Figure 9.** Evolution of the percolation parameters as a function of the system size for the Geo-DFN,  $l_c$ -open,  $\alpha$ -open and  $pp$ -open models. The expression of  $p$  is given by (10). The dashed rectangle indicates the percolation threshold for the prescribed orientations.

and that  $l_c$ -open models are better connected than  $\alpha$ -open and  $pp$ -open model.

The Hydro-DFN is defined by assigning a transmissivity for each fracture. We define 4 different transmissivity models:

- all fractures have the same transmissivity ( $T = 1$ )—in this case, the resulting permeability is called the effective connectivity [de Dreuzy et al., 2001a]—;
- fracture transmissivity is proportional to the fracture size ( $T = l$ ), which is the standard model studied by Frampton and Cvetkovic [2010] for similar geological formations;
- fracture transmissivity is a function of the normal stress applied to the fracture ( $T = f(\sigma_n)$ ); it mimics the closure of fractures by the normal stress to the fracture plane  $\sigma_n$ ; in practice,  $\sigma_n$  is calculated from the projection of the remote stress tensor  $\bar{\sigma}$  given in Glamheden et al. [2007]; the transmissivity-stress function is derived from Follin and Stigsson [2014] as:

$$f(\sigma_n) = \exp\left(-\frac{\sigma_n}{\sigma_c}\right)$$

with  $\sigma_c = 4$  MPa;



- fracture transmissivity is assumed to be the product of the two previous relationships, i.e., fracture size and stress:  $T = lf(\sigma_n)$ .

This list does not cover the whole range of model possibilities but it gives good examples of the types of transmissivity relationships that could be investigated for a site modeling. Note that, in these relations, the transmissivities are relative with respect to a reference value that must be calibrated to the data. The following analysis is independent of this reference.

## 5.2. Comparison models/data, and sensitivity analysis

Flow simulations are performed with the computational suite “DFN.lab” [Le Goc et al., 2019], according to the setup described in Appendices A and B. Results are averaged from 10 realizations for  $K_g$  (6) and the value is normalized by the permeability calculated first for a scale of 0.5 m:  $K_g^*(L) = K_g(L)/K_g(L=0.5)$ .

Figure 10 shows the effects of the transmissivity model on the evolution of  $K_g^*$ . The first part of the curve is similar for all the transmissivity models with a  $L^{-1}$  decrease. After the percolation threshold,  $K_g^*$  stabilizes around a constant value for  $T = 1$  (constant fracture transmissivity), and it grows for all other transmissivity models. The permeability increase after threshold is primarily induced by a size dependency of fracture transmissivity (see the models  $T = l$  and  $T = l \cdot f(\sigma)$ ); it is also induced, even though to a lesser degree, by a dependence of the transmissivity on stress ( $T = f(\sigma)$ ) and thus on orientation. The combination of fracture size and stress/orientation dependency of fracture transmissivity ( $T = l \cdot f(\sigma)$ ) has an effect on permeability scaling that is more than each of the property effects taken separately. This illustrates that the flow variability inducing the permeability increase is not a linear juxtaposition of effects.

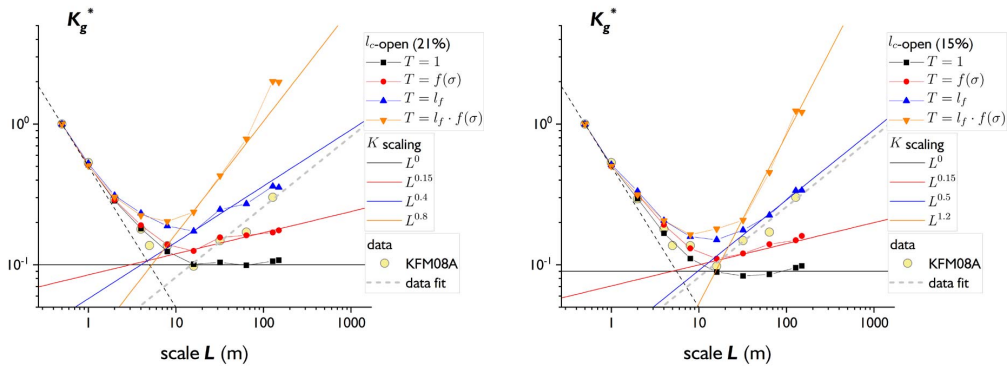
The transition scale between both regimes has been estimated from the intersection of the end-member regimes at small and large scales:  $K \propto L^{-1}$  at small scales, and  $K \propto L^\alpha$  at large scales with  $\alpha$  dependent on the transmissivity model. The transition is about independent of the transmissivity model around 4–4.5 m for the  $l_c$ -open model with an open fraction of 21% (Figure 10a) and around 10 m with 15% (Figure 10b). It is slightly larger than—but still

consistent with—the percolation scale for this model estimated from Figure 9.

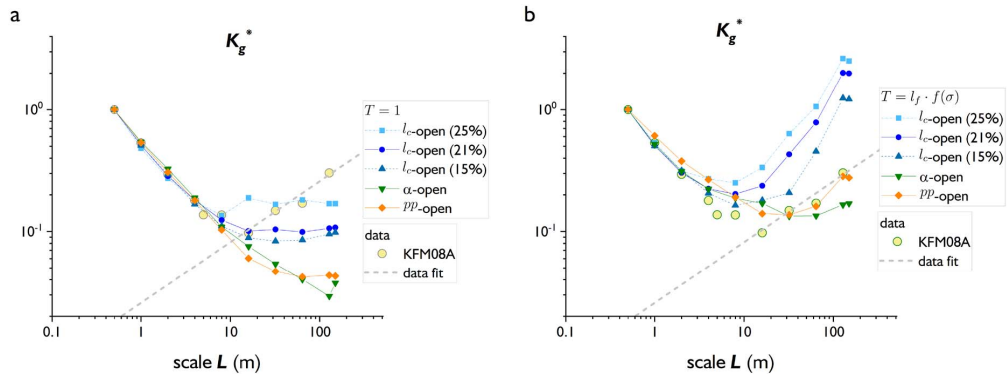
These simulations are merely indications of the possible effects associated with the transmissivity models. A more complete study is needed to derive general laws. Note also that the normalization by the first point of the curve removes the variability from one model to another and from one realization to another. This point is discussed later.

