REVIEWS

Community ecology in the age of multivariate multiscale spatial analysis

S. Dray,[1,16] R. Pélissier,[2,3] P. Couteron,[2] M.-J. Fortin,[4] P. Legendre,[5] P. R. Peres-Neto,[6] E. Bellier,[7,8] R. Bivand,[9] F. G. Blanchet,[10] M. De Cáceres,[11] A.-B. Dufour,[1] E. Heegaard,[12] T. Jombart,[1,13] F. Munoz,[2] J. Oksanen,[14] J. Thioulouse,[1] and H. H. Wagner[15]

[1]*Université de Lyon, F-69000, Lyon; Université Lyon 1; CNRS, UMR5558, Laboratoire de Biométrie et Biologie Evolutive, F-69622, Villeurbanne, France*
[2]*Institut de Recherche pour le Développement (IRD), Université Montpellier 2, UMR Botanique et Bioinformatique de l'Architecture des Plantes (AMAP), Boulevard de la Lironde, TA A-51/PS2, F-34398 Montpellier cedex 5, France*
[3]*Institut Français de Pondichéry, UMIFRE 21 CNRS-MAEE, 11 St Louis Street, Puducherry 605 001 India*
[4]*Department of Ecology and Evolutionary Biology, University of Toronto, 25 Harbord Street, Toronto, Ontario M5S 3G5 Canada*
[5]*Département de Sciences Biologiques, Université de Montréal, C.P. 6128, Succursale Centre-ville, Montréal, Québec H3C 3J7 Canada*
[6]*Département des Sciences Biologiques, Université du Québec à Montréal, C.P. 8888, Succursale Centre-ville, Montréal, Québec H3C 3P8 Canada*
[7]*Unité Biostatistique et Processus Spatiaux, INRA Avignon, Domaine Saint-Paul, Site Agroparc 84914 Avignon cedex 9, France*
[8]*Norwegian Institute for Nature Research, NINA-NO-7485, Trondheim, Norway*
[9]*Department of Economics, NHH Norwegian School of Economics, Helleveien 30, N-5045 Bergen, Norway*
[10]*Department of Renewable Resources, University of Alberta, 751 General Services Building, Edmonton, Alberta T6G 2H1 Canada*
[11]*Biodiversity and Landscape Ecology Laboratory, Forest Science Center of Catalonia, Ctra. Antiga St Llorenç km 2, E-25280, Solsona, Catalonia, Spain*
[12]*Norwegian Forest and Landscape Institute, Fanaflaten 4, N-5244 Fana, Norway*
[13]*Department of Infectious Disease Epidemiology, Imperial College London, London W21PG United Kingdom*
[14]*Department of Biology, University of Oulu, FI-90014 Oulu, Finland*
[15]*Department of Ecology and Evolutionary Biology, University of Toronto, 3359 Mississauga Road, Mississauga L5L 1C6 Canada*

Abstract. Species spatial distributions are the result of population demography, behavioral traits, and species interactions in spatially heterogeneous environmental conditions. Hence the composition of species assemblages is an integrative response variable, and its variability can be explained by the complex interplay among several structuring factors. The thorough analysis of spatial variation in species assemblages may help infer processes shaping ecological communities. We suggest that ecological studies would benefit from the combined use of the classical statistical models of community composition data, such as constrained or unconstrained multivariate analyses of site-by-species abundance tables, with rapidly emerging and diversifying methods of spatial pattern analysis. Doing so allows one to deal with spatially explicit ecological models of beta diversity in a biogeographic context through the multiscale analysis of spatial patterns in original species data tables, including spatial characterization of fitted or residual variation from environmental models. We summarize here the recent progress for specifying spatial features through spatial weighting matrices and spatial eigenfunctions in order to define spatially constrained or scale-explicit multivariate analyses. Through a worked example on tropical tree communities, we also show the potential of the overall approach to identify significant residual spatial patterns that could arise from the omission of important unmeasured explanatory variables or processes.

Key words: ecological community; multivariate spatial data; ordination; spatial autocorrelation; spatial connectivity; spatial eigenfunction; spatial structure; spatial weight.

INTRODUCTION

A major concern of ecology is the identification and explanation of the spatial patterns of ecological structures (species distributions, composition, or diversity [Legendre 1993]). The presence and abundance of

individual species vary through space in a nonrandom way, displaying spatial structures. Hence, community composition is usually not random, so that beta diversity, defined as the variation in community composition (Whittaker 1972), displays spatial patterns. Considering the fact that there is no single natural scale at which ecological processes occur, Levin (1992) suggested that attention should focus on the notion of scale as a means to understanding patterns in natural variation. It is now well established that patterns observed in communities at a given scale are often the consequence of a complex interplay between various processes occurring at multiple scales (Menge and Olson 1990). For instance, the interactions between individual organisms and their local environment, including immediate neighbors, influence community composition (Agrawal et al. 2007), while regional and historical processes can profoundly influence local community structure (Ricklefs 1987, Borcard and Legendre 1994). An important research goal in ecology is thus to identify the most relevant scales to account for compositional variation and model species–environment interactions, which may change with the scale of observation (Dungan et al. 2002, Legendre et al. 2009). Searching for and analyzing spatial structures at multiple scales allows researchers to test hypotheses about the mechanisms through which species diversity evolves or is maintained in ecosystems, a topic of key importance, for instance, for the development of conservation policies and the design of biodiversity protection areas (e.g., Seiferling et al., *in press*).

