



Petrology, Geochemistry

## Comparing sequential Gaussian and turning bands algorithms for cosimulating grades in multi-element deposits

Shahrokh Paravarzar<sup>a,b</sup>, Xavier Emery<sup>a,\*</sup>, Nasser Madani<sup>a,b,c</sup><sup>a</sup> Department of Mining Engineering, University of Chile, Santiago, Chile<sup>b</sup> Advanced Mining Technology Center, University of Chile, Santiago, Chile<sup>c</sup> CSIRO-Chile International Center of Excellence in Mining and Mineral Processing, Santiago, Chile

## ARTICLE INFO

## Article history:

Received 19 March 2015

Accepted after revision 25 May 2015

Available online 17 July 2015

Handled by Marguerite Godard

## Keywords:

Geological heterogeneity

Multivariate modeling

Turning bands

Sequential simulation

Collocated cokriging

## ABSTRACT

Stochastic simulation is increasingly used to map the spatial variability in the grades of elements of interest and to assess the uncertainty in the mineral resources and ore reserves. The practical implementation requires specifying a stochastic model, which describes the spatial distribution of the grades, and an algorithm to construct realizations of these grades, viewed as different possible outcomes or scenarios. In the case of the Gaussian random field model, a variety of algorithms have been proposed in the past decades, but their ability to reproduce the model statistics is often unequal. In this paper, we compare two such algorithms, namely the turning bands and the sequential algorithms. The comparison is held through a synthetic case study and a real case study in a porphyry copper deposit located in southeastern Iran, in which it is of interest to jointly simulate the copper, molybdenum, silver, lead and zinc grades. Statistical testing and graphical validations are realized to check whether or not the realizations reproduce the features of the true grades, in particular their direct and cross variograms. Sequential simulation based on collocated cokriging turns out to poorly reproduce the cross variograms, while turning bands proves to be accurate in all the analyzed cases.

© 2015 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

### 1. Introduction

Geostatistical simulation is widely used in the evaluation of mineral resources and ore reserves to map geological heterogeneity at different spatial scales and to assess the uncertainty in the unknown values of coregionalized variables, such as the grades of elements of interest, petrophysical properties of the subsoil, or geometallurgical properties (work index, acid consumption, metal recoveries) (Boisvert et al., 2013; Rossi and Deutsch, 2014). Its practical implementation requires specifying a stochastic model, which describes the spatial

distribution of the coregionalized variables (*what* should be simulated), and an algorithm, which aims at constructing realizations of the prescribed model (*how* it should be simulated) (Chilès and Delfiner, 2012; Lantuéjoul, 2002).

When the coregionalized variables can be modeled (up to some nonlinear transformation) by Gaussian random fields, a few exact algorithms, such as the matrix decomposition (Davis, 1987; Myers, 1989), discrete spectral (Chilès and Delfiner, 1997; Dietrich and Newsam, 1993; Le Ravalec et al., 2000 Pardo-Igúzquiza and Chica-Olmo, 1993) and moving average (Black and Freyberg, 1990) algorithms, perfectly reproduce their joint distribution and spatial correlation structure, but such algorithms are limited, either because they cannot be used for large-scale problems or because they require the data and target locations to be regularly spaced.

\* Corresponding author.

E-mail address: xemery@ing.uchile.cl (X. Emery).

To overcome these limitations, approximate algorithms can be applied, allowing dealing with large numbers of data on unstructured grids. In this category, one finds the sequential Gaussian (Deutsch and Journel, 1998), continuous spectral (Shinozuka and Jan, 1972) and turning bands (Matheron, 1973) algorithms. Sequential Gaussian simulation has been widely used in practice due to its simplicity and straightforwardness in a variety of areas (Alabert and Massonnat, 1990; Ravenscroft, 1994), but the accuracy of this method is not always ensured (Emery, 2004; Emery and Peláez, 2011; Gómez-Hernández and Cassiraga, 1994; Omre et al., 1993; Tran, 1994) and its applicability to multivariate cases may be challenging and require simplifications (Almeida and Journel, 1994; Gómez-Hernández and Journel, 1993). An alternative to obtain good-quality realizations is the turning bands approach proposed by Matheron (1973). In a nutshell, this method performs simulation in a multi-dimensional space through a series of one-dimensional simulations. The algorithm allows fast calculations and, in theory, yields an accurate reproduction of the spatial correlation structure (in univariate and multivariate cases), although the resulting distributions may slightly differ from the target ones due to the use of a finite number of one-dimensional simulations (stripping effect) (Emery and Lantuéjoul, 2006).

The purpose of this paper is to assess the performance and check the accuracy of sequential Gaussian and turning bands simulation, through actual and synthetic case studies.

## 2. Theory of joint simulation

It is often of interest to construct numerical models that reproduce the joint distribution of several coregionalized variables at unsampled locations, conditionally to the information available at sampling locations (conditional cosimulation). Because the variables are usually spatially cross-correlated, it is not sufficient to simulate each variable separately. Instead, a multivariate approach has to be used.

In the case of representing the coregionalized variables of interest by Gaussian random fields, the problem of cosimulation consists in constructing realizations of a vector Gaussian random field, say  $\mathbf{Y} = \{\mathbf{Y}(\mathbf{x}) : \mathbf{x} \in \mathbf{D}\}$ , where  $\mathbf{D}$  is the domain of interest and  $\mathbf{x}$  represents a generic location in  $\mathbf{D}$ . For the sake of simplicity, further assume that the random field has zero mean and that its spatial correlation structure can be represented by a linear coregionalization model (Journel and Huijbregts, 1978; Wackernagel, 2003):

$$\mathbf{C}(\mathbf{h}) = \sum_{s=1}^S \mathbf{B}_s \rho_s(\mathbf{h}) \quad (1)$$

where  $\{\rho_s : s = 1, \dots, S\}$  is a set of auto-correlation functions (basic nested structures)

$\{\mathbf{B}_s : s = 1, \dots, S\}$  is a set of real-valued, positive semi-definite matrices (coregionalization matrices)

$\mathbf{C}(\mathbf{h})$  is a matrix containing the direct (diagonal terms) and cross (off-diagonal) covariance functions of the

components of  $\mathbf{Y}$  for a given separation vector  $\mathbf{h}$ :  $\mathbf{C}(\mathbf{h}) = E\{\mathbf{Y}(\mathbf{x}) \cdot \mathbf{Y}(\mathbf{x} + \mathbf{h})^T\}$ .