How the permeability scaling reveals the open-model is presented in Figure 11 for two transmissivity models:  $T = 1$  (left) and  $T = lf(\sigma)$  (right). The trends are similar to those presented in Figure 10, and the differences between the different open models mostly lie in the scale at which the transition between both two end-member regimes occurs. The increasing order of the different transition scales is in agreement with the evolution of the percolation parameter with the scale:  $l_c$ -open models first, then  $pp$ -open model and  $\alpha$ -open. Even the crossing of the curves of the last two models observed on  $K_g^*$  is predicted by the percolation parameter. A comparison with the borehole data of KFM08A is made in Figure 11 (right). None of the presented open models correctly reproduce the increase of permeability with scale after the transition scale for the two transmissivity models presented,  $T = 1$  and  $T = lf(\sigma)$ . The  $\alpha$ -open model is too weakly connected at large scale irrespective of the fracture transmissivity model, while  $pp$ -open model is slightly too connected at small scales. The  $l_c$ -open model with an open fraction of 15% and fracture transmissivity that scales with the fracture size is a good candidate to reproduce the permeability scaling at small and large scales (Figure 10b).

In Figure 12, we present an attempt to estimate the model variability (third validation indicator in Section 4.4). We calculate the lognormal standard deviation of the geometric average permeability  $\sigma_{\ln(K_g)}$  for 10 realizations for models, and in-between boreholes for data.  $\sigma_{\ln(K_g)}$  is dimensionless and does not depend on the unit of the fracture transmissivity, which makes it a useful indicator for comparing models and data.  $\sigma_{\ln(K_g)}$  has been calculated for 10 realizations for each combination between an open model and a transmissivity model. Figure 12 shows the evolution of  $\sigma_{\ln(K_g)}$  with scale for different transmissivity models and  $l_c$ -open 15% (left), and for different open models and  $T = lf(\sigma)$  (right). Not sur-



**Figure 10.** Evolution of the geometric averaged permeability as a function of scale for the  $l_c$ -open model 21% (left) and 15% (right) with the various transmissivity models (see the list above). An approximate power law fit for the different regimes of scaling is provided for all the models. The fit exponent is an indication of the intensity of the permeability with scale for large scales. The permeability scaling measured at KFM08A (yellow dots) and the fit for large scaling (grey dashed lines) are also indicated for comparison.

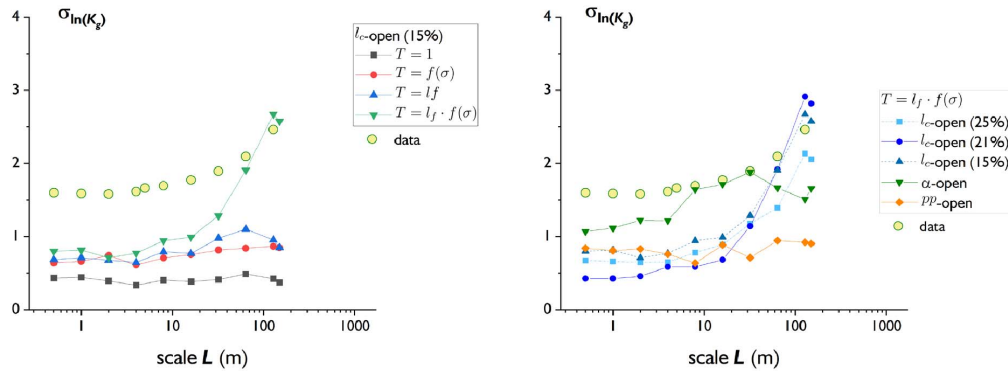


**Figure 11.** Evolution of the normalized geometric-averaged permeability for the different open-models and two transmissivity models:  $T = 1$  (left) and  $T = l_f \cdot f(\sigma)$  (right). The symbol colors and line types are similar to the one used in the Figure 9. The yellow dots on the right plot indicate the KFM08A scale analysis.

prisingly, the transmissivity model with the largest variability is  $T = l_f f(\sigma)$ , i.e., when the transmissivity depends both on stress and fracture size. This is the only model that shows an increase in variability with scale above 10–20 m. This trend is also observed in the data but to a lesser degree. For all other transmissivity models, the variability is about constant with scale and it is not very different from one model to another except for the model with a constant transmissivity which has a lower variability (Figure 12a). This result must be confirmed by a larger number of simulations.

The impact of the open model on  $\sigma_{\ln(K_g)}$  is presented in Figure 12b. The  $\alpha$ -open model has the largest variability, probably because it is largely controlled by a few big fractures. The  $pp$ -open model is the only one where the variability does not increase with scale for this transmissivity model.

Note that the comparison between model and data is not fair because the boreholes mix domains that can be quite heterogeneous, in particular, in terms of open fraction, while models are built with a constant open fraction. The  $\sigma_{\ln(K_g)}$  values are therefore likely to be overestimated for the data compared



**Figure 12.** Evolution of the standard deviation for 10 realizations of the lognormal distribution of permeability as a function of scale. Left: comparison of the different open-DFN models, all with the transmissivity model  $T = l$ ; right: comparison of the different transmissivity models, all with the  $l_c$ -open DFN model.

to the models and a correction may be necessary to make a true comparison.

## 6. Discussion

The four indicators defined in the Section 4.4 turn out to be useful to validate or calibrate models. All the models presented are consistent with the data, but no all of them can reproduce the data indicators. We discuss here the ability of the indicators to constrain both the structure that carries the flow and the transmissivity model that is applied to that structure.

All the models—again, calibrated on the data from Forsmark—reproduce the V-shape scaling curve of the permeability geometric average observed in data.

- The first indicator is the scale which marks the transition between the  $L^{-1}$  scaling at small scale and the transmissivity scaling at large scale. This scale is closely related to a connectivity threshold of the open-DFN structure. Indeed, for open-DFN, whose distribution of the largest fractures follows a power law, the connectivity increases with scale and the percolation threshold is reached at a given scale which can be predicted by the evolution of the percolation parameter with the scale. For all the tested models, the transition/percolation scale seems to be independent of the transmissivity model.

- The second indicator is the increase of the geometrically-averaged permeability with scale. It can be modeled by a power-law scaling whose exponent gives the increase in intensity. This indicator reveals the transmissivity model and a large increase is likely associated with a dependency of fracture transmissivity with fracture size. A high variability in fracture transmissivity, even without this transmissivity scaling effect, can also create an increase in permeability but to a lesser extent.
- The third indicator is the variability of geometrically-averaged permeability from one well to another (data) or from one realization to another (models). It has been quantified as the standard deviation of the logarithm of permeability, which is a dimensionless parameter that relies on the likely lognormal permeability distribution. This indicator is much more sensitive to the structure model (open-DFN) than the transmissivity model with the exception of the extreme transmissivity model  $T = l \cdot f(\sigma)$ , where the variability is high and scale dependent above the transition scale (i.e., 2nd indicator) as it is the case for the data. We do not want to draw definitive conclusions on this indicator because it would require many more models and realizations, and a comparison taken on more homogeneous

data.