Inferring processes underlying community structure depends heavily on our capacity to detect patterns in species distribution data, given that manipulative (field and laboratory) experiments cannot, in most instances, be extensive enough to detect the effects of mechanisms that influence communities at large spatial scales (Currie 2007). However, "because changes in process intensity can create different patterns, and because several different processes can generate the same pattern signature" (Fortin and Dale 2005:3), community ecology faces the challenge to draw clear links between patterns and processes (Vellend 2010). Recently, McIntire and Fajardo (2009) proposed a conceptual framework to enhance inference in ecology using space as a surrogate for uncovering unmeasured or unmeasurable ecological processes through the analysis of spatial patterns and/or spatial residuals. They argued that the analysis of spatial patterns and associated properties (e.g., intensity, patchiness, scale, clustering) can help to infer ecological processes under the condition that precise alternative hypotheses are well defined a priori and that appropriate statistical methods are used to select the most likely hypothesis. In this context, technological advances to acquire, manage, and store large georeferenced data sets (e.g., as obtained routinely through remote sensing, geographic information system, etc.) and recent methodological developments opened new avenues to test hypotheses concerning the spatial organization of ecological structures at different spatial scales. While McIntire and Fajardo (2009) mainly developed the conceptual and heuristic aspects of drawing inferences from spatial analysis, we focus here on the most recent methods and tools for the multiscale analysis of spatial structures observed in ecological communities.

Traditionally, questions about the structure and determinants of ecological communities have been tackled using multivariate analyses (Gauch 1982). The last decade has seen several methodological developments designed to make the multivariate analysis of species assemblages more spatially explicit and to generalize analyses of spatial distributions to handle multi-species communities. Now that large-scale georeferenced data sets, sophisticated statistical methods, and adequate computing power are available, the modeling of ecological patterns can be done in much greater detail. Depending on the nature of the data considered, such models are highly relevant (see Fig. 1). They may be used for (1) detecting and characterizing spatial patterns (e.g., is a community organized into fairly discrete homogeneous regions (patches) or distributed along smooth gradients?); (2) determining whether spatial variation in community composition can be explained by measured environmental factors (e.g., is community variation explained by environmental variables? Do we observe significant remaining spatial structures that are not explained by measured environmental descriptors?); (3) identifying characteristic scales of spatial structures (e.g., at which spatial scales is the variation in community composition well modeled by environmental factors? At which other scales do we observe spatial structures that are not explained by environmental descriptors?)