### 2.1. Sequential Gaussian cosimulation

Consider that  $\mathbf{D}$  is composed of  $n$  locations:  $\mathbf{D} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ . The sequential algorithm aims at simulating the vector random field  $\mathbf{Y}$  at each location successively. Specifically, at location  $\mathbf{x}_i$  (with  $i = 1, \dots, n$ ), the simulated vector is obtained as follows:

$$\mathbf{Y}(\mathbf{x}_i) = \mathbf{Y}^{\text{SCK}}(\mathbf{x}_i) + \sqrt{\Sigma^{\text{SCK}}(\mathbf{x}_i)} \mathbf{U}_i, \quad (2)$$

where

- $\mathbf{Y}^{\text{SCK}}(\mathbf{x}_i)$  is the simple cokriging prediction of  $\mathbf{Y}(\mathbf{x}_i)$ , obtained by using the pre-existing data as well as  $\mathbf{Y}(\mathbf{x}_1), \dots, \mathbf{Y}(\mathbf{x}_{i-1})$  as conditioning data
- $\Sigma^{\text{SCK}}(\mathbf{x}_i)$  is the variance-covariance matrix of the associated cokriging errors
- $\sqrt{\Sigma^{\text{SCK}}(\mathbf{x}_i)}$  is the principal square root of  $\Sigma^{\text{SCK}}(\mathbf{x}_i)$  (alternatively, the Cholesky factor of  $\Sigma^{\text{SCK}}(\mathbf{x}_i)$  could be used instead of the principal square root)
- $\mathbf{U}_i$  is a standard Gaussian vector with independent components, independent of  $\mathbf{U}_1, \dots, \mathbf{U}_{i-1}$ .

The sequential approach is applicable to simulate any vector Gaussian random field, even when its correlation structure is not a linear coregionalization model Eq. (1) (Marcotte, 2012), and, at least in theory, is perfectly accurate. However, in practice, some simplifications are required because the cokriging is computationally prohibitive when the number of data is too large. This happens when either the number of pre-existing data or the number of locations targeted for simulation ( $n$ ) is large. In this context, the following approximations are often used.

#### 2.1.1. Full cokriging in a moving neighborhood

Instead of cokriging with all the previously simulated vectors  $\mathbf{Y}(\mathbf{x}_1), \dots, \mathbf{Y}(\mathbf{x}_{i-1})$  and all the pre-existing conditioning data, one can restrict to the vectors and data that are located in a neighborhood of the target point  $\mathbf{x}_i$ . The design of such a neighborhood often considers a maximal search radius around the target point, as well as the definition of a maximum number of data and previously simulated vectors to search for (Deutsch and Journel, 1998; Goovaerts, 1997; Pebesma, 2004). The use of a local neighborhood is often combined with a multiple-grid strategy, which consists in visiting the target grid nodes according to a set of nested grids (starting from a coarse grid and following with finer ones), in order to better reproduce the spatial correlation at different scales (Tran, 1994).

#### 2.1.2. Collocated cokriging in a moving neighborhood

Here, simulation is performed in a hierarchical manner: the first component of  $\mathbf{Y}$  is simulated first, using univariate kriging in a moving neighborhood to determine the successive simulated values. The second component of  $\mathbf{Y}$  is then simulated using cokriging, conditionally to the

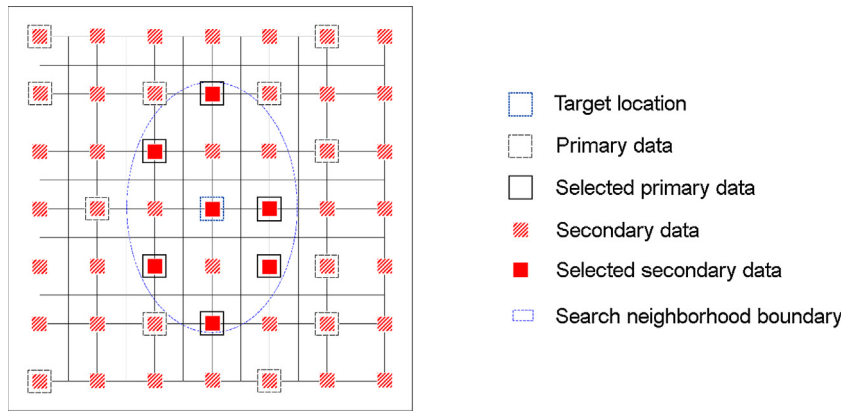


Fig. 1. (Color online.) Data selection for collocated cokriging (case of two variables, in which it is of interest to simulate the primary variable at the target location).

information of this component (i.e. original data and previously simulated values) located in the neighborhood of the target point, as well as the values of the first component at the target point  $\mathbf{x}_i$  and at the selected points with information of the second component. This variant of cokriging is known in the literature as multi-collocated (Rivoirard, 2001, 2004; Wackernagel, 2003), although in the following it will be referred just as collocated cokriging (Fig. 1). The process follows with the next components, by considering each time only the collocated information of the previously simulated components.

## 2.2. Turning bands cosimulation

### 2.2.1. Factorizing the target vector random field

To jointly simulate the components of  $\mathbf{Y}$ , one can split this vector random field as follows:

$$\forall \mathbf{x} \in \mathbf{D}, \mathbf{Y}(\mathbf{x}) = \sum_{s=1}^S \mathbf{Y}_s(\mathbf{x}), \quad (3)$$

In accordance with the linear coregionalization model,  $\mathbf{Y}_1, \dots, \mathbf{Y}_s$  are independent vector Gaussian random fields, with  $\mathbf{B}_1 \rho_1, \dots, \mathbf{B}_s \rho_s$  as their respective matrices of direct and cross covariance functions. Since each coregionalization matrix is positive semi-definite, it can be decomposed as follows:

$$\mathbf{B}_s = \mathbf{Q}_s \Delta_s \mathbf{Q}_s^T = \mathbf{A}_s \mathbf{A}_s^T, \quad (4)$$

with  $\mathbf{Q}_s$  an orthogonal matrix of eigenvectors,  $\Delta_s$  a diagonal matrix of eigenvalues and  $\mathbf{A}_s = \mathbf{Q}_s \sqrt{\Delta_s}$ .

Let  $\mathbf{W}_s$  be a vector random field with independent components, each with covariance function  $\rho_s$ . Then, it can be shown (Emery, 2008b) that the random field  $\mathbf{A}_s \mathbf{W}_s$  has the same correlation structure as  $\mathbf{Y}_s$ . The simulation of the target vector random field  $\mathbf{Y}$  therefore boils down to simulating independent scalar random fields (the components of  $\mathbf{W}_s$  for  $s = 1, \dots, S$ ) with covariance functions equal to the basic nested structures used in the linear coregionalization model.

### 2.2.2. Turning bands univariate simulation

The simulation of the components of  $\mathbf{W}_s$  can be performed via any multivariate Gaussian simulation algorithm. In the present case, the turning bands algorithm is chosen for its accuracy and computational efficiency. In a nutshell, this algorithm consists in drawing many lines in space, with random or, preferably, quasi-regular orientations, and simulating a one-dimensional random field along each line (Lantuéjoul, 1994, 2002). By adequately choosing the covariance function of such one-dimensional random fields, their superposition provides a multi-dimensional random field with the target covariance function, the distribution of which is practically Gaussian by virtue of the central limit theorem. The application of the method is therefore controlled by the number of lines used and by the way to simulate the basic one-dimensional random fields. The reader is referred to Emery and Lantuéjoul (2006) for mathematical details on these aspects.