Although the analysis is far from exploring all the structure and transmissivity models of DFN, we can draw few preliminary conclusions on the capacity of the chosen models to reproduce the indicators.

- In the DFN modeling workflow, an under-constrained critical point is the way in which the open fracture fraction is distributed in the network. Only the total open fraction cannot be retrieved from flow logs. If the open fraction is uniformly distributed over all fractures regardless of their size or orientation, as it is for the  $\alpha$ -open model, the network is less connected than other models, and probably the natural networks, whatever the assigned transmissivity model may be (see Figure 8a). On the other hand, if the open fraction is large at small scale, as it is for the  $pp$ -open model, the network is more connected than other models and data at these scales. This highlights the primary importance of properly defining what fractures are sealed and the proportion of open fractures per fracture size.
- The transmissivity model  $T = 1$  (i.e., constant transmissivity for each fracture) cannot not reproduce the V-shape of  $K_g$  vs scale. This model is obviously unrealistic, however we suspect that any transmissivity model, whose variability from one fracture to another is too low could not reproduce the permeability increase with scale as observed in data above  $\sim 10$  m. On the other hand, the transmissivity model combining fracture size and normal stress  $T = lf(\sigma)$  predicts permeability increase with scale too sharp compared to data. It is not certain to find an unequivocal solution between the transmissivity model and the observed permeability scaling by playing on the fracture size and orientation/stress dependence, but this indicator still reduces the field of possibilities.
- It is difficult to know at present whether a model should be rejected altogether (e.g., the  $\alpha$ -open or  $pp$ -open models) or whether it is still possible to modify some of their parameters to fit the data. This exercise is

however necessary to limit the field of possible models.

- The best-fit model among those tested is  $l_c$ -open with a transmissivity model  $T = l$  but it is not totally satisfactory, particularly in terms of the variability it induces (3rd indicator), which is too low and does not depend on the scale.

Note that the fourth indicator, which is the existence of a representative elementary volume, is not observed in the data where the permeability geometric average increases continuously with the scale.

## 7. Conclusion

The objective of the paper is to better understand and quantify the flow structure in fractured rocks from PFL flow logs, and also to propose relevant indicators for validating, calibrating or even rejecting hydrological models. The data consists of a series of inflows along each borehole with an equivalent transmissivity measured from pumping tests. We first studied what the inflow distribution tells us about the permeability structure from a series of analysis: distribution of transmissivities as a function of depth, proportion of flowing sections as a function of section scale, scaling of the arithmetically-averaged and geometrically-averaged permeability. We argue that the permeability scaling evolution provides more information on the flow structures than macroscopic values (i.e., large scale permeability), the latter being too integrative.

We thus define three indicators of the flow structure from the scale evolution of the permeability geometric average that highlight intrinsic properties of the flow structure:

- a percolation scale  $l_s$ ,
- the way permeability increases with scale above  $l_s$ ,
- the permeability variability, measured as the standard deviation of the permeability logarithm for different boreholes or model realizations, as a function of scale.

A 4th indicator on the representative elemental volume could in principle be defined but the data show that this volume/scale is beyond the 300 m investigated.

We tested a series of numerical models by running flow simulations on generated DFNs to compute the same indicators on synthetic data. Most of the models are built from the DFN methodology with three steps: the geo-DFN based on the observed fracture network, the open-DFN which is the part of the geo-DFN where fractures are open, and a transmissivity model applying on each fracture of the open-DFN. The analysis of the models showed that the percolation scale is controlled by the open-DFN structure and that the percolation scale can be predicted from a scale analysis of the percolation parameter (basically, the third moment of the fracture size distribution that provides a measure of the network connectivity). For the same geo-DFN, the distribution of the open fraction with fracture size is the critical parameter that controls the percolation parameter and the percolation threshold. The way permeability increases with scale above the percolation threshold is controlled by the transmissivity model. We have tested several of them with a dependence of the fracture transmissivity on either the orientation of the fractures via a stress-controlled transmissivity or their size or both. Although size dependence induces the greatest effect on permeability increase, the variability due to orientation (i.e., stress) dependence also has a non-negligible effect that is enhanced when the two dependencies are combined.

The comparison with data on the first two indicators shows that a model that matches the characteristics of the geo-DFN with the measured open fraction of 15% adequately fits the data, provided that the large fractures remain open and that the fracture transmissivity model is well selected. Most of the other models shows unacceptable differences with data, however other models or model combinations still have to be explored before rejecting them.

The third indicator on model variability is still problematic since the natural data show a higher variability than the models but the open fraction is also much more variable in the data than in the models.

By characterizing the flow structure in terms of permeability scaling and variability, and by defining three indicators that describe some fundamental characteristics of the flow/permeability whatever the scale may be, we bring a new way to calibrate and/or validate models or at least to reduce the field of model possibilities.

## Conflicts of interest

Authors have no conflict of interest to declare.

## Acknowledgements

The authors acknowledge the two anonymous referees who helped improve the manuscript. Philippe Davy is particularly grateful to Professor Ghislain de Marsily for the help and encouragement to develop the fundamental research on the flow properties of fractured media.

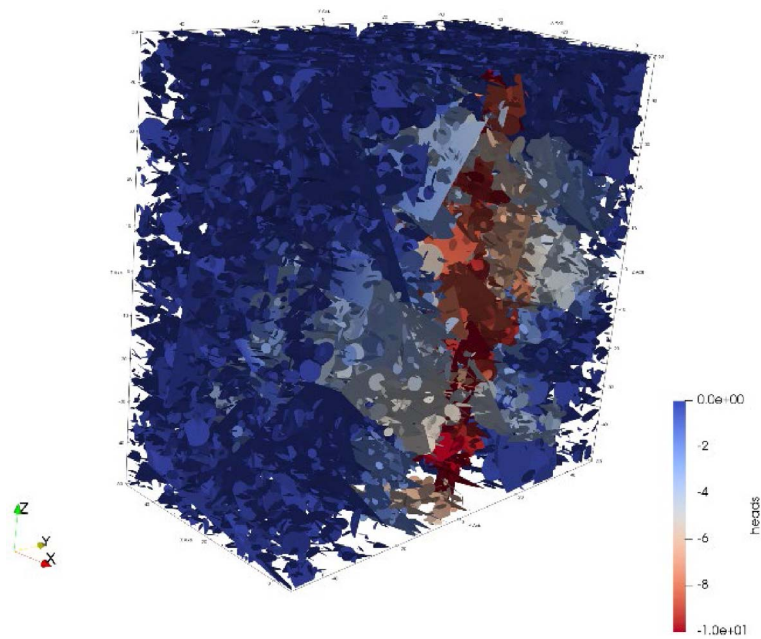
## Appendix A. DFN geometrical model definition

We consider the fractured rock mass as a set of fractures represented by discrete elements embedded in a linear elastic media, as commonly described by the Discrete Fracture Network (DFN) methodology. In 3D, fractures are represented by surfaces (usually disks), embedded in a volume. We hereafter refer to two types of DFN models, the *Poisson Point process-based* and the *genetic (Universal Fracture Model) UFM* ones.

