



ELSEVIER

Contents lists available at ScienceDirect

## Comptes Rendus Biologies

www.sciencedirect.com



Biodiversity/Biodiversité

## Clustering of ant communities and indicator species analysis using self-organizing maps

Sang-Hyun Park<sup>a,\*</sup>, Shingo Hosoishi<sup>a</sup>, Kazuo Ogata<sup>a</sup>, Yuzuru Kuboki<sup>b</sup><sup>a</sup> Institute of Tropical Agriculture, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka, Japan<sup>b</sup> Yuzuru Kuboki, 2-12-2, Kiyota, Yahata-higashi, Kitakyushu, Japan

## ARTICLE INFO

## Article history:

Received 22 January 2014

Accepted after revision 8 July 2014

Available online 10 August 2014

## Keywords:

Species composition

Habitat

SOM

DCA

Indicator

## ABSTRACT

To understand the complex relationships that exist between ant assemblages and their habitats, we performed a self-organizing map (SOM) analysis to clarify the interactions among ant diversity, spatial distribution, and land use types in Fukuoka City, Japan. A total of 52 species from 12 study sites with nine land use types were collected from 1998 to 2012. A SOM was used to classify the collected data into three clusters based on the similarities between the ant communities. Consequently, each cluster reflected both the species composition and habitat characteristics in the study area. A detrended correspondence analysis (DCA) corroborated these findings, but removal of unique and duplicate species from the dataset in order to avoid sampling errors had a marked effect on the results; specifically, the clusters produced by DCA before and after the exclusion of specific data points were very different, while the clusters produced by the SOM were consistent. In addition, while the indicator value associated with SOMs clearly illustrated the importance of individual species in each cluster, the DCA scatterplot generated for species was not clear. The results suggested that SOM analysis was better suited for understanding the relationships between ant communities and species and habitat characteristics.

© 2014 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

## 1. Introduction

Ants are considered to be one of the most suitable taxa for assessing ecological integrity in terrestrial ecosystems [1–3]. Ant community characteristics are affected by a variety of environmental factors that act on habitats [4–7]. Furthermore, ants are typically relatively sedentary and are thus sensitive to small-scale spatiotemporal changes in the environment [8]. Ants are therefore efficient indicators

of the impact of environmental changes in an integrative and continuous manner.

Although ant community characteristics reflect ecological integrity fairly well, these community data are highly complex and difficult to analyze, primarily because communities consist of numerous species and also because they respond in highly complex manners to natural and anthropogenic pressures in their environments. A variety of ordination analyses have therefore been employed to describe ant communities. Davis and Zigler [9] used detrended correspondence analysis (DCA) to show that ant communities differed significantly between different habitat types and that species composition over time. Sarty et al. [10] clarified the influence of invasion by *Anoplolepis gracilipes* on ant community composition and species abundance using

\* Corresponding author.

E-mail addresses: [shpark@agr.kyushu-u.ac.jp](mailto:shpark@agr.kyushu-u.ac.jp), [logospsh@daum.net](mailto:logospsh@daum.net) (S.-H. Park).

non-metric multidimensional scaling (NMDS). Indeed, DCA and NMDS are used extensively for the analysis of community data, but the methods sometimes fail to show clear clustering and habitat preference of species [11,12].

Self-organizing maps (SOMs), which are another method used in ordination analysis to mine nonlinear data, have been used extensively to characterize patterns in communities since the 1990s [13,14]. The SOM method can efficiently classify a bi-dimensional projection of habitat characteristics; for example, SOMs can illustrate the community response to anthropogenic disturbance at a site and clarify how such disturbance affects species occurrence [15–17].

This study aimed to:

- characterize the distributional patterns of ant communities in various habitats;
- compare SOM and DCA ordination methods;
- visualize distribution patterns of ant species;
- quantitatively estimate the relative importance of each ant species in each habitat.

We applied the SOM and DCA methods to characterize ant community structure and distribution patterns according to land use type. In addition, we employed an indicator value (IndVal) to quantify the importance of each species identified using the SOM map. This approach was proposed by Park et al. [18] and has recently been employed to better explain ecological data [19,20].

## 2. Materials and methods

### 2.1. Study site

Habitat and ant surveys were carried out at 12 study sites in Fukuoka City, Kyushu, Japan, from 1998 to 2012 (Fig. 1). The 12 study sites were classified into nine land use types, which were grouped into one of two categories: open areas (open urban parks, coastal areas, and agricultural

fields) and forested areas (evergreen forests, shrine forests, urban park forests, deciduous forests, bamboo forests, and cedar forests) (Table 1). Each land use type presents ants with a variety of structurally different habitats and biotic factors. In addition, the characteristics of each habitat have been maintained by surrounding environmental conditions and/or continuous management. For instance, the management of open urban parks is directed at preventing the free growth of planted trees, which limits the provision of habitats containing large quantities of decaying nesting material. We resampled two sites (S05 and S06, Table 1) to clarify whether ant communities were influenced by land use type or time.

### 2.2. Sampling method

We used time-unit sampling (TUS), which is a direct sampling method that is well suited for making a comprehensive inventory of species at different study sites. TUS involves collecting as many species as possible by visually and manually searching for ants on the ground surface, under stones or under wood bark, and around the bases of tree trunks. Although direct sampling methods are effective for capturing the majority of species at a site [25], such methods are not well suited for studies involving quantitative comparisons between sites [26]. TUS is designed to address this weakness by repeatedly sampling sites over a fixed period of time. The tools employed to capture the ants were aspirators, sifters, pans, and vials containing 70% ethanol [27].