### 2.2.3. Conditioning to data

In order to make the simulated vector random field conditional to a set of pre-existing data, one can post-process the realizations obtained by turning bands (non-conditional simulation) in order to turn them conditional. This step is based on cokriging the difference between the simulation at the data locations and the actual conditioning data values (Chilès and Delfiner, 2012; Emery, 2008b; Journel and Huijbregts, 1978). Note that cokriging can be performed even when the random fields to be simulated are not observed at the same locations (Wackernagel, 2003), so that the conditioning process does not lose information whatever the design of the conditioning dataset, either isotopic or heterotopic.

## 3. Synthetic analysis

One approach to assess the quality of a simulation algorithm is to compare the experimental statistics of a set of realizations (e.g., the mean value or the variogram calculated for given lag separation distances) with the underlying model statistics (Leuangthong et al., 2004).

However because the realizations are constructed over a bounded domain  $\mathbf{D}$  (represented by a finite set of spatial locations), some deviations or “fluctuations” between experimental and model statistics are expected, even if the simulation algorithm is exact. In this context, statistical testing can be used to reach a solid conclusion on whether or not the realization statistics agree with the model statistics (i.e., whether or not the fluctuations are within acceptable limits). In particular, one can check the reproduction of the mean value and of the variogram at specific lag distances (Emery, 2008a).

The synthetic case study shown in this section consists in simulating a single Gaussian random field with zero mean and isotropic spherical variogram on a two-dimensional regular grid  $\mathbf{D}$  with  $100 \times 100$  nodes. Two cases are considered: in the first one, the variogram has a range of 15 units and no nugget effect, while in the second case the variogram has a range of 60 units and 30% relative nugget effect. Both the turning bands and the sequential Gaussian algorithms are used to produce  $N = 100$  independent non-conditional realizations, using publicly available codes (Deutsch and Journel, 1998; Emery and Lantuéjoul, 2006). For sequential simulation, a moving neighborhood containing up to 64 previously simulated points (8 per octant) and a multiple-grid strategy with three nested grids are considered, whereas for turning bands, one thousand lines are used for simulation. The other parameters (variogram model, number of realizations and output grid) are the same for both the sequential and turning bands algorithms.

### 3.1. Testing the reproduction of the mean value

Consider the average of the Gaussian random field over the whole grid:

$$\mathbf{Y}(\mathbf{D}) = \frac{1}{10,000} \sum_{i=1}^{10,000} \mathbf{Y}(\mathbf{x}_i) \quad (5)$$

and denote by  $\bar{\mathbf{Y}}(\mathbf{D})$  and by  $S^2(\mathbf{D})$  the sample mean and sample variance of  $\mathbf{Y}(\mathbf{D})$ , calculated over the available  $N$  realizations. Then, one expects the following (Emery, 2008a):

$$\frac{\bar{\mathbf{Y}}(\mathbf{D})}{S(\mathbf{D})} \sqrt{N-1} \sim T_{N-1}, \quad (6)$$

where  $T_{N-1}$  is a Student random variable with  $N-1$  degrees of freedom. Accordingly, if the absolute value of the left-hand side member is greater than a critical value (for a pre-specified significance level), one would reject the null

**Table 1**

Student test on the mean for synthetic non-conditional simulations.

Simulation algorithm	Variogram range	Absolute value of T statistics
Turning bands	15	0.9617
	60	0.92259
Sequential simulation	15	0.17286
	60	0.64154

Theoretical mean value = 0; significance level = 0.05, critical value = 1.9842.

hypothesis that the mean value of the simulated field is zero. In the contrary case, the hypothesis of a zero mean is accepted and the simulation algorithm is validated, at least in what refers to the mean value. This test has been performed on the realizations obtained by turning bands and sequential Gaussian simulation. The results (Table 1) demonstrate that the hypothesis of a zero mean value is acceptable in all the cases with a significance level of 0.05.

### 3.2. Testing the reproduction of the variogram

Fig. 2 compares the experimental variograms of the realizations with the underlying variogram model. It is seen that, on average over the realizations, the variogram is well reproduced with the turning bands method, whereas a bias is perceptible for the sequential algorithm, with an exaggerated range for the variograms of the realizations and a different behavior at short distances in the case of the model with nugget effect. This visual appraisal can be confirmed by statistical testing.

For a given lag separation vector  $\mathbf{h}$ , let us denote by  $\Gamma(\mathbf{h})$  the experimental variogram of the simulated random field and by  $\gamma(\mathbf{h})$  the theoretical variogram model. Let  $\bar{\Gamma}(\mathbf{h})$  and  $S^2(\mathbf{h})$  be the sample mean and sample variance of  $\Gamma(\mathbf{h})$ , calculated over  $N$  independent realizations. For sufficiently high  $N$ , both  $\bar{\Gamma}(\mathbf{h})$  and  $S^2(\mathbf{h})$  are approximately normally distributed (due to the central limit theorem) and independent, so that one should have (Emery, 2008a):

$$\frac{\bar{\Gamma}(\mathbf{h}) - \gamma(\mathbf{h})}{S(\mathbf{h})} \sqrt{N-1} \sim T_{N-1} \quad (7)$$

As for the case of the mean value, if the absolute value of the left-hand side member is greater than a critical value (for a pre-specified significance level), one will reject the null hypothesis that the variogram of the simulated field at lag  $\mathbf{h}$  is equal to  $\gamma(\mathbf{h})$ . Otherwise, the null hypothesis is accepted and the simulation algorithm is validated in what refers to variogram reproduction. The test has been performed on the realizations obtained by turning bands

**Table 2**

Student test on variogram for synthetic non-conditional simulations.

Simulation algorithm	Variogram range	Absolute value of T-statistics				
		Lag 10	Lag 20	Lag 30	Lag 40	Lag 50
Turning bands	15	0.47958	0.13415	0.10243	1.1527	0.8572
	60	0.13575	0.37503	0.2262	0.11128	0.3551
Sequential simulation	15	<b>7.5313</b>	1.0281	0.68336	0.69996	1.4693
	60	<b>14.5126</b>	0.79557	<b>2.6556</b>	<b>3.4024</b>	<b>3.2084</b>

Significance level = 0.05, critical value = 1.9842. Bold numbers indicate rejection of the test.