The Poisson Point process-based DFN model—named *Poisson* DFN model—relies only on statistics-based fracture generation. The DFN properties are predefined by statistical distributions (density distributions of fracture sizes and orientations) and fractures are generated from a random sampling of geometrical properties (Poisson Point process-based), where each fracture position, size, and orientation are independent of the other fractures, but the average density is predetermined.

The *UFM* DFN model introduced by Davy et al. [2013] belongs to the range of genetic DFN models, introducing a fracture hierarchy between fractures by mimicking the fracturing process. The UFM model produces fracture size distributions and fracture intersections that are consistent with observations, using simple rules (N–G–A) for:

- Nucleation: fractures are introduced in the system at a constant nucleation rate  $n'_N$ , with a nuclei size of  $l_n$ .
- Growth: once created, fractures grow following a power-law relationship that describes the crack tip velocity  $v(l) = Cl^{a_G}$ , with  $l$  the



**Figure 13.** Clipped view along a vertical central plane going through the vertical borehole for PFL-like boundary conditions (a) flow rate (in  $\text{m}^3/\text{s}$ ) and (b) head (in m) fields.

fracture size,  $C$  the growth rate, and  $a_G$  the growth exponent.

- Arrest: while growing, fractures stop if they encounter a larger one, which is a first-order approximation of mechanical interactions.

While the DFN fracture size distribution is imposed from the beginning in a Poisson DFN model, it naturally arises from the growth and arrest processes in the UFM, as part of a spatially organized structure. For constant nucleation rate and growing fractures, there is a stationary solution for the fracture size distribution (referred to as the dilute regime) described by  $n_G(l) = \dot{n}_N/C \cdot l^{-a_G}$ , while the arrest rule is responsible for the emergence of a quasi-universal self-similar fracture size distribution (referred to as the dense regime) described by  $n_A(l) = D\gamma^D l^{-D+1}$ , with  $D$  the topological dimension associated to fracture centers, and  $\gamma$ , a geometrical parameter dependent on fracture orientations. The model thus results in a two power-law size distribution, where the transition size  $l_c$  between  $n_G$  and  $n_A$  is both the scale at which the network is connected and the average size of fracture blocks. Moreover, the arrest rule is responsible for the emergence of T intersections, where a fracture

abuts on another one. The fracture spatial organization introduced by the UFM framework has an impact on both the network topology [Hope *et al.*, 2015] and its hydrogeological behavior [Maillot *et al.*, 2016].

## Appendix B. Numerical set-up

The DFN and flow simulation volume is a generic 3D cubic domain. Boundary conditions are assigned all over the domain sides, either fixed heads or no flow and possibly along a generic borehole (like a 1D line) added in the cube. Steady-state flow tests are defined with PFL-like conditions to mimic the in-situ Posiva Flow Log (PFL) tests done at the Forsmark site and described in Section 2. A head drawdown of  $-10$  is imposed at a vertical borehole at the center of the domain to simulate the constant head pumping, while zero heads are imposed on the vertical sides and the top and no-flow on the bottom. The steady-state flow-field was calculated using DFN.lab, and the inflows to the borehole at each connected open fracture intersected by the borehole were retrieved. The borehole itself was represented by a vertical line.

The first step of the flow simulation is to identify the set of fractures of the DFN that constitute a

connected path between the defined active hydraulic boundaries. If there is no connected path, there is simply no flow simulation. Once the set of connected fractures is identified, it is cleaned up from its dead-ends (we assume that a fracture with only one intersection to the connected cluster or hydraulic boundary does not contribute to flow). This cleaned fracture cluster can be called a flow backbone. Finally, fractures of the flow backbone are meshed, and flow is solved using a mixed-hybrid finite-element scheme in steady-state conditions. (Figure 13).

For each intersection of the borehole with a flowing fracture  $f$ , one defines the observed fracture transmissivity  $T_{\text{PFL},f}$  as:

$$T_{\text{PFL},f} = \frac{Q_{\text{PFL},f}}{2\pi\Delta h} \cdot \ln\left(c \cdot \frac{R}{r_m}\right) \quad (11)$$

with  $Q_{\text{PFL},f}$  the flow rate at the borehole intersection,  $r_m/c$  (with  $c$  equal to 10) the equivalent borehole radius and  $R$  the radial distance between the borehole and the boundary conditions.  $c$  is a correction factor that accounts for numerical artifacts relative to the borehole type condition and meshes element shape (triangle). With this definition, the transmissivity of an isolated fracture connected to the borehole and the system limits can be inferred from the simulated values of  $Q_{\text{PFL},l}$  and  $\Delta h$ . If the intercepted fracture is not isolated, measured equivalent transmissivity rather reflects the flow capacity of the closely connected cluster of fractures around the intercepted one rather than the transmissivity of the intercepted fracture itself.

The transmissivity  $T_{\text{PFL}}(s)$ , of a borehole section of size  $s$ , including potentially several flowing fractures. It is defined from (11) with  $Q_{\text{PFL}}(s)$  the sum of the flow rates over the individual fractures of the section.

The equivalent hydraulic conductivity of the fracture network (in m/s) estimated from a borehole section is equal to the sum of inflow transmissivities divided by the section length.