The time unit and the number of TUS times were considered based on the conditions at each study site. Of the 12 study sites identified, 2 were sampled 10 times, 6 were sampled 8 times, and 6 were sampled 16 times (Table 1). These 164 samples were used to analyze ant community and occurrence patterns of indigenous ant species according to land use type. In this study, we defined unique and duplicate species as those that were only recorded once or twice among the 164 samples.

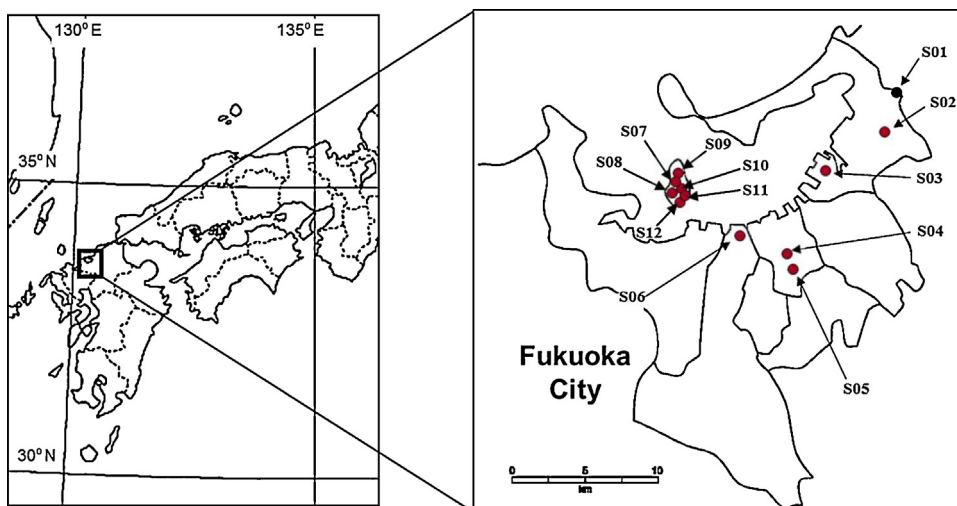


Fig. 1. Map of the location at the 12 study sites in Fukuoka city, Japan.

**Table 1**  
Description of site codes used in Fig. 1.

Site code	Data code	Locality	Land use type	Vegetation type	Repetition of TUS	Reference
S01	A	Mt. Tachibana	Natural forest	Evergreen forest	15 min × 10 times	[21]
S02	B	Kashii-gu Shrine	Shrine	Evergreen forest	15 min × 10 times	[21]
S03	C	Hakozaki Futuo Kouen	Urban open park	Ornamental plant	15 min × 8 times	[21]
S04	D	Gokoku Jinja Shrine	Shrine	Evergreen forest	15 min × 8 times	[21]
S05	E	Minami Kouen	Urban forest park (1998)	Evergreen forest	30 min × 8 times	[22]
	F	Minami Kouen	Urban forest park (2012)	Evergreen forest	30 min × 8 times	[23]
S06	G	Momochi Chuou Kouen	Urban open park (1998)	Ornamental plant	30 min × 8 times	[22]
	H	Momochi Chuou Kouen	Urban open park (2012)	Ornamental plant	30 min × 8 times	[23]
S07	I	Nokonoshima	Seacoast	Grass land	15 min × 16 times	[24]
S08	J	Nokonoshima	Secondary forest	Deciduous forest	15 min × 16 times	[24]
S09	K	Nokonoshima	Agricultural land	Paddy	15 min × 16 times	[24]
S10	L	Nokonoshima	Secondary forest	Evergreen forest	15 min × 16 times	[24]
S11	M	Nokonoshima	Secondary forest	Bamboo forest	15 min × 16 times	[24]
S12	N	Nokonoshima	Secondary forest	Cedar forest	15 min × 16 times	[24]

### 2.3. Data analysis

We used the SOM algorithm described by Kohonen to characterize the distribution patterns of ant communities [28]. SOMs are used to approximate the probability density function of input data, and have frequently been used for clustering, visualization, and abstraction of ecological data [29]. The structure of a SOM consists of an input data layer and an output layer, with each layer connected by connection intensities (or weights). The input layer is composed of 52 neurons (one per ant species) connected to the 14 sampling datasets. We employed a batch training algorithm as it is considerably faster and does not require learning factor rates to be specified [30]. Two approaches were employed to define the map size [29,31]. According to Vesanto's heuristic rule, the actual neuron numbers should be close to  $5\sqrt{\text{sampling sites}}$ . However, Park et al. [29] proposed the use the quantization error (QE) and the topographic error (TE) to assess the map size; in the present study the QE and TE errors were 1.43 and 0, respectively. However, the final decision on the actual number of output neurons in the map used in this study was influenced by applying both the heuristic rule, and calculating both QE and TE. Even so, despite the number of recommended rules obtained in this way, we eventually used a smaller map ( $4 \times 3$ ) as there were a large number of empty output neurons [32]. After training the SOM, we applied Ward's linkage method with the Euclidean distance measure to define the cluster boundaries between different SOM units. In addition, the SOM map employed shades of grey to show the probability of occurrence for each species in a given area, as well as predicting the probability of occurrence for each species at sampling sites where they were not collected consistently at the time of sampling [33]. Training of the SOM and the clustering procedures were performed using Matlab [34]. To test for significant differences among the clusters, we used a multi-response permutation procedure (MRPP), which is a nonparametric method used to test for differences between groups implemented in PC-ORD (version 5.31) [35].