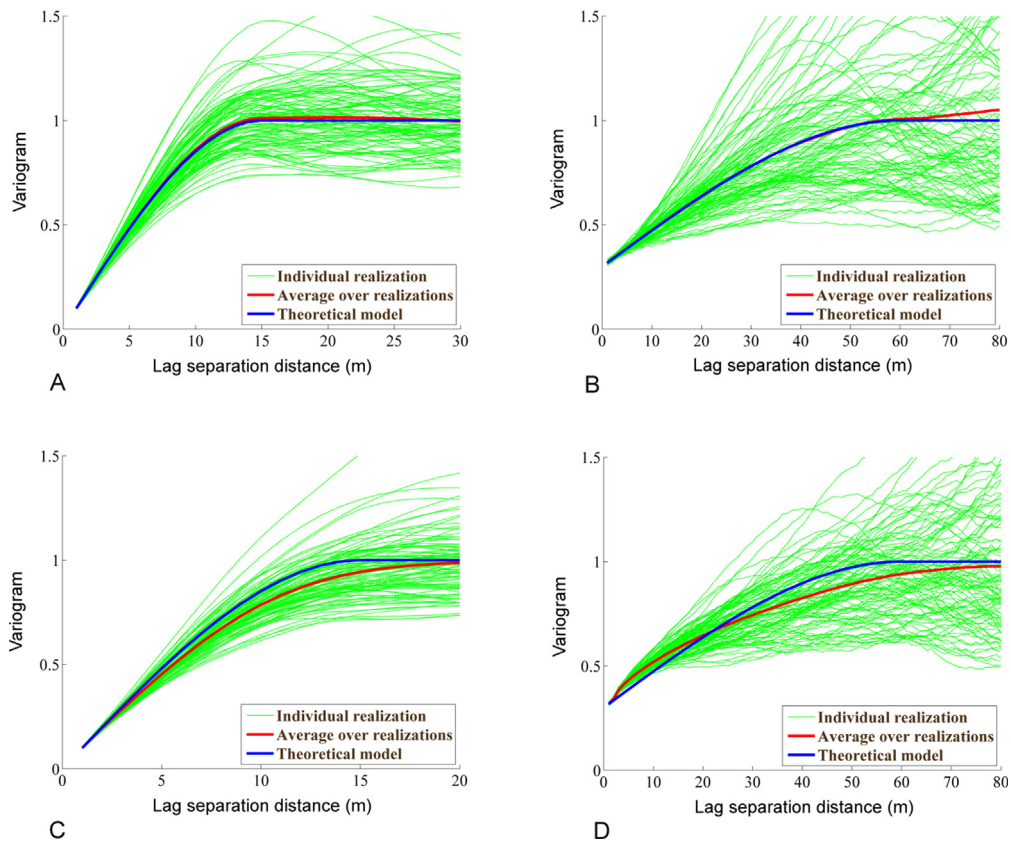


Fig. 2. (Color online.) Variogram reproduction for spherical variogram of range 15 and no nugget effect (left) and spherical variogram of range 60 and 30% relative nugget effect (right). A, B. Turning bands. C, D. Sequential simulation.

and sequential simulation for five lags (10, 20, 30, 40 and 50) and a significance level of 0.05. The results (Table 2) do not present any case of rejection for turning bands simulation, indicating that the variogram is indeed accurately reproduced, whereas sequential simulation has a rejection rate of 50%.

The previous results concern the reproduction of the variogram at five different lags considered separately. One can also test for the reproduction at the same five lags simultaneously, by using a Hotelling test that can be seen as a multivariate extension of the Student test (Emery, 2008a). The results (Table 3) confirm the accuracy of variogram reproduction for turning bands simulation and the bias for sequential simulation.

Although limited to a single variable, these experiments demonstrate that the choice of the simulation algorithm

has an impact on the quality of the results. In particular, beyond the visual inspection of the variograms of realizations, the Student and Hotelling tests prove that there is statistical evidence that sequential Gaussian simulation fails in reproducing the target spatial correlation structure (variogram), whereas turning bands proves to be accurate. The problem of sequential simulation stems from the iterative nature of the algorithm: the error made at each stage by using a moving neighborhood approximation propagates at the next stages, insofar as the value simulated at one location is used as a conditioning data for all subsequent locations (Emery and Peláez, 2011). Let us now examine what happens in a multivariate case, through a case study on real data.

#### 4. Case study: Dar-Alu porphyry copper deposit

##### 4.1. Data presentation

The Dar-Alu copper deposit is located at 29°24'46.6"N and 57°05'56.9"E, at about 135 km away from Kerman city, southeastern Iran. It is situated in the Urumieh–Dokhtar magmatic arc belt (Shafiei and Shahabpour, 2008) and consists of rhyodacite porphyry, silicified granodiorite porphyry, microgranodiorite, tonalite and rhyolite rocks. The classical alteration for copper ore deposits, such as

Table 3  
Hotelling test on variogram for synthetic non-conditional simulations.

Simulation algorithm	Variogram range	Hotelling statistics
Turning bands	15	3.0394
	60	4.3194
Sequential simulation	15	<b>120.2051</b>
	60	<b>815.0394</b>

Significance level = 0.05, critical value = 12.0375. Bold numbers indicate rejection of the test.

**Table 4**  
Basic statistics of grade data (Dar-Alu deposit).

Variable	Mean	Standard deviation	Minimum	Lower quartile	Median	Upper quartile	Maximum
Cu (%)	0.24	0.22	0.00	0.06	0.19	0.35	3.32
Mo (ppm)	48.18	83.00	0.00	11.61	29.31	58.18	3686.52
Ag (ppm)	0.56	0.63	0.00	0.18	0.41	0.75	22.73
Pb (ppm)	16.70	106.00	0.00	6.00	9.87	17.00	8564.29
Zn (ppm)	76.35	115.03	0.00	41.26	63.56	92.09	8016.48

sericitization and chloritization, prevails in the area and the mineralization mainly consists of pyrite, chalcopyrite, sphalerite, bornite, covellite, chalcocite, magnetite, hematite and limonite, distributed in oxide, leach, supergene and hypogene zones (Salehian and Ghaderi, 2010).

A data set has been obtained from an exploration drilling campaign, where sample cores from 69 drill holes were analyzed for Cu (%), Mo (ppm), Ag (ppm), Pb (ppm), and Zn (ppm). For statistical analyses, the data have been composited at a regular length of two meters. The main statistics and experimental distributions of the analyzed grades are presented in Table 4.

#### 4.2. Geostatistical modeling

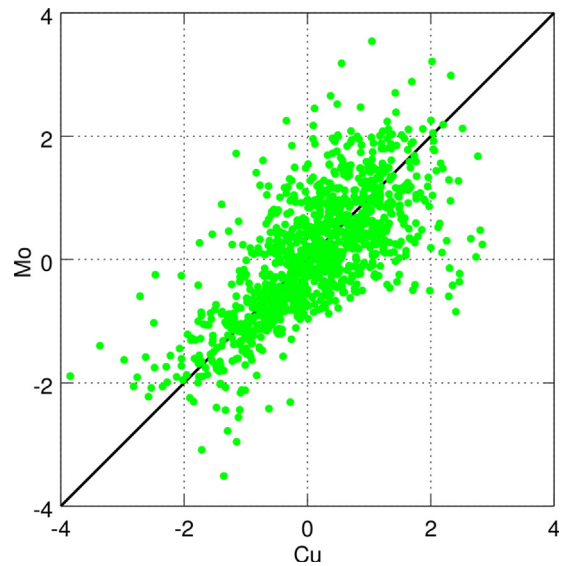
The modeling stage consists of the following steps:

- in order to have representative distributions of the original grade data, cell declustering is performed to correct for the irregular sampling design (Pyrzcz et al., 2006);
- each grade variable is transformed into a standard Gaussian variable (Deutsch and Journel, 1998). The different Gaussian variables appear to be cross-correlated (Table 5), especially between copper and molybdenum and between lead and zinc. These correlations can be explained because of similar geochemical and paragenetic properties in arc-related copper molybdenum deposits (Taylor et al., 2012);
- the marginal bivariate normality of the pairs of transformed variables is validated by examining their scatter diagrams, which exhibit an elliptic shape (Fig. 3). In turn, the spatial bivariate normality of the transformed variables is validated by checking that their experimental madograms are approximately proportional to the square roots of their experimental variograms (Fig. 4) (Emery, 2005);
- variogram analysis is performed on the transformed data in order to model their joint correlation structure. Since no obvious anisotropy was detected, omnidirectional sample variograms were calculated and fitted with