## References

- Berkowitz, B., Bour, O., Davy, P., and Odling, N. (2000). Scaling of fracture connectivity in geological formations. *Geophys. Res. Lett.*, 27(14), 2061–2064.
- Bonnet, E., Bour, O., Odling, N. E., Davy, P., Main, I., Cowie, P., and Berkowitz, B. (2001). Scaling of fracture systems in geological media. *Rev. Geophys.*, 39(3), 347–383.
- Bour, O. and Davy, P. (1997). Connectivity of random fault networks following a power law fault length distribution. *Water Resour. Res.*, 33(7), 1567–1583.
- Bour, O. and Davy, P. (1998). On the connectivity of three-dimensional fault networks. *Water Resour. Res.*, 34(10), 2611–2622.
- Cacas, M. C., Ledoux, E., de Marsily, G., Barbreau, A., Calmels, P., Gaillard, B., and Margritta, R. (1990a). Modeling fracture flow with a stochastic discrete fracture network: Calibration and validation: 2. The transport model. *Water Resour. Res.*, 26(3), 491–500.
- Cacas, M. C., Ledoux, E., de Marsily, G., Tillie, B., Barbreau, A., Durand, E., Feuga, B., and Peaudecerf, P. (1990b). Modeling fracture flow with a stochastic discrete fracture network: calibration and validation: 1. The flow model. *Water Resour. Res.*, 26(3), 479–489.
- Charlaix, E., Guyon, E., and Rivier, N. (1984). A criterion for percolation threshold in a random array of plates. *Solid State Commun.*, 50(11), 999–1002.
- Charlaix, E., Guyon, E., and Roux, S. (1987). Permeability of a random array of fractures of widely varying apertures. *Transp. Porous Media*, 2, 31–43.
- Chen, T., Clauser, C., Marquart, G., Willbrand, K., and Hiller, T. (2018). Upscaling permeability for three-dimensional fractured porous rocks with the multiple boundary method. *Hydrogeol. J.*, 26(6), 1903–1916.
- Chen, T., Clauser, C., Marquart, G., Willbrand, K., and Mottaghy, D. (2015). A new upscaling method for fractured porous media. *Adv. Water Resour.*, 80, 60–68.
- Clauser, C. (1992). Permeability of crystalline rocks. *Eos, Trans. Am. Geophys. Union*, 73(21), 233–238.
- Darcel, C., Bour, O., and Davy, P. (2003a). Stereological analysis of fractal fracture networks. *J. Geophys. Res.*, 108(B9), 13–1–13–14.
- Darcel, C., Bour, O., Davy, P., and de Dreuz, J. R. (2003b). Connectivity properties of two-dimensional fracture networks with stochastic fractal correlation. *Water Resour. Res.*, 39(10), article no. 1272.
- Darcel, C., Davy, P., Le Goc, R., Bour, O., and

- de Dreuzy, J. R. (2009). Statistical methodology for discrete fracture models – including fracture size, orientation uncertainty together with intensity uncertainty and variability. Technical Report R-09-38. Retrieved from Stockholm: <http://www.skb.com/publication/1983611/R-09-38.pdf>.
- Davy, P., Bour, O., De Dreuzy, J.-R., and Darcel, C. (2006). Flow in multiscale fractal fracture networks. In *Fractal Analysis for Natural Hazards*, volume 261 of *Geol. Soc. London, Spec. Publ.*, pages 31–45. Geological Society of London.
- Davy, P., Darcel, C., Le Goc, R., Munier, R., Selroos, J.-O., and Mas Ivars, D. (2018). DFN, why, how and what for, concepts, theories and issues. In *Paper presented at the 2nd International Discrete Fracture Network Engineering Conference*. OnePetro, Richardson, TX.
- Davy, P., Le Goc, R., and Darcel, C. (2013). A model of fracture nucleation, growth and arrest, and consequences for fracture density and scaling. *J. Geophys. Res. Solid Earth*, 118(4), 1393–1407.
- Davy, P., Le Goc, R., Darcel, C., Bour, O., de Dreuzy, J.-R., and Munier, R. (2010). A likely universal model of fracture scaling and its consequence for crustal hydromechanics. *J. Geophys. Res.*, 115(B10), 1–13.
- de Dreuzy, J. R., Davy, P., and Bour, O. (2000). Percolation parameter and percolation-threshold estimates for three-dimensional random ellipses with widely scattered distributions of eccentricity and size. *Phys. Rev. E*, 62(5), 5948–5952.
- de Dreuzy, J. R., Davy, P., and Bour, O. (2001a). Hydraulic properties of two-dimensional random fracture networks following a power law length distribution 1. Effective connectivity. *Water Resour. Res.*, 37(8), 2065–2078.
- de Dreuzy, J. R., Davy, P., and Bour, O. (2001b). Hydraulic properties of two-dimensional random fracture networks following a power law length distribution 2. Permeability of networks based on log-normal distribution of apertures. *Water Resour. Res.*, 37(8), 2079–2096.
- de Dreuzy, J. R., Davy, P., and Bour, O. (2002). Hydraulic properties of two-dimensional random fracture networks following power law distributions of length and aperture. *Water Resour. Res.*, 38(12), 12–1–12–9.
- de Dreuzy, J. R., de Boiry, P., Pichot, G., and Davy, P. (2010). Use of power averaging for quantifying the influence of structure organization on permeability upscaling in on-lattice networks under mean parallel flow. *Water Resour. Res.*, 46(8), 1–11.
- De Dreuzy, J. R., Pichot, G., Poirriez, B., and Erhel, J. (2013). Synthetic benchmark for modeling flow in 3D fractured media. *Comput. Geosci.*, 50, 59–71.
- Dershowitz, W. S. and Herda, H. H. (1992). Interpretation of fracture spacing and intensity. In *Paper presented at the The 33th US Symposium on Rock Mechanics (USRMS)*. OnePetro, Richardson, TX.
- Dewandel, B., Lachassagne, P., Wyns, R., Marechal, J. C., and Krishnamurthy, N. S. (2006). A generalized 3-D geological and hydrogeological conceptual model of granite aquifers controlled by single or multiphase weathering. *J. Hydrol.*, 330(1–2), 260–284.
- Doolaege Wehowsky, D. (2021). *Colmatage des réseaux de fractures, modèles et conséquences hydrologiques*. PhD thesis, University of Rennes 1, Rennes.
- Follin, S. (2008). Bedrock hydrogeology Forsmark, Site descriptive modelling, SDM-Site Forsmark. Technical Report R-08-95, Svensk Kärnbränslehantering AB (SKB), Stockholm, Sweden.
- Follin, S., Hartley, L., Rhén, I., Jackson, P., Joyce, S., Roberts, D., and Swift, B. (2014). A methodology to constrain the parameters of a hydrogeological discrete fracture network model for sparsely fractured crystalline rock, exemplified by data from the proposed high-level nuclear waste repository site at Forsmark, Sweden. *Hydrogeol. J.*, 22(2), 313–331.
- Follin, S., Leven, J., Hartley, L., Jackson, P., Joyce, S., Roberts, D., and Swift, B. (2007). Hydrogeological characterisation and modelling of deformation zones and fracture domains, Forsmark modelling stage 2.2. Technical Report R-07-48, Svensk Kärnbränslehantering AB (SKB), Stockholm, Sweden.
- Follin, S. and Stigsson, M. (2014). A transmissivity model for deformation zones in fractured crystalline rock and its possible correlation to in situ stress at the proposed high-level nuclear waste repository site at Forsmark, Sweden. *Hydrogeol. J.*, 22(2), 299–311.
- Fox, A., La Pointe, P., Hermanson, J., and Öhman, J. (2007). Statistical geological discrete fracture network model. Forsmark modelling stage 2.2. Technical Report R-07-46, Svensk Kärnbränslehantering AB (SKB), Stockholm, Sweden.
- Frampton, A. and Cvetkovic, V. (2010). Inference