Since SOM analysis does not generate any statistical information that each species is responsible for the

obtained clusters [36], we employed indicator species analysis to identify the most important indicator species in the clusters generated by the SOM analysis [37]. Indicator species were selected by calculating their IndVal, which is the product of the relative abundance of a species and the relative frequency of that species at a site; IndVal ranges from 0 (no indication) to 100 (perfect indication). We considered indicator species to be those for which  $\text{IndVal} > 25$  (i.e., individuals of a particular species that were present in more than 50% of the samples in one cluster, with the relative abundance of that indicator species in the cluster exceeding 50%) [37]. The significance of the indicator values for each species was tested using a Monte Carlo permutation test (999 random permutations).

In addition, we performed a DCA on the same dataset to compare the results obtained using the two methods. DCA is a multivariate statistical technique that is widely used in ecological studies to identify patterns in ecological community data [9,38] and it is also an eigenvector ordination technique based on a correspondence analysis that displays both species and sample units simultaneously [39]. In this study, we used the PC-ORD package to generate DCA scatter plots where each transect was represented by a single point, and then was displayed for clusters of points as an indication of similar ant communities [35].

## 3. Results

### 3.1. Patterning ant assemblages

Using 14 datasets containing the occurrence and distribution data for 52 ant species collected from 12 study sites, we trained the SOM to cluster this information according to similarities in community composition (Table 2, Fig. 2). After training the model, SOM units were separated into two main clusters, X (five datasets) and Y (nine datasets). These clusters were classified based on the dendrogram generated using Ward's linkage method (Fig. 2(B)), which showed that the community composition of the two clusters was significantly different (MRPP,  $A = 0.19$ ,  $P < 0.001$ ).

**Table 2**  
Ants collected by TUS at the 12 sample sites.

Code	Species	Sampling sites																		
		S01		S02		S03		S04		S05		S06		S07		S08	S09	S10	S11	S12
		A	B	C	D	E	F	G	H	I	J	K	L	M	N					
<b>Amblyoponinae</b>																				
01	<i>Amblyopone silvestrii</i>	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1	
<b>Dolichoderinae</b>																				
02	<i>Ochetellus glaber</i>	0	1	0	0	0	0	4	1	8	0	6	0	0	0	0	0	0	0	
<b>Formicinae</b>																				
03	<i>Camponotus japonicus</i>	0	1	3	0	0	1	7	6	0	1	1	0	1	0	0	0	0	0	
04	<i>Camponotus keihitai</i>	0	3	0	0	0	1	0	0	0	2	0	0	0	0	0	0	0	0	
05	<i>Camponotus nipponicus</i>	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
06	<i>Camponotus quadrinotatus</i>	0	0	0	0	0	0	0	0	0	6	0	0	7	0	0	0	0	0	
07	<i>Camponotus vitiosus</i>	0	5	5	0	1	7	4	3	5	3	5	1	0	0	0	0	0	0	
08	<i>Formica hayashi</i>	2	4	0	0	3	4	0	0	0	0	0	0	4	0	0	0	0	0	
09	<i>Formica japonica</i>	0	0	3	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	
10	<i>Lasius alienus</i>	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	
11	<i>Lasius japonicus</i>	6	4	2	0	3	4	7	5	4	8	3	0	4	0	0	0	0	0	
12	<i>Lasius spathepus</i>	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	
13	<i>Nylanderia amia</i>	0	0	0	0	0	0	8	3	0	0	0	0	0	0	0	0	0	0	
14	<i>Nylanderia flavipes</i>	8	10	3	5	8	4	1	0	0	3	3	10	10	5	0	0	0	0	
15	<i>Nylanderia sakurae</i>	0	0	5	0	0	0	7	6	4	0	1	0	0	0	0	0	0	0	
16	<i>Polyrhachis lamellidens</i>	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	
17	<i>Polyrhachis moesta</i>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	
<b>Myrmicinae</b>																				
18	<i>Aphaenogaster famelica</i>	4	1	0	0	8	6	0	0	0	5	0	6	0	6	0	0	6	0	
19	<i>Carebara yamatensis</i>	2	1	0	2	0	0	0	0	0	2	0	1	3	0	0	0	0	0	
20	<i>Crematogaster matsumurai</i>	0	0	4	2	0	4	6	6	3	5	5	0	0	0	0	0	0	0	
21	<i>Crematogaster nawai</i>	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
22	<i>Crematogaster osakensis</i>	2	8	0	0	8	6	0	0	0	4	2	2	0	1	0	0	0	1	
23	<i>Crematogaster teranishii</i>	0	0	0	0	0	2	0	0	0	0	0	2	0	1	0	0	0	1	
24	<i>Crematogaster vagula</i>	0	2	0	0	0	0	0	0	0	6	2	0	0	1	0	0	0	1	
25	<i>Monomorium chinense</i>	0	0	7	0	0	2	7	6	5	0	6	0	0	0	0	0	0	0	
26	<i>Monomorium intrudens</i>	0	0	0	0	5	5	0	2	0	1	2	0	9	1	0	0	0	1	
27	<i>Solenopsis japonica</i>	0	0	1	0	0	0	5	1	2	0	1	0	2	0	0	0	0	0	
28	<i>Strumigenys lewisi</i>	5	5	0	6	1	4	0	0	0	8	0	10	5	10	0	0	0	0	
29	<i>Temnothorax congruus</i>	0	0	1	0	0	0	0	2	0	2	0	0	0	0	0	0	0	0	
30	<i>Temnothorax koreanus</i>	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
31	<i>Temnothorax spinosior</i>	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
32	<i>Tetramorium bicarinatum</i>	0	0	0	0	0	0	8	2	0	0	1	0	0	0	0	0	0	0	
33	<i>Tetramorium nipponense</i>	0	8	0	1	3	3	0	0	0	2	0	0	3	0	0	0	0	0	
34	<i>Tetramorium tsushimae</i>	0	0	5	0	0	0	8	8	8	1	8	0	0	0	0	0	0	0	
35	<i>Pheidole fervida</i>	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
36	<i>Pheidole noda</i>	0	6	7	0	7	7	5	3	7	5	6	1	6	3	0	0	0	0	
37	<i>Pheidole pieli</i>	0	0	0	3	0	4	0	0	0	0	0	0	0	0	0	0	0	1	
38	<i>Pristomyrmex punctatus</i>	3	10	6	2	4	4	6	4	2	6	2	2	4	1	0	0	0	0	
39	<i>Pyramica benten</i>	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	
40	<i>Pyramica canina</i>	0	0	0	0	1	0	0	0	0	4	0	3	0	0	0	0	0	0	
41	<i>Pyramica hexamera</i>	0	0	0	2	0	0	0	0	0	1	0	0	1	0	0	0	0	0	
42	<i>Pyramica membranifera</i>	0	1	1	0	0	0	2	0	1	0	0	0	0	0	0	0	0	0	
43	<i>Vollenhovia benzai</i>	0	5	0	0	6	1	0	0	0	2	0	0	0	0	0	0	0	3	
44	<i>Vollenhovia emeryi</i>	0	1	0	0	0	0	0	0	0	1	0	9	1	2	0	0	0	0	
<b>Ponerinae</b>																				
45	<i>Cryptopone sauteri</i>	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
46	<i>Discothyrea sauteri</i>	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
47	<i>Hypoponera beppin</i>	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	
48	<i>Hypoponera sauteri</i>	1	4	0	0	2	1	0	0	0	1	1	8	1	2	0	0	0	0	
49	<i>Pachycondyla chinensis</i>	6	5	3	5	6	3	1	6	0	7	0	4	3	5	0	0	0	0	
50	<i>Ponera scabra</i>	5	0	0	1	1	0	0	0	0	2	0	5	2	2	0	0	0	0	
<b>Proceratiinae</b>																				
51	<i>Proceratium itoi</i>	4	3	0	1	1	0	0	0	0	2	0	3	0	0	0	0	0	0	
52	<i>Proceratium watasei</i>	0	0	0	0	2	1	0	0	0	0	0	1	0	1	0	0	0	1	
Total		61	89	56	38	74	76	88	65	49	92	56	71	67	46	0	0	0	0	