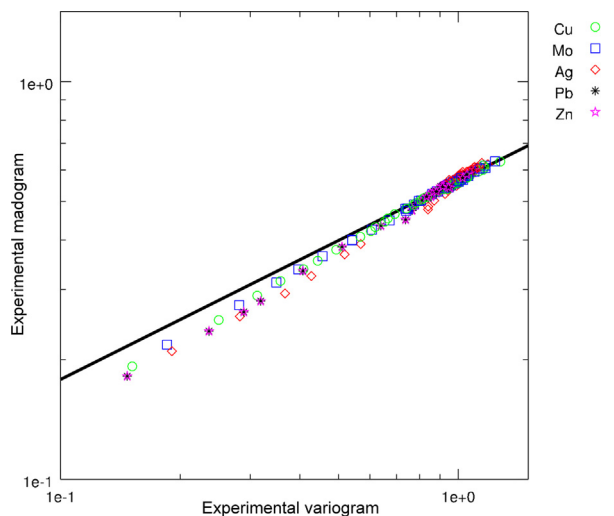
**Table 5**  
Experimental correlation matrix for transformed (Gaussian) variables (Dar-Alu deposit).

	Cu	Mo	Ag	Pb	Zn
Cu	1				
Mo	0.69	1			
Ag	0.42	0.31	1		
Pb	0.10	0.03	-0.17	1	
Zn	0.33	0.05	0.12	0.43	1

isotropic models, using mixtures of six basic nested structures: a nugget effect and isotropic spherical models with ranges 30 m, 100 m, 150 m, 300 m, and 600 m, respectively. The fitting of the coregionalization



**Fig. 3.** (Color online.) An example of scatter plot between Gaussian variables (Cu vs. Mo).



**Fig. 4.** (Color online.) Experimental madograms of the Gaussian variables as a function of their experimental variograms. In case of bivariate normality, the points should be distributed along the thick solid line.

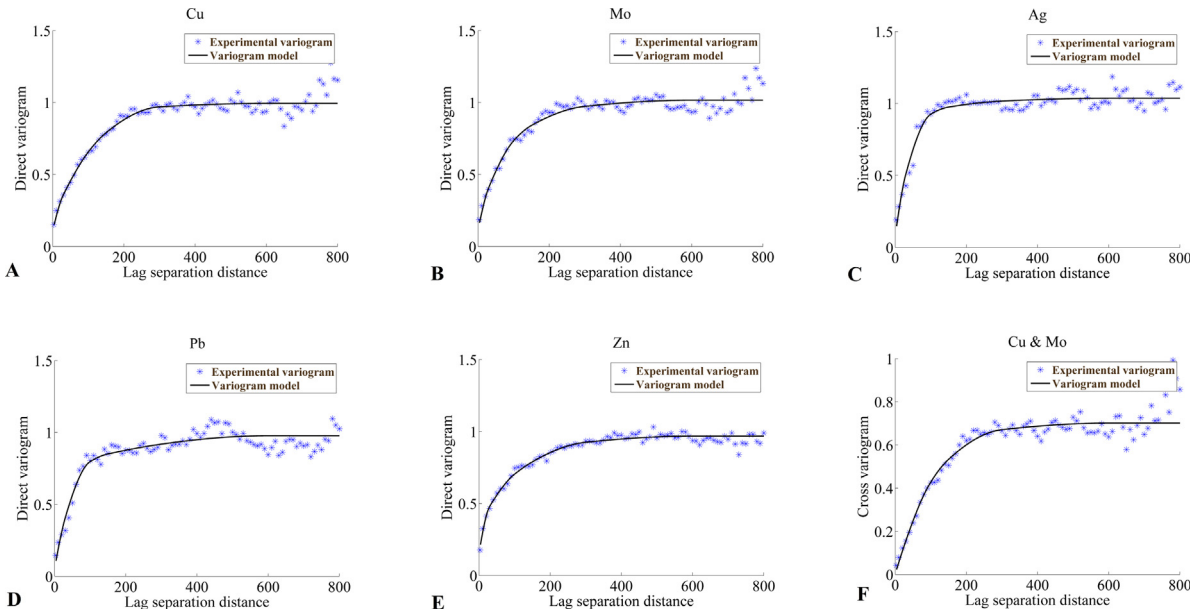


Fig. 5. (Color online.) Sample (crosses) and modeled (solid lines) direct and cross variograms of Gaussian variables. For brevity, only the Cu–Mo cross-variogram is displayed, which corresponds to the pair of variables with the highest cross-correlation, although the fitting has been achieved with all the cross-variograms.

matrices was achieved by recourse to a semi-automated fitting algorithm (Emery, 2010; Goulard and Voltz, 1992). The results of the fitting are partially displayed in Fig. 5. Positive semi-definiteness of the coregionalization matrices (each of size  $5 \times 5$ ) is imposed by the fitting algorithm, which ensures the validity of the coregionalization model.

4.3. Joint simulation of grades

Cosimulation is performed on a regular grid with a mesh of  $10 \text{ m} \times 10 \text{ m} \times 12 \text{ m}$ . In both algorithms (sequential and turning bands), simple cokriging is used with a moving neighborhood of horizontal radius 500 m and vertical radius 250 m, containing up to eight original data points in each octant and, for sequential simulation, up to eight previously simulated points in each octant, in addition to the original data (one may therefore have up to 64 selected points for turning bands and up to 128 selected points for sequential simulation, each point consisting of a

vector with five components). Furthermore, the same conditioning data set (original drill hole data), Gaussian transformation functions, direct and cross variogram models and numbers of realizations (100) are considered for both algorithms. Two variants are tested for sequential simulation: full cokriging and collocated cokriging, as explained in Section 2, in both cases using a multiple-grid strategy with three nested grids. For turning bands, one thousand lines are used, as in the synthetic case study.

In order to validate the results, the distributions of the simulated grades are compared with the distributions of original data, through quantile-quantile plots. It is expected that, on average over many realizations, the plots fit the diagonal lines, indicating an accurate reproduction of the data distributions. This is what effectively happens for the realizations obtained with both turning bands and sequential simulation (Fig. 6).