- of field-scale fracture transmissivities in crystalline rock using flow log measurements. *Water Resour. Res.*, 46(11), article no. W11502.
- Glamheden, R., Fälth, B., Jacobsson, L., Harrström, J., Berglund, J., and Bergkvist, I. (2010). Counterforce applied to prevent spalling. Technical Report TR-10-37, Svensk Kärnbränslehantering AB (SKB), Stockholm, Sweden.
- Glamheden, R., Fredriksson, A., Roeshoff, K., Karlsson, J., Hakami, H., and Christiansson, R. (2007). Rock mechanics Forsmark. Site descriptive modelling Forsmark stage 2.2. Technical report, Swedish Nuclear Fuel and Waste Management, Sweden.
- Hope, S. M., Davy, P., Maillot, J., Le Goc, R., and Hansen, A. (2015). Topological impact of constrained fracture growth. *Front. Phys.*, 3, article no. 75.
- Hunt, A. G. (2003a). Reply to comment by S. P. Neuman on “Some comments on the scale dependence of the hydraulic conductivity in the presence of nested heterogeneity”. *Adv. Water Resour.*, 26(11), 1215.
- Hunt, A. G. (2003b). Some comments on the scale dependence of the hydraulic conductivity in the presence of nested heterogeneity. *Adv. Water Resour.*, 26(1), 71–77.
- Illman, W. A. (2006). Strong field evidence of directional permeability scale effect in fractured rock. *J. Hydrol.*, 319(1–4), 227–236.
- Le Goc, R., Pinier, B., Darcel, C., Lavoine, E., Doolaege, D., De Simone, S., de Dreuzy, J.-R., and Davy, P. (2019). DFN.lab: software platform for Discrete Fracture Network models. In *Paper presented at the AGU Fall Meeting 2019*. American Geophysical Union, Washington, DC.
- Lei, Q., Latham, J.-P., and Tsang, C.-F. (2017). The use of discrete fracture networks for modelling coupled geomechanical and hydrological behaviour of fractured rocks. *Comput. Geotech.*, 85, 151–176.
- Long, J. and Billaux, D. M. (1987). From field data to fracture network modeling: an example incorporating spatial structure. *Water Resour. Res.*, 23(7), 1201–1216.
- Long, J. C. S. and Witherspoon, P. A. (1985). The relationship of the degree of interconnection to permeability in fracture networks. *J. Geophys. Res.*, 90(B4), 3087–3098.
- Maillot, J., Davy, P., Le Goc, R., Darcel, C., and de Dreuzy, J. R. (2016). Connectivity, permeability, and channeling in randomly distributed and kinematically defined discrete fracture network models. *Water Resour. Res.*, 52(11), 8526–8545.
- Mandelbrot, B. B. (1982). *The Fractal Geometry of Nature*. W.H. Freeman, New-York.
- Martin, D. and Follin, S. (2011). Review of possible correlations between in situ stress and PFL fracture transmissivity data at Forsmark. Technical Report R-08-69, Svensk Kärnbränslehantering AB (SKB), Stockholm, Sweden.
- Martinez-Landa, L. and Carrera, J. (2005). An analysis of hydraulic conductivity scale effects in granite (Full-scale Engineered Barrier Experiment (FEBEX), Grimsel, Switzerland). *Water Resour. Res.*, 41(3), article no. W03006.
- Meier, P. M., Carrera, J., and Sanchez-Vila, X. (1999). A numerical study on the relationship between transmissivity and specific capacity in heterogeneous aquifers. *Groundwater*, 37(4), 611–617.
- Neuman, S. (2005). Trends, prospects and challenges in quantifying flow and transport through fractured rocks. *Hydrogeol. J.*, 13(1), 124–147.
- Neuman, S. P. (1994). Generalized scaling of permeabilities. *Geophys. Res. Lett.*, 21(5), 349–352.
- Neuman, S. P. (2003). Comment on “Some comments on the scale dependence of the hydraulic conductivity in the presence of nested heterogeneity” by A. G. Hunt. *Adv. Water Resour.*, 26(11), 1213.
- Neuman, S. P. and Di Federico, V. (2003). Multifaceted nature of hydrogeologic scaling and its interpretation. *Rev. Geophys.*, 41(3), 4–1–4–31.
- Oda, M. (1985). Permeability tensor for discontinuous rock masses. *Géotechnique*, 35(4), 483–495.
- Öhberg, A. and Rouhiainen, P. (2000). Posiva groundwater flow measuring techniques. Technical Report 2000-12, Posiva Oy, Helsinki, Finland. Retrieved from Helsinki.
- Olofsson, I., Simeonov, A., Stephens, M., Follin, S., Nilsson, A., Röshoff, K., Lindberg, U., Lanaro, F., Fredriksson, A., and Persson, L. (2007). Site descriptive modelling Forsmark, stage 2.2. Technical Report R-07-15, Svensk Kärnbränslehantering AB (SKB), Stockholm, Sweden.
- Park, B. Y., Kim, K. S., Kwon, S., Kim, C., Bae, D. S., Hartley, L. J., and Lee, H. K. (2002). Determination of the hydraulic conductivity components using a three-dimensional fracture network model in volcanic rock. *Eng. Geol.*, 66(1), 127–141.

- Pavón-Domínguez, P. and Moreno-Pulido, S. (2022). Sandbox fixed-mass algorithm for multifractal unweighted complex networks. *Chaos Solitons Fractals*, 156, article no. 111836.
- Perrochet, P. (2005). A simple solution to tunnel or well discharge under constant drawdown. *Hydrogeol. J.*, 13(5), 886–888.
- Piggott, A. (1997). Fractal relations for the diameter and trace length of disc-shaped fractures. *J. Geophys. Res.*, 102(B8), 18121–18125.
- Ren, S., Zhang, Y., Jim Yeh, T. C., Wang, Y., and Carr, B. J. (2021). Multiscale hydraulic conductivity characterization in a fractured granitic aquifer: the evaluation of scale effect. *Water Resour. Res.*, 57(9), article no. e2020WR028482.
- Renard, P. and Marsily, G. d. (1997). Calculating equivalent permeability: a review. *Adv. Water Resour.*, 20(5–6), 253–278.
- Rouhiainen, P. and Pöllänen, J. (2003). Forsmark site investigation Difference flow logging of borehole KFM01A. Technical Report P-03-28, Posiva Oy, Helsinki, Finland.
- Rouhiainen, P., Pöllänen, J., and Tec-Oy, P. (2004). Forsmark site investigation: Difference flow logging in borehole KFM06A. Technical Report P-05-15, Posiva Oy, Helsinki, Finland.
- Selroos, J.-O., Ivars, D. M., Munier, R., Hartley, L., Libby, S., Davy, P., Darcel, C., and Trincherio, P. (2022). Methodology for discrete fracture network modelling of the Forsmark site. Part 1—concepts, data and interpretation methods. Technical Report R-20-11, Svensk Kärnbränslehantering AB (SKB), Stockholm, Sweden.
- SKB (2011). Long-term safety for the final repository for spent nuclear fuel at Forsmark. Main Report of the SR-Site Project. Technical Report TR-11-01, Svensk Kärnbränslehantering AB (SKB), Stockholm, Sweden.
- Tél, T., Fülöp, A., and Vicsek, T. (1989). Determination of fractal dimensions for geometrical multifractals. *Phys. A: Stat. Mech. Appl.*, 159(2), 155–166.
- Wen, X.-H. and Gómez-Hernández, J. J. (1996). Upscaling hydraulic conductivities in heterogeneous media: An overview. *J. Hydrol.*, 183(1), ix–xxxii.
- Zou, L. and Cvetkovic, V. (2020). Inference of transmissivity in crystalline rock using flow logs under steady-state pumping: impact of multiscale heterogeneity. *Water Resour. Res.*, 56(8), article no. e2020WR027254.
- Zou, L. and Cvetkovic, V. (2021). Evaluation of flow-log data from crystalline rocks with steady-state pumping and ambient flow. *Geophys. Res. Lett.*, 48(9), article no. e2021GL092741.