These clusters reflected the land use types at each of the sampling sites. For example, land use types in cluster X corresponded to open land such as coastal areas, agricultural land and urban parks, while the datasets in cluster Y

corresponded to different forest types. Cluster Y could further be subdivided into two sub-clusters, Y<sub>1</sub> and Y<sub>2</sub>, each of which contained three and six datasets, respectively. The communities collected from evergreen forests and

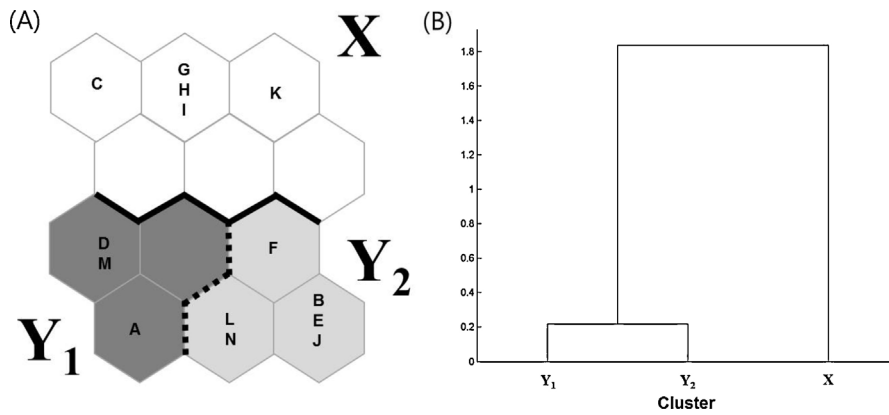


Fig. 2. A. Distribution of the 14 sampled sites on SOM based on the occurrence of 52 ant species. B. Dendrogram for three clusters based on the Ward's linkage method.

shrine forests were occurred in both sub-clusters Y<sub>1</sub> and Y<sub>2</sub> (Fig. 2A). In addition, the community compositions of some study sites (e.g., S05 and S06) did not change over time.

DCA scatter plots also revealed three distinct groups of sampling sites that could be correlated with land use types (Fig. 3), corroborating the findings obtained by the SOM. However, because unique and duplicate species were occasionally collected during the 164 collection times, we re-analyzed the data after removing these samples from our analysis. The results showed that the clusters generated by the DCA changed, whereas those generated by the SOM remained unaffected (Fig. 4).