To check that the realizations also reproduce the dependence relationships between grades, one can calculate the correlation coefficients between the Gaussian transforms of simulated grades in each realization. It is

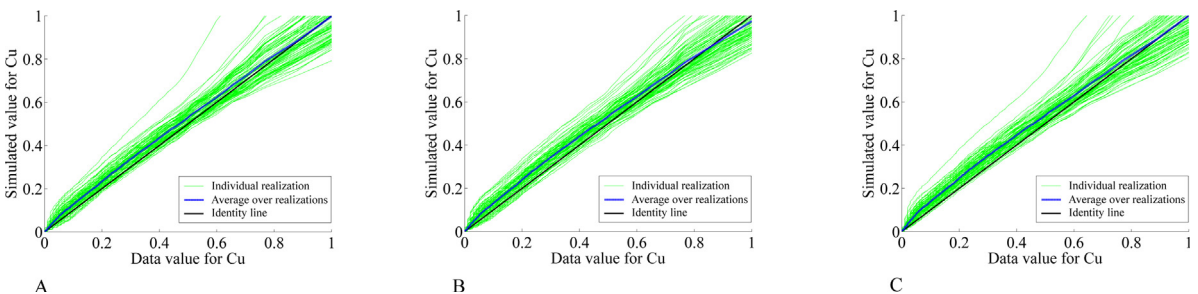


Fig. 6. (Color online.) Quantile–quantile plots between original copper grades and simulated copper grades, for turning bands cosimulation (A), sequential cosimulation with full cokriging (B) and sequential cosimulation with collocated cokriging (C).

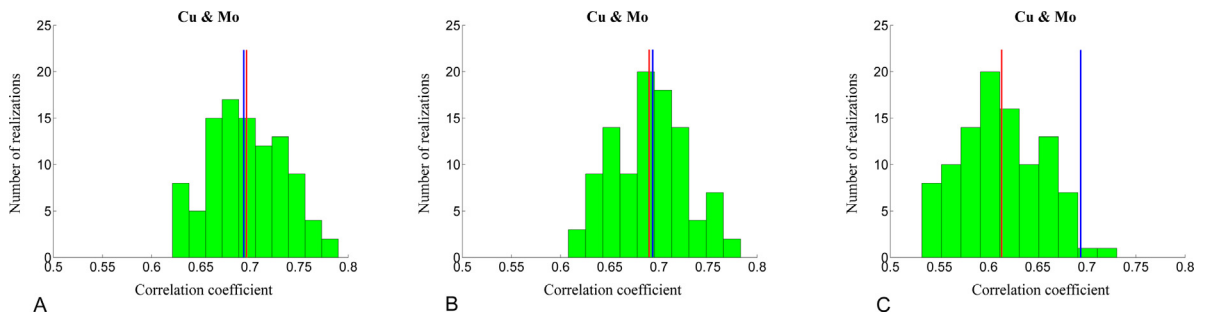


Fig. 7. (Color online.) Correlation coefficients between Gaussian random fields simulated with turning bands (A), sequential cosimulation with full cokriging (B) and sequential cosimulation with collocated cokriging (C). For brevity, only the correlation between copper and molybdenum is displayed. The average correlation is represented with a red line, while the true data correlation is represented with a blue line.

expected that these coefficients fluctuate around the experimental correlations (Table 5), for instance, 0.69 for Cu–Mo. This situation happens for turning bands and sequential cosimulation with full cokriging, but not for sequential cosimulation with collocated cokriging (Fig. 7). In the latter case, the correlations reproduced by the realizations are, on average, lower than the expected correlations. This bias can be explained because, in comparison with full cokriging, collocated cokriging is likely to discard many data of the covariates (Fig. 1), which entails a poor reproduction of the dependence between the variables to be simulated (proof in Appendix, supplementary data).

Another criterion for assessing the quality of each simulation algorithm is to check the reproduction of the direct and cross variograms of the simulated Gaussian

random fields. The results indicate that the direct variograms are well reproduced by all the algorithms (Fig. 8) and the cross variograms are well reproduced with turning bands and, to a lesser extent, with sequential simulation using full cokriging. In contrast, sequential simulation using collocated cokriging leads to an inaccurate reproduction of the cross variograms (Fig. 9), which can be explained because, unless for very specific coregionalization models, such a cokriging is usually a poor approximation of full cokriging (Chilès and Delfiner, 2012; Rivoirard, 2001, 2004). To improve the reproduction of the cross variograms with sequential simulation, an alternative would be to follow the same workflow as for turning bands (Section 2.2), i.e. to (i) factorize the target vector random field, (ii) sequentially simulate the factors, (iii) recombine the simulated factors and finally (iv)

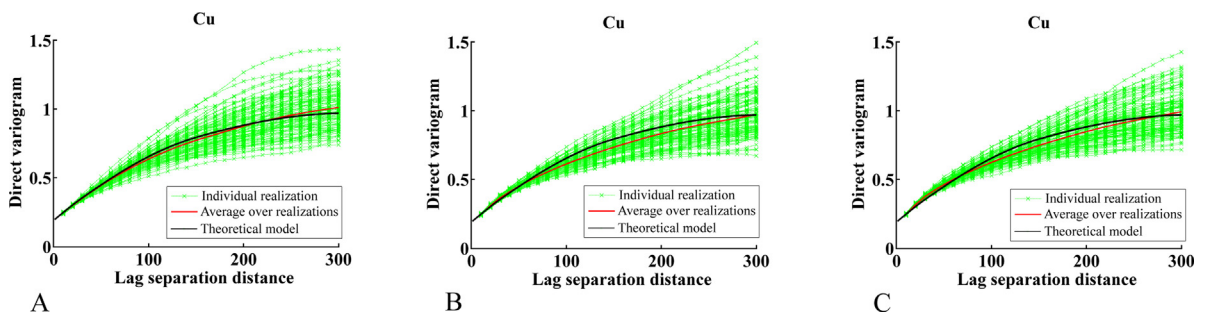


Fig. 8. (Color online.) Direct variograms of Gaussian random fields simulated with turning bands (A), sequential cosimulation with full cokriging (B) and sequential cosimulation with collocated cokriging (C). For brevity, only the variograms for copper are displayed.

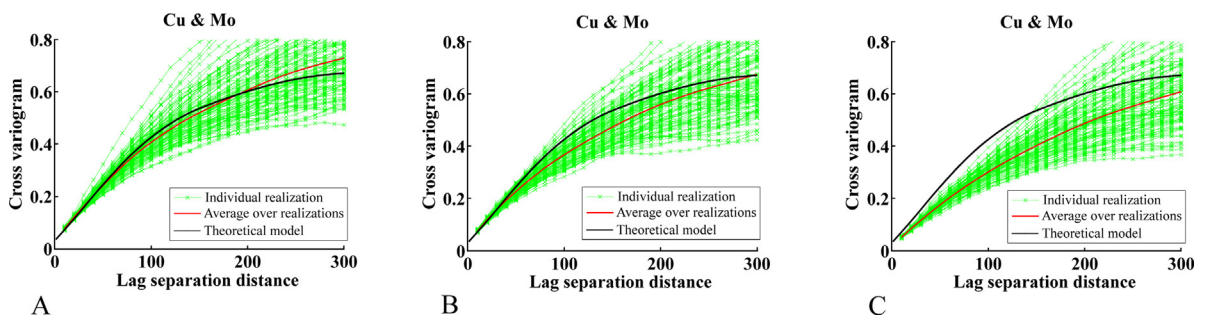


Fig. 9. (Color online.) Cross variograms of Gaussian random fields simulated with turning bands (A), sequential cosimulation with full cokriging (B) and sequential cosimulation with collocated cokriging (C). For brevity, only the cross variograms between copper and molybdenum are displayed.