# *Comptes Rendus*

---

## *Géoscience*

### **Objet de la revue**

Les *Comptes Rendus Géoscience* sont une revue électronique évaluée par les pairs de niveau international, qui couvre l'ensemble des domaines des sciences de la Terre et du développement durable. Ils publient des articles originaux de recherche, des articles de revue, des mises en perspective historiques, des textes à visée pédagogique ou encore des actes de colloque, sans limite de longueur, en anglais ou en français. Les *Comptes Rendus Géoscience* sont diffusés selon une politique vertueuse de libre accès diamant, gratuit pour les auteurs (pas de frais de publication) comme pour les lecteurs (libre accès immédiat et pérenne).

**Directeur de la publication :** Étienne Ghys

**Rédacteurs en chef :** Éric Calais, Michel Campillo, François Chabaux, Ghislain de Marsily

**Comité éditorial :** Jean-Claude André, Pierre Auger, Mustapha Besbes, Sylvie Bourquin, Yves Bréchet, Marie-Lise Chanin, Philippe Davy, Henri Décamps, Sylvie Derenne, Michel Faure, François Forget, Claude Jaupart, Jean Jouzel, Eric Karsenti, Amaëlle Landais, Sandra Lavorel, Yvon Le Maho, Mickaele Le Ravalec, Hervé Le Treut, Benoit Noetinger, Carole Petit, Valérie Plagnes, Pierre Ribstein, Didier Roux, Bruno Scaillet, Marie-Hélène Tusseau-Vuillemin, Élisabeth Vergès

**Secrétaire éditoriale :** Adenise Lopes

### **À propos de la revue**

Toutes les informations concernant la revue, y compris le texte des articles publiés qui est en accès libre intégral, figurent sur le site <https://comptes-rendus.academie-sciences.fr/geoscience/>.

### **Informations à l'attention des auteurs**

Pour toute question relative à la soumission des articles, les auteurs peuvent consulter le site <https://comptes-rendus.academie-sciences.fr/geoscience/>.

### **Contact**

Académie des sciences  
23, quai de Conti, 75006 Paris, France  
Tél. : (+33) (0)1 44 41 43 72  
CR-Geoscience@academie-sciences.fr



Les articles de cette revue sont mis à disposition sous la licence  
Creative Commons Attribution 4.0 International (CC-BY 4.0)  
<https://creativecommons.org/licenses/by/4.0/deed.fr>

# COMPTES RENDUS DE L'ACADÉMIE DES SCIENCES

## *Géoscience* *Sciences de la Planète*

Volume 355, n° S1, 2023

### Special issue / Numéro thématique

Geo-hydrological Data & Models / *GDM - Géo-hydrologie, données et modèles*

### Guest editors / Rédacteurs en chef invités

Vazken Andréassian (INRAE, France), Valérie Plagnes (Sorbonne Université, France), Craig Simmons (Flinders University, Australia)  
Pierre Ribstein (Sorbonne Université, France)

### Cover illustration / Illustration de couverture

Portrait de Ghislain de Marsily, d'après une photographie de Roberto Bruno prise le 12/12/2008 à l'université de Bologne (tous droits réservés). Figure d'après la fig. 1 p. 86 du polycopié de cours de Ghislain de Marsily « Cours d'Hydrogéologie de l'Université Paris VI, version de septembre 2004, M2 HHGG ».

### Vazken Andréassian, Valérie Plagnes, Craig Simmons, Pierre Ribstein

A hydrogeological acoustic: in honour of Ghislain de Marsily ..... 1-8

## INVERSE METHODS

### Jeremy White, Marsh Lavenue

Advances in the pilot point inverse method: Où En Sommes-Nous maintenant? ..... 9-17

### Przemysław Juda, Julien Straubhaar, Philippe Renard

Comparison of three recent discrete stochastic inversion methods and influence of the prior choice ..... 19-44

### Philippe Ackerer, Jesus Carrera, Frédéric Delay

Identification of aquifer heterogeneity through inverse methods ..... 45-58

### Frederick Delay, Jean-Luc Mari, Gilles Porel, François Chabaux, Philippe Ackerer

Is subsurface geophysics as seismic and acoustic investigations a rescue to groundwater flow inversion? ..... 59-78

## SURFACE HYDROLOGY

### Nicolas Le Moine

On the computation of the Basal Envelope Surface of Talwegs using the Analytic Element Method ..... 79-97

### Ludovic Oudin, Morgane Lalonde

Pitfalls of space-time trading when parametrizing a land use dependent hydrological model ..... 99-115

### Thibault Mathevet, Nicolas Le Moine, Vazken Andréassian, Hoshin Gupta, Ludovic Oudin

Multi-objective assessment of hydrological model performances using Nash–Sutcliffe and Kling–Gupta efficiencies on a worldwide large sample of watersheds ..... 117-141

### Deniz Kilic, Agnès Rivière, Nicolas Gallois, Agnès Ducharne, Shuaitao Wang, Philippe Peylin, Nicolas Flipo

Assessing water and energy fluxes in a regional hydrosystem: case study of the Seine basin ..... 143-163

### Daniel Zimmer, Julien Tournéze, Sami Bouarfa, Cyril Kao, Benoît Lesaffre

Land drainage functioning and hydrological impacts in rural catchments: model development and field experiments ..... 165-183