### 3.2. Distribution of species

The probability of occurrence for each indicator species (i.e., species with significant IndVal values) illustrates

how the distribution of ant species contributed to each cluster on SOM map (Fig. 5). Among 52 species identified in total, 17 were selected as indicator species ( $P < 0.05$ ). In cluster X, seven species (*Ochetellus glaber*, *Nylanderia sakurae*, *Crematogaster matsumurai*, *Monomorium chinense*, *Solenopsis japonica*, *Tetramorium bicarinatum*, and *Tetramorium tsushimae*) were selected as indicator species. Of the seven species in sub-cluster Y<sub>1</sub>, two species (*Carebara yamatonis* and *Cryptopone sauteri*) were indicator species. In sub-cluster Y<sub>2</sub>, *Amblyopone silvestrii*, *Aphaenogaster famelica*, *Crematogaster osakensis*, *Vollenhovia benzai*, and *Proceratium watasei* were indicator species. *Ponera scabra* was represented as an indicator species in both sub-clusters Y<sub>1</sub> and Y<sub>2</sub>. *Camponotus vitosus* and *Pheidole noda* were represented in both cluster X and sub-cluster Y<sub>2</sub>.

### 4. Discussion

According to the similarity of community compositions, the 14 sampling sites examined in this study were divided into three clusters by training the SOM. The different clusters not only reflected differences in species composition, but also differences in land use types between sampling sites (Table 1, Fig. 2). Based on the obtained clusters, the differences between cluster X (open environments) and cluster Y (forested environments) were intuitive within the context of land use type (Fig. 2). However, at higher levels of resolution (e.g., the different forest types contained within sub-cluster Y<sub>1</sub> and Y<sub>2</sub>), the effect of land use type was less apparent as there was some overlap between species in areas of evergreen and shrine forest (Fig. 2A). Although sub-clusters Y<sub>1</sub> and Y<sub>2</sub> were characterized as having the same land use type and similar vegetation type, the environmental conditions in these areas were different. We therefore examined instances in which land use type was the same but the samples were split between sub-clusters Y<sub>1</sub> and Y<sub>2</sub> in greater detail. The findings showed that, for example, study site S01 was located in the saddle of a mountain ridge while site S04 was located on topographically flat ground in the center of an urban area, and site S11 had no weeds on the forest

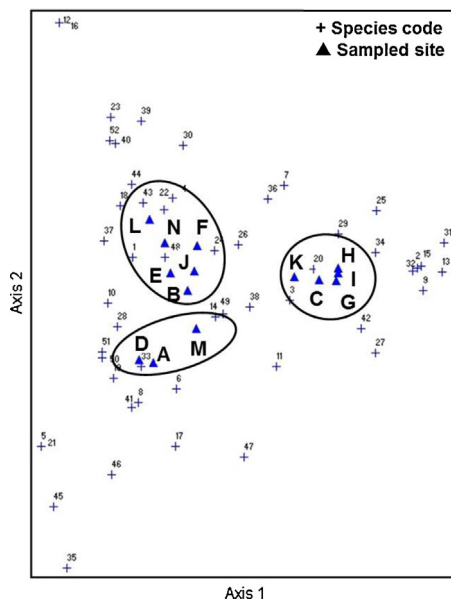
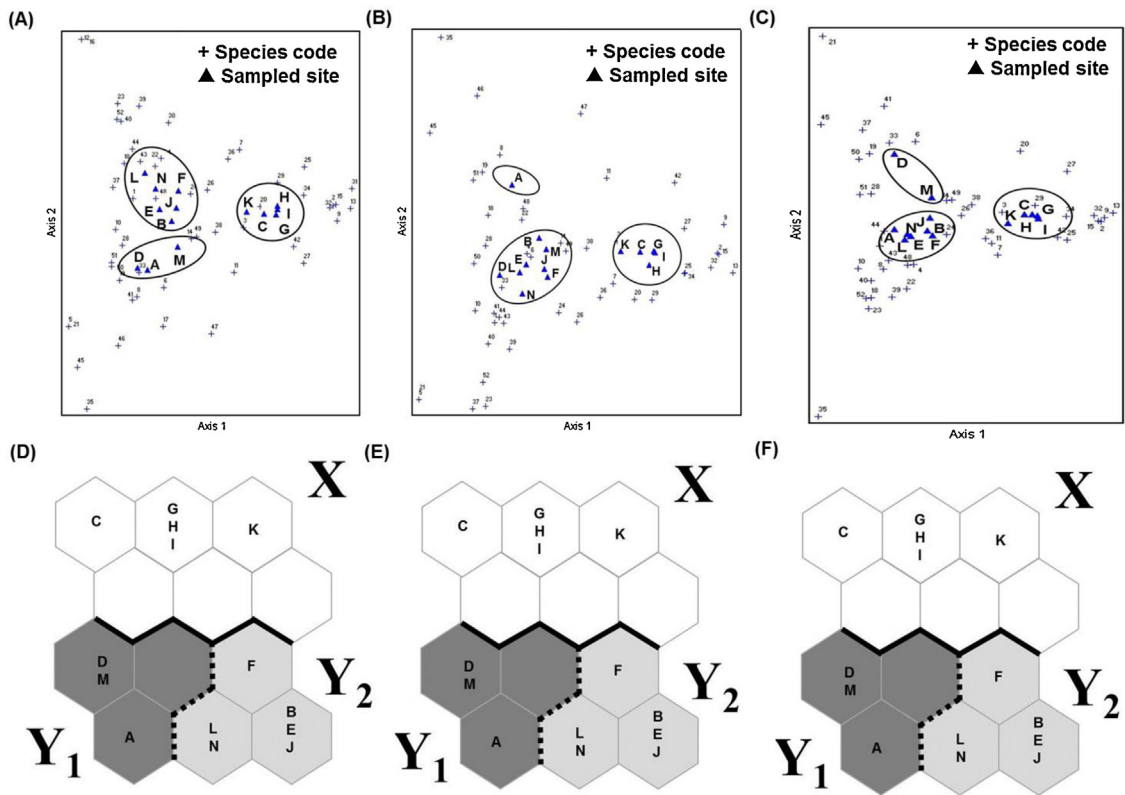


Fig. 3. Distribution of the 14 sampled sites based on DCA of occurrence of 52 ant species.