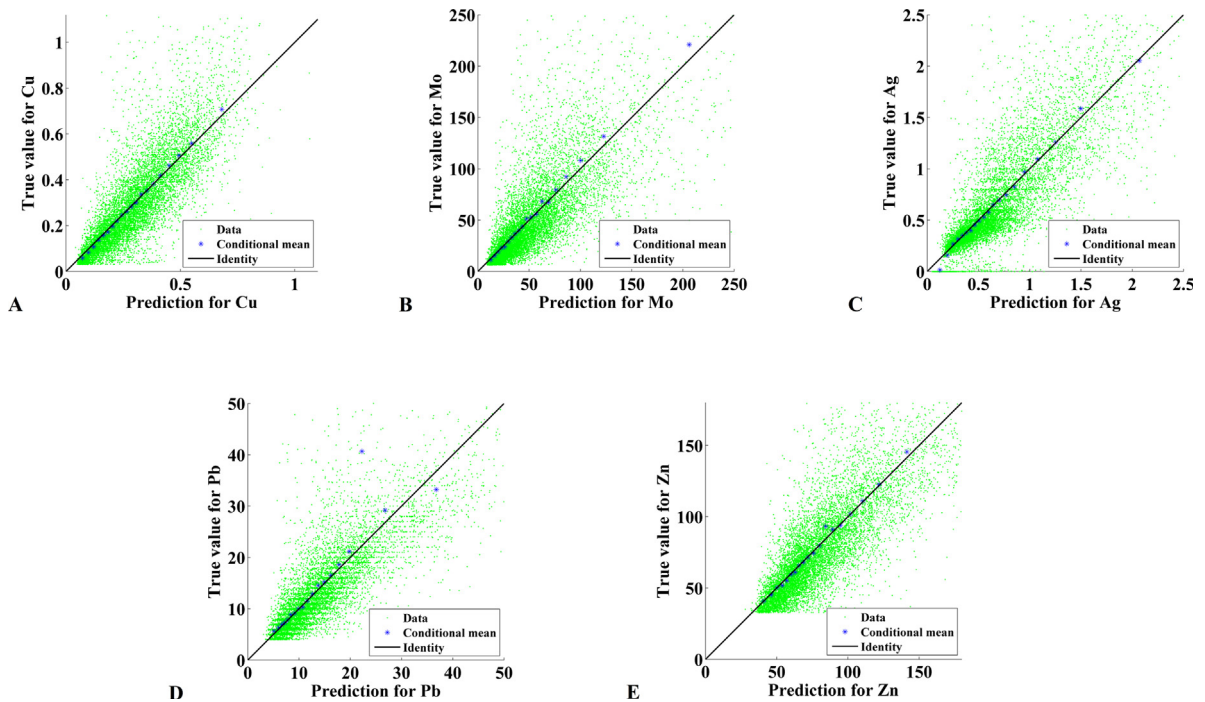


Fig. 10. (Color online.) Scatter plot between true data grades (ordinate) and predicted grades (abscissa) for (A) copper, (B) molybdenum, (C) silver, (D) lead and (E) zinc. Predicted grades are calculated as the average of 100 realizations obtained with turning bands.

condition to the original data by cokriging. This would avoid using cokriging in the sequential simulation stage, insofar as the factors would be simulated separately at step (ii), so that the errors due to a moving neighborhood implementation would not propagate from one factor to another one. Provided that the spatial correlation of the factors is accurately reproduced, so will be the direct and cross variograms of the target vector Gaussian random field.

#### 4.4. Cross validation

The previous analyses allowed assessing the quality of the simulation algorithm (turning bands or sequential) but do not provide much insight into the chosen stochastic model (namely, a Gaussian random field model with the spatial correlation structure given by the variograms fitted in Fig. 5). To corroborate the suitability of this model to the data, a leave-one-out cross-validation exercise is proposed, which consists in cosimulating the grades at each data location, conditionally to the information available at the other locations only, then in comparing the actual data grades with the average of the simulated values (considered as the best prediction of the true grades) at the data locations. For the sake of simplicity, the turning bands algorithm is used for cosimulation.

It is seen (Fig. 10) that, for each element, the average of simulated grades fluctuates around the true data grade and that the conditional regression between true and simulated grades is close to the identity, which reflects conditional unbiasedness (Rivoirard, 1987; Vann et al., 2003). These results indicate that the Gaussian model used to represent

the grade data is adequate in this case study, as the results of simulation are in agreement with the actually observed data.

## 5. Conclusions

This study aimed at assessing the performance of two popular algorithms (turning bands and sequential) for jointly simulating coregionalized variables, through a synthetic case study (univariate) and a real case study (multivariate) from a multi-element deposit. The algorithms are fed with the same inputs: original conditioning data and Gaussian transformation functions in the real case study, variogram models, moving neighborhoods, output grids and numbers of realizations. The only differences are the use of previously simulated points in sequential simulation and the use of a finite number of lines in turning bands simulation, which is explained by the specificities of each algorithm.

In both cases, turning bands accurately reproduces the spatial correlation structure, measured by the direct and cross variograms of the coregionalized variables under consideration, while sequential simulation produces some biases. Such biases are explained by a propagation of errors caused by the use of a moving neighborhood to determine the successive conditional distributions. This problem is more severe in the multivariate case when using collocated cokriging, in which case the reproduction of the cross variograms turns out to be poor. Accordingly, the choice of the simulation algorithm is relevant to accurately reproduce regionalized properties and the turning bands

algorithm appears as an alternative of the sequential algorithm for the modeling of multi-element deposits.

## Acknowledgements

The authors thank the reviewers, Dr. Denis Marcotte and Dr. Joao Felipe Costa, for their recommendations that helped to improve the paper, and to Mr. Esfahanipour R. from the National Iranian Copper Industry Co. (NICICO) for granting access to the Dar-Alu copper deposit exploration data. The last two authors also acknowledge the funding by the Chilean Commission for Scientific and Technological Research, through Project Conicyt/Fondecyt/Regular/No. 1130085.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.crte.2015.05.008>.