## GROUNDWATER HYDROLOGY

François Chabaux, Philippe Négrel, Frederick Gal, Daniel Viville, Coralie Ranchoux, Jérôme Van der Woerd, Frederick Delay, Julien Ackerer, Luc Aquilina, Marc Diraison, Frédéric Masson, Anne-Sophie Mériaux, Thierry Labasque, Yann Lucas, Philippe Adrien Pezard, Thierry Reuschlé, Anne-Désirée Schmitt, Peter Stille, Robert Wyns, Catherine Lerouge, Chrystel Dezayes

Hard rock aquifer architecture and water circulation levels in the Strengbach critical zone observatory (France) ..... 185-206

Marc Dumont, Valérie Plagnes, Patrick Lachassagne, Roger Guérin, Bayu Nugraha, Febriwan Mohamad, Ludovic Oudin, Arif Fadillah, Danièle Valdès, Gilles Brocard, Jean-Luc Bonjour, Mohamed Saadi, Anne-Sophie Esneu, Aswar Muhammad, Hendarmawan, Nathalie Dörfliger

Water cycle modelling strengthened by probabilistic integration of field data for groundwater management of a quite unknown tropical volcanic hydrosystem ..... 207-229

### Arnaud Malard, Nathalie Dörfliger, Pierre-Yves Jeannin, Silvan Laube

Contribution of the dynamic visualization of a 3D hydrogeological conceptual model to the participatory management of groundwater resources ..... 231-244

<b>Anne Jost, Shuaitao Wang, Thomas Verbeke, François Colleoni, Nicolas Flipo</b> Hydrodynamic relationships between gravel pit lakes and aquifers: brief review and insights from numerical investigations .....	245-269
<b>Antoine Pelletier, Vazken Andréassian</b> An underground view of surface hydrology: what can piezometers tell us about river floods and droughts? .....	271-280
<b>Julio Gonçalves, Alexis Nutz, Pierre Séraphin, Amine Chekireb, Lahcen Kabiri, Pierre Deschamps</b> Dealing with hydrologic data scarcity: the case of the Tindouf basin .....	281-300

## WATER IN THE ANTHROPOCENE

<b>Gilles Billen, Josette Garnier</b> The water-agro-food system: upscaling from the Seine river basin to the global scale .....	301-315
<b>Michel Meybeck, Gabrielle Bouleau, Catherine Carré, Josette Garnier, Laurence Lestel</b> Rivers help us to quantify the socio-ecological functioning of their basin at the Anthropocene: the Seine example (1850–2020) .....	317-335

## MODELLING PHILOSOPHY & LAW

<b>Vazken Andréassian</b> On the (im)possible validation of hydrogeological models .....	337-345
<b>Maria-Theresia Schafmeister</b> High-level radioactive waste repository: How Geology combined with societal principles can lead to public acceptance—the German experiment .....	347-361
<b>Anne Coudrain, Matthieu Le Duff, Danielle Mitja</b> The Anthropocene is shifting the paradigm of geosciences and science .....	363-380

## CLIMATE CHANGE

<b>Philémon Autin, Jean Emmanuel Sicart, Antoine Rabatel, Regine Hock, Vincent Jomelli</b> Climate reconstruction of the Little Ice Age maximum extent of the tropical Zongo Glacier using a distributed energy balance model .....	381-398
<b>Chantal Gascuel-Odoux, Ophélie Fovet, Mikaël Fauchoux, Jordy Salmon-Monviola, Laurent Strohmeier</b> How to assess water quality change in temperate headwater catchments of western Europe under climate change: examples and perspectives .....	399-409
<b>Roland Yonaba, Lawani Adjadi Mounirou, Fowé Tazen, Mahamadou Koïta, Angelbert Chabi Biaou, Cheick Oumar Zouré, Pierre Queloz, Harouna Karambiri, Hamma Yacouba</b> Future climate or land use? Attribution of changes in surface runoff in a typical Sahelian landscape .....	411-438
<b>Justine Tirogo, Anne Jost, Angelbert Biaou, Youssouf Koussoubé, Pierre Ribstein, Denis Dakouré</b> Impacts of climate change and pumping on groundwater resources in the Kou River basin, Burkina Faso .....	439-464
<b>Mustapha Besbes, Jamel Chahed</b> Predictability of water resources with global climate models. Case of Northern Tunisia .....	465-486

## KARSTS & SUBMARINE SPRINGS

<b>Perrine Fleury, Séverin Pistre, Michel Bakalowicz</b> Coastal karst aquifers and submarine springs: what future for their water resources? .....	487-500
<b>Esteban Sanz, Jesús Carrera, Carlos Ayora, Alfredo Barón, Concha González</b> Assessing the salinization mechanisms of coastal brackish springs .....	501-522
<b>David Labat, Thomas Guilloteau, Sébastien Lamagnère, Thaïs Dumareau, Gabin Haryouli, Sandra Béranger, Franck Granouillac, Vianney Sivelles</b> New insights from an enhancement of hydrogeological archives of a French karst watershed over the last 50 years .....	523-531
<b>Guy Vasseur</b> Geothermal heat advected by the recharge of underground conduit. Case study of the karstic spring of Lez (Hérault, France) .....	533-557

## GEOSTATISTICS

<b>Benoît Noetinger</b> Random fields and up scaling, towards a more predictive probabilistic quantitative hydrogeology .....	559-572
<b>Hayet Chihi, Mohamed Amin Hammami, Imen Mezni, Habib Belayouni, Abdallah Ben Mammou</b> Multiscale modeling of reservoir systems using geostatistical methods .....	573-603
<b>Chantal de Fouquet, Jérémy Nos</b> Cokriging of transmissivity from head measurements revisited in the case of two quasi steady state flows .....	605-615
<b>Vanessa Teles, Youri Hamon, Rémy Deschamps, Sébastien Rohais, Fadi H. Nader, Elodie Heckenmeyer, Marta Gasparini, Mickael Barbier, Olivier Lerat, Philippe Joseph, Brigitte Doligez</b> Modelling the coupled heterogeneities of the lacustrine microbialite-bearing carbonate reservoir of the Yacoraite Formation (Salta, Argentina) .....	617-636

## GEOLOGY

**Emmanuel Ledoux, Emmanuel Hertz, Jean-Charles Robinet, Pierre Combes**

Reflections on the role of chemical osmosis mechanisms on the long-term behavior of a collapsed salt cavity ..... 637-665

**Philippe Davy, Romain Le Goc, Caroline Darcel, Jan-Olof Selroos**

Scaling of fractured rock flow. Proposition of indicators for selection of DFN based flow models ..... 667-690

**COMPTES RENDUS**  
**DE L'ACADÉMIE DES SCIENCES**

***Géosciences***

Volume 355, n° S1, 2023