**Fig. 4.** Comparison of SOM and DCA after removal of the unique and duplicate species. A. DCA of all species. B. DCA after removal of unique species. C. DCA after removal of unique and duplicate species. D. SOM analysis of all species. E. SOM analysis after removal of unique species. F. SOM analysis after removal of unique and duplicate species.

floor. Compared to the forested sites in sub-cluster  $Y_2$ , all three of these sites were relatively dry. It therefore seems likely that the differences between sub-clusters  $Y_1$  and  $Y_2$  may be related to microhabitat conditions in forests.

The results of the SOM analysis corroborated the findings of the DCA based on the same data. However, given the likelihood that unique and duplicate species could easily be overlooked in the total inventory for each sampling site, we re-analyzed the data after excluding such species. The clustering produced by DCA changed, while that produced by the SOM method remained unaffected. We therefore consider that SOM clustering is more robust than DCA clustering.

In addition, DCA scatterplots of species clusters were less well defined, with the data for species being highly dispersed in the multivariate state space of the scatterplot. It was therefore difficult to understand the relationship between the obtained clusters and distinct species [12]. However, the SOM method with IndVal values clearly distinguished between the different species in each cluster displaying statistical values (Fig. 5).

Among the indicator species in cluster X, *Tetramorium tsushimae* was shown to prefer open land [Fig. 5], which has been reported previously [40]. Similarly, *Tetramorium bicarinatum*, an invasive species with a global distribution, has been reported to be particularly common in open

habitats, such as farmland, urban parks, and coastal areas [Fig. 5] [41,42]. In addition, *Pheidole noda*, an indicator species in cluster X and sub-cluster  $Y_2$ , has been recorded in open areas and woodland [Fig. 5] [40]. Thus, previous reports corroborated our results based on IndVal values and SOM clusters [40–42]. Other indicator species, such as *N. sakurae*, *V. benzai*, and *C. sauteri*, were present in all habitats, albeit in limited numbers. Taken together, these results indicated that ant species were not randomly distributed and that the distribution patterns of ant species were related to habitat suitability [43].

In conclusion, the SOM was found to be an effective analytical tool for understanding the responses of ant communities and individual species to different habitat characteristics, particularly given the consistency of SOM results and the ability to assess the relative importance of each species in each habitat. Finally, the ant community structure in a habitat within the same land use type was not affected by different collectors or collection times, indicating that ants are particularly well suited for monitoring ecosystem conditions. The application of SOM methods to analyze ant communities could therefore be used as an ecological map with which to illustrate changes in community distribution in response to environmental stressors and to develop a general strategy for conserving biodiversity in urban environments.