## References

- Alabert, F.G., Massonnat, G.J., 1990. Heterogeneity in a complex turbiditic reservoir: stochastic modelling of facies and petrophysical variability. In: 65th Annual Technical Conference and Exhibition of the Society of Petroleum Engineers. SPE-20604, 775–790.
- Almeida, A.S., Journel, A.G., 1994. Joint simulation of multiple variables with a Markov-type coregionalization model. *Math. Geol.* 26 (5), 565–588.
- Black, T.C., Freyberg, D.L., 1990. Simulation of one-dimensional correlated fields using a matrix-factorization moving average approach. *Math. Geol.* 22 (1), 39–62.
- Boisvert, J.B., Rossi, M.E., Ehrig, K., Deutsch, C.V., 2013. Geometallurgical modeling at Olympic Dam mine, South Australia. *Math. Geosci.* 45, 901–925.
- Chilès, J.P., Delfiner, P., 1997. Discrete exact simulation by the Fourier method. In: Baafi, E.Y., Schofield, N.A. (Eds.), *Geostatistics Wollongong*, 96. Kluwer, Dordrecht, The Netherlands, pp. 258–269.
- Chilès, J.P., Delfiner, P., 2012. *Geostatistics: modeling spatial uncertainty*, Second edition. Wiley, New York.
- Davis, M., 1987. Production of conditional simulations via the LU triangular decomposition of the covariance matrix. *Math. Geol.* 19 (2), 91–98.
- Deutsch, C.V., Journel, A.G., 1998. *GSLIB: Geostatistical Software Library and User's Guide*. Oxford University Press, New York.
- Dietrich, C.R., Newsam, G.N., 1993. A fast and exact method for multidimensional Gaussian stochastic simulations. *Water Resour. Res.* 29 (8), 2861–2869.
- Emery, X., 2004. Testing the correctness of the sequential algorithm for simulating Gaussian random fields. *Stoch. Env. Res. Risk. A* 18 (6), 401–413.
- Emery, X., 2005. Variograms of order  $\omega$ : a tool to validate a bivariate distribution model. *Math. Geol.* 37 (2), 163–181.
- Emery, X., 2008a. Statistical tests for validating geostatistical simulation algorithms. *Comput. Geosci.* 34 (11), 1610–1620.
- Emery, X., 2008b. A turning bands program for conditional co-simulation of cross-correlated Gaussian random fields. *Comput. Geosci.* 34 (12), 1850–1862.
- Emery, X., 2010. Iterative algorithms for fitting a linear model of coregionalization. *Comput. Geosci.* 36 (9), 1150–1160.
- Emery, X., Lantuéjoul, C., 2006. TBSIM: a computer program for conditional simulation of three-dimensional Gaussian random fields via the turning bands method. *Comput. Geosci.* 32 (10), 1615–1628.
- Emery, X., Peláez, M., 2011. Assessing the accuracy of sequential Gaussian simulation and cosimulation. *Comput. Geosci.* 15 (4), 673–689.
- Gómez-Hernández, J.J., Cassiraga, E.F., 1994. Theory and practice of sequential simulation. In: Armstrong, M., Dowd, P.A. (Eds.), *Geostatistical simulations*. Kluwer, Dordrecht, The Netherlands, pp. 111–124.
- Gómez-Hernández, J.J., Journel, A.G., 1993. Joint sequential simulation of multigaussian fields. In: Soares, A. (Ed.), *Geostatistics Troia'92*. Kluwer, Dordrecht, The Netherlands, pp. 85–94.
- Goovaerts, P., 1997. *Geostatistics for natural resources evaluation*. Oxford University Press, New York.
- Goulard, M., Voltz, M., 1992. Linear coregionalization model: tools for estimation and choice of cross-variogram matrix. *Math. Geol.* 24 (3), 269–286.
- Journel, A.G., Huijbregts, C.J., 1978. *Mining geostatistics*. Academic Press, London.
- Lantuéjoul, C., 1994. Non conditional simulation of stationary isotropic multigaussian random functions. In: Armstrong, M., Dowd, P.A. (Eds.), *Geostatistical simulations*. Kluwer, Dordrecht, The Netherlands, pp. 147–177.
- Lantuéjoul, C., 2002. *Geostatistical simulation: models and algorithms*. Springer, Berlin.
- Le Ravalec, M., Noetinger, B., 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.
- Leuangthong, O., McLennan, J.A., Deutsch, C.V., 2004. Minimum acceptance criteria for geostatistical realizations. *Natural Resour. Res.* 13 (3), 131–141.
- Marcotte, D., 2012. Revisiting the linear model of coregionalization. In: Abrahamsen, P., Hauge, R., Kolbjørnsen, O. (Eds.), *Geostatistics Oslo 2012*. Springer, Dordrecht, The Netherlands, pp. 67–78.
- Matheron, G., 1973. The intrinsic random functions and their applications. *Advances in applied probability* 439–468.
- Myers, D.E., 1989. Vector conditional simulation. In: Armstrong, M. (Ed.), *Geostatistics*. Kluwer, Dordrecht, The Netherlands, pp. 283–293.
- Omre, H., Sølna, K., Tjelmeland, H., 1993. Simulation of random functions on large lattices. In: Soares, A. (Ed.), *Geostatistics Tróia'92*. Kluwer, Dordrecht, The Netherlands, pp. 179–199.
- Pardo-Igúzquiza, E., Chica-Olmo, M., 1993. The Fourier integral method: an efficient spectral method for simulation of random fields. *Math. Geol.* 25 (2), 177–217.
- Pebesma, E.J., 2004. Multivariable geostatistics in S: the Gstat package. *Comput. Geosci.* 30 (7), 683–691.
- Pyrzc, M.J., Gringarten, E., Frykman, P., Deutsch, C.V., 2006. Representative input parameters for geostatistical simulation. In: Coburn, T.C., Yarus, J.M., Chambers, R.L. (Eds.), *Stochastic modeling and geostatistics: principles, methods, and case studies*, volume II, 5. AAPG Computer applications in geology, pp. 123–137.
- Ravenscroft, P.J., 1994. Conditional simulation for mining: practical implementation in an industrial environment. In: Armstrong, M., Dowd, P.A. (Eds.), *Geostatistical simulations*. Kluwer, Dordrecht, The Netherlands, pp. 79–87.
- Rivoirard, J., 1987. Two key parameters when choosing the kriging neighborhood. *Math. Geol.* 19 (8), 851–856.
- Rivoirard, J., 2001. Which models for collocated cokriging? *Math. Geol.* 33 (2), 117–131.
- Rivoirard, J., 2004. On some simplifications of cokriging neighborhood. *Math. Geol.* 36 (8), 899–915.
- Rossi, M.E., Deutsch, C.V., 2014. *Mineral resource estimation*. Springer, Dordrecht, The Netherlands.
- Salehian, M., Ghaderi, M., 2010. Fluid inclusion evolution at Daralu porphyry copper deposit, southeast Iran. In: *Tectonic Crossroads: Evolving Orogens of Eurasia–Africa–Arabia*. Geological Society of America.
- Shafiei, B., Shahabpour, J., 2008. Gold distribution in porphyry copper deposits of Kerman region, southeastern Iran. *J. Sci. Islam. Repub. Iran* 19 (3), 247–260.
- Shinozuka, M., Jan, C.M., 1972. Digital simulation of random processes and its applications. *J. Sound Vib.* 25 (1), 111–128.
- Taylor, R.D., Hammarstrom, J.M., Piatok, N.M., Seal I.I., R.R., 2012. Arc-related porphyry molybdenum deposit model. Chapter D of mineral deposit models for resource assessment. U.S. Geological Survey Scientific Investigations Report 2010-5070-D, 64 p.
- Tran, T., 1994. Improving variogram reproduction on dense simulation grids. *Comput. Geosci.* 20 (7), 1161–1168.
- Vann, J., Jackson, S., Bertoli, O., 2003. Quantitative kriging neighbourhood analysis for the mining geologist – a description of the method with worked case examples. In: Dominy, S. (Ed.), *Proceedings of the 5th International Mining Geology Conference*. Australasian Institute of Mining and Metallurgy, Melbourne, Australia, pp. 215–223.
- Wackernagel, H., 2003. *Multivariate geostatistics: an introduction with applications*. Springer, Berlin.