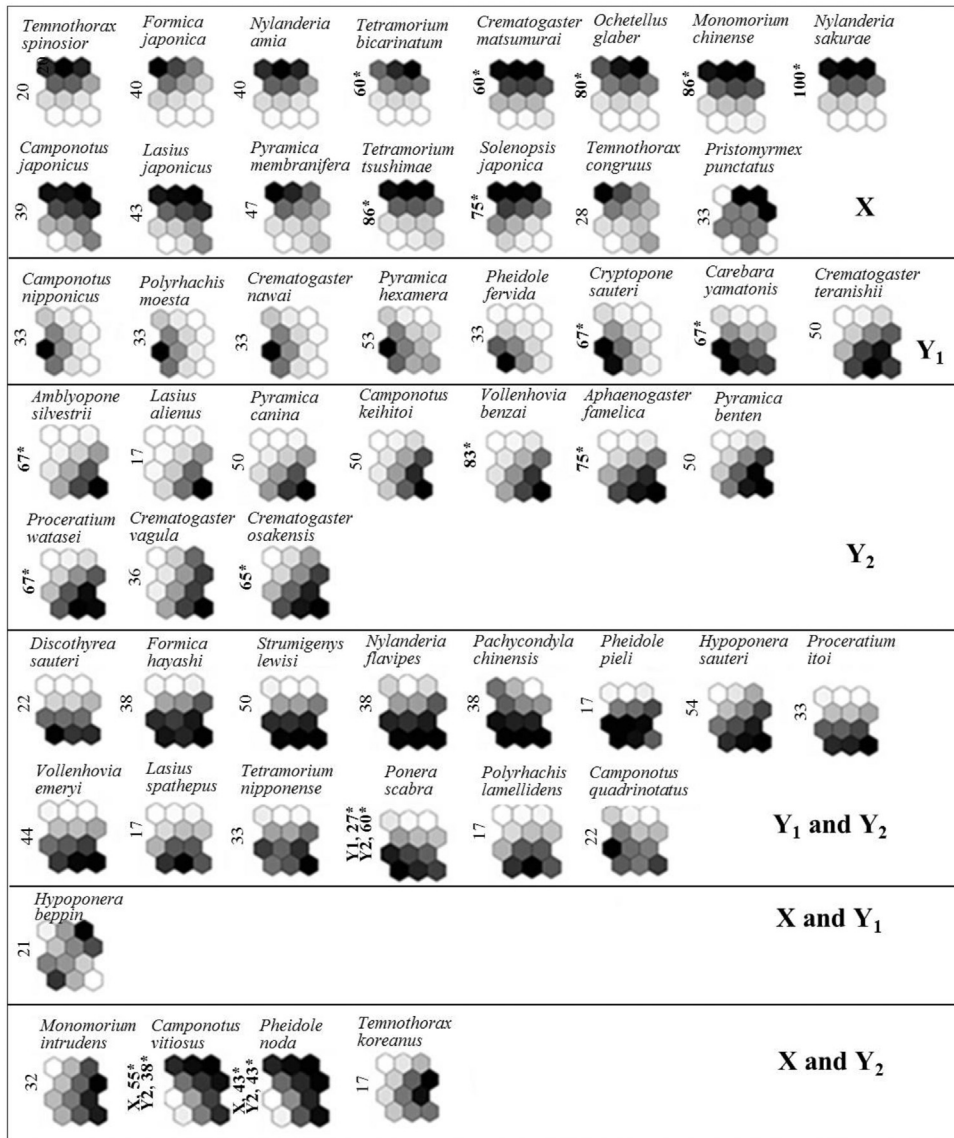


Fig. 5. Probability of occurrence of each species on the trained SOM, visualized by a shaded scale (dark=high probability of occurrence, light=low probability of occurrence). IndVal and significance level (\* $P < 0.05$ ) obtained from Monte Carlo permutations for each ant species for the three different clusters defined in the SOM.

## References

- [1] A.N. Andersen, J.D. Majer, Ants show the way Down Under: invertebrates as bioindicators in land management, *Front. Ecol. Environ.* 2 (2004) 291–298.
- [2] E.C. Underwood, B.L. Fisher, The role of ants in conservation monitoring: if, when, and how, *Biol. Conserv.* 132 (2006) 166–182.
- [3] T.O. Crist, Biodiversity, species interactions, and functional roles of ant (Hymenoptera: Formicidae) in fragmented landscapes: a review, *Myrmecol. News* 12 (2009) 3–13.
- [4] H. Gibb, D.F. Hochuli, Colonisation by a dominant ant facilitated by anthropogenic disturbance: effects on ant assemblage composition, biomass and resource use, *Oikos* 103 (2003) 469–478.
- [5] D.A. Holway, S.V. Suarez, Homogenization of ant communities in Mediterranean California: the effects of urbanization and invasion, *Biol. Conserv.* 127 (2006) 319–326.
- [6] L. Barrow, C.L. Parr, J.L. Kohen, Habitat type influences fire resilience of ant assemblages in the semi-arid tropics of northern Australia, *J. Arid Environ.* 69 (2007) 80–95.
- [7] S. Groc, J.H.C. Delabie, F. Fernández, M. Leponce, J. Orivel, R. Silvestre, H.L. Vasconcelos, A. Dejean, Leaf-litter ant communities (Hymenoptera: Formicidae) in a pristine Guianese rain-forest: stable functional structure versus high species turnover, *Myrmecol. News* 19 (2014) 43–51.
- [8] B. Bestelmeyer, J.A. Wiens, The effects of land use on the structure of ground-foraging ant communities in the Argentine Chaco, *Ecol. Appl.* 6 (1996) 1225–1240.
- [9] R.A. Davis, K.S. Zigler, Ant (Hymenoptera: Formicidae) communities of the southern Cumberland Plateau, *Ann. Entomol. Soc. Am.* 105 (3) (2012) 484–492.
- [10] M. Sarty, K.L. Abbott, P.J. Lester, Community level impacts of an ant invader and food mediated coexistence, *Insect. Soc.* 54 (2007) 166–173.
- [11] M.W. Palmer, Ordination methods for ecologists, 2006 <http://ordination.okstate.edu/overview.htm>.
- [12] A. Kruk, S. Lek, Y.-S. Park, T. Peczek, Fish assemblages in the large lowland Narew River systems (Poland): Application of the self-organizing map algorithm, *Ecol. Model.* 203 (2007) 45–61.
- [13] T.-S. Chon, Y.S. Park, K.H. Moon, E.Y. Cha, Patternizing communities by using an artificial neural network, *Ecol. Model.* 90 (1996) 69–78.

- [14] R. Céréghino, C. Leroy, A. Dejean, B. Corbara, Ants mediate the structure of phytotelm communities in an ant-garden bromeliad, *Ecology* 91 (2010) 1549–1556.
- [15] M.-Y. Song, H.-J. Hwang, I.-S. Kwak, C.W. Ji, Y.-N. Oh, B.J. Youn, T.-S. Chon, Self-organization mapping of benthic macroinvertebrate communities implemented to community assessment and water quality evaluation, *Ecol. Model.* 203 (2007) 18–25.
- [16] Y.-S. Park, J. Tison, S. Lek, J.-L. Giraudel, M. Coste, F. Delmas, Application of a self-organizing map to select representative species in multivariate analysis: a case study determining diatom distribution patterns across France, *Ecol. Inform.* 3 (2006) 247–257.
- [17] J.H.C. Delabie, R. Céréghino, S. Groc, A. Dejean, M. Gibernau, B. Corbara, A. Dejean, Ants as biological indicators of Wayana Amerindian land use in French Guiana, *C.R. Biologies* 332 (2009) 673–684.
- [18] Y.S. Park, T. Oberdorff, S. Lek, Patterning riverine fish assemblages using an unsupervised neural network, in: S. Lek, M. Scardi, P. Verdonschot, J.P. Verdonschot, Y.S. Park (Eds.), *Modelling community structure in freshwater ecosystems*, Springer, Berlin, 2005, pp. 43–53.
- [19] E. Lasne, B. Bergerot, S. Lek, P. Laffille, Fish zonation and indicator species for the evaluation of the ecological status of rivers: example of the Loire basin (France), *River Res. Appl.* 23 (2007) 877–890.
- [20] T. Penczak, L. Glowacki, A. Kruk, W. Galicka, Implementation of a self-organizing map for investigation of impoundment impact of fish assemblages in a large, lowland river: long-term study, *Ecol. Model.* 227 (2012) 64–71.
- [21] Y. Kuboki, Biodiversity assessment of ant communities – ants as bioindicators –, Kyushu University, 2010 Ph.D. dissertation [In Japanese].
- [22] K. Ogata, Y. Takematsu, S. Urano, Species diversity of ants in two urban parks (Hymenoptera: Formicidae), *Bull. Inst. Trop. Agr. Kyushu Univ.* 21 (1998) 57–66.
- [23] S.-H. Park, S. Hosoishi, K. Ogata, E. Kasuya, Changes of species diversity of ants over time: a case study in two urban parks, *J. Fac. Agr., Kyushu Univ.* 59 (2014) 71–76.
- [24] S. Hosoishi, Ant fauna of Noko Island, in: *Univ. Kyushu, Bio-Researching Club (Eds.), The flora and fauna of Noko Island*, Ishizaki Press, Fukuoka, Japan, 2006, pp. 99–107 [In Japanese].
- [25] J.R. King, S.D. Porter, Evaluation of sampling methods and species richness estimators for ants in upland ecosystems in Florida, *Environ. Entomol.* 34 (2005) 1566–1578.
- [26] B.T. Bestelmeyer, D. Agosti, L.E. Alonso, C.R.F. Brandão, W.L. Brown Jr., J.H.C. Delabie, R. Silvestre, Field techniques for the study of ground-dwelling ants: and overview, description, and evaluation, in: D. Agosti, J.D. Majer, L.E. Alonso, T.R. Schultz (Eds.), *Ants: standard methods for measuring and monitoring biodiversity*, Smithsonian Institution Press, Washington, D.C., USA, 2000, pp. 122–144.
- [27] K. Ogata, Time unit sampling: a protocol, *Anet. Newslett.* 3 (2001) 18–19.
- [28] T. Kohonen, *Self-Organizing Maps*, third edition, Springer, Berlin, 2001.
- [29] Y.-S. Park, R. Céréghino, A. Compin, S. Lek, Applications of artificial neural networks for patterning and predicting aquatic insect species richness in running waters, *Ecol. Model.* 160 (2003) 265–280.
- [30] R. Céréghino, Y.-S. Park, Review of the self-organizing map (SOM) approach in water resources: commentary, *Environ. Model. Softw.* 24 (2009) 945–947.
- [31] J. Vesanto, E. Alhoniemi, Clustering of the self-organizing map, *IEEE Trans. Neural Netw.* 11 (2000) 586–600.
- [32] T. Penczak, L. Glowacki, A. Kruk, W. Galicka, Implementation of a self-organizing map for investigation of impoundment impact on fish assemblages in a large, lowland river: long-term study, *Ecol. Model.* 227 (2012) 64–71.
- [33] A. Ruggiero, R. Céréghino, J. Figuerola, P. Marty, S. Angélibert, Farm ponds make a contribution to the biodiversity of aquatic insects in a French agricultural landscape, *C.R. Biologies* 331 (2008) 298–308.
- [34] MathWorks, *The Matlab*, Version 6.1, The MathWorks, Inc., 2001.
- [35] E.W. Mielke, K.J. Berry, E.S. Johnson, Multi-response permutation procedures for a priori classifications, *Commun. Stat. Theory* 5 (1976) 1409–1424.
- [36] Y.S. Park, M. Geyrev, S. Lek, J.-L. Giraudel, Evaluation of relevant species in communities: development of structuring indices for the classification of communities using a self-organizing map, in: S. Lek, M. Scardi, P. Verdonschot, J.-P. Descy, Y.S. Park (Eds.), *Modelling community structure in Freshwater Ecosystems*, Springer, Berlin, 2005, pp. 369–380.
- [37] M. Dufrière, P. Legendre, Species assemblages and indicator species: the need for a flexible asymmetrical approach, *Ecol. Monogr.* 67 (1997) 345–366.
- [38] B.T. Bestelmeyer, J.A. Wiens, The effects of land use on the structure of ground-foraging ant communities in the argentine Chaco, *Ecol. Appl.* 6 (4) (1996) 1225–1240.
- [39] M.O. Hill, H.G. Gauch, Detrended correspondence analysis: an improved ordination technique, *Vegetatio* 42 (1980) 47–58.
- [40] Japanese ant database, group, Japanese ant image database, 2003 <http://ant.edb.miyakyo-u.ac.jp/E/index.html>.
- [41] F. Hita Garcia, B.L. Fisher, The ant genus *Tetramorium* Mayr (Hymenoptera: Formicidae) in the Malagasy region-introduction, definition of species groups, and revision of the *T. bicarinatum*, *T. obseum*, *T. sericeiventris* and *T. tosi* species groups, *Zootaxa* 3039 (2011) 1–72.
- [42] J.K. Wetterer, Worldwide spread of the Penny ant, *Tetramorium bicarinatum* (Hymenoptera: Formicidae), *Sociobiology* 54 (3) (2009) 1–20.
- [43] J.-P. Lessard, C.M. Buddle, The effects of urbanization on ant assemblages (Hymenoptera: Formicidae) associated with the Molson Nature Reserve, Quebec, *Can. Entomol.* 137 (2005) 215–225.