As multivariate spatially explicit methods have recently been developed and diversified (e.g., Borcard et al. 1992, Borcard and Legendre 2002, Wagner 2004, Dray et al. 2006, Ferrier and Guisan 2006, Lichstein 2007, Soininen et al. 2007, Blanchet et al. 2008a), we aim to highlight here how they can be used to address the above questions. While much effort has focused on the nuisance aspect of spatially dependent observations in statistical models (see Dormann et al. [2007] for a review), we adopt another viewpoint based on the idea that the presence of any nonrandom spatial structure in species data has a biological, historical or environmental cause. Hence, following McIntire and Fajardo (2009), we suggest that explicitly introducing the spatial component as a proxy (space as a surrogate) into analyses of community composition data helps to identify potential underlying processes that may be difficult to measure directly from field studies. In many situations, species assemblages may display spatial dependence induced by responses to spatially structured explanatory variables, and also spatial autocorrelation due to the population dynamics of the response variables themselves (Peres-Neto and Legendre 2010).
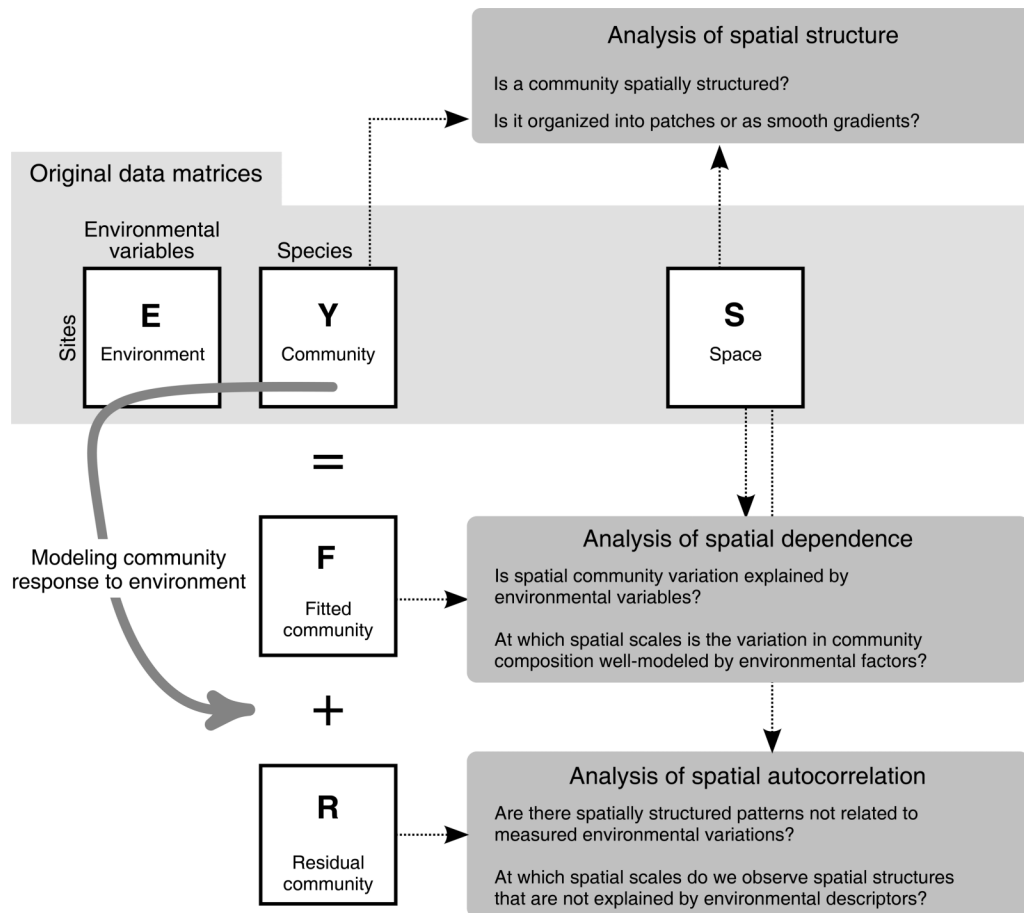
FIG. 1. Relationships between data, questions, and methods involved in the analysis of the spatial structures of ecological community data.

It is generally assumed that broad-scaled spatial structures in the species response data correspond to the scale (sometimes referred to as wavelength) of environmental drivers, whereas spatial autocorrelation generated by community dynamics is usually finer scaled. However, this classical dichotomy is probably too naïve (there are plenty of counterexamples of fine-scale environmental drivers and broad-scale population dynamics) and may arise simply because community ecologists usually measure and include broad-scaled environmental variables, ignoring or failing to measure those structured at finer scales. While it may appear oversimplistic to oppose niche and neutral influences on community structuring (Currie 2007, Smith and Lund-holm 2010), we suggest that identifying the characteristic spatial scales of variation displayed by the response variables that are explained (or not) by the measured environmental variables is a first step toward disentan-gling the various processes that might act to structure the spatial distributions of species (Smith and Lundholm 2010). To achieve this goal, the hypotheses about the processes one is positing as influential should be clearly stated in the form of competing models explaining

variation in community composition. As a consequence, the choice of sampling and analytical procedures should become more oriented toward comparing and contrast-ing relevant models and hypotheses. A causal model can then be built efficiently using a combination of several statistical tests corresponding to different alternative hypotheses regarding the determinants of (spatial) structures of ecological communities (Legendre and Legendre 2012: Section 13) (see Cottenie [2005] for an illustration).

After introducing some basic but critical notions, we review some recent developments in spatially explicit multivariate methods. Here we will restrict our review to ecological studies involving community data and envi-ronmental variables, measured in space, illustrating how these methods can help to illuminate the important factors structuring communities and their turnover or variation (beta diversity), using a tropical forest data set.

RELATING COMMUNITY STRUCTURES TO
ENVIRONMENTAL VARIABILITY

A traditional approach in community ecology consists in combining data from multiple species distributions

with environmental descriptors to obtain a summary of community structures and their relationships with environmental variability. From a methodological point of view, this community–environment modeling comprises two steps: (1) summarizing community structures and (2) relating them to environmental variability (the "assemble" and "predict" steps of Ferrier and Guisan [2006]), both of which can be combined in different ways. Usually, the first step requires the synthesis of complex information on a large number of species into a simpler form. The use of diversity indices, for example, is an extreme simplification that reduces multiple species information to a single synthetic value (e.g., species richness). Multivariate analysis (classification and ordination methods) represents a much richer alternative that takes better advantage of the multi-dimensional nature of ecological data. These two options share some similarities as ordination methods and some classical diversity indices are intimately linked (Pélissier et al. 2003). Hereinafter, we will refer to the multiple species distribution data as a site-by-species abundance matrix or community table, **Y**, and to the environmental data as a site-by-variables matrix or environment table, **E**, where both of the matrices **Y** and **E** have matching sites (Fig. 1). Ordination has long been a central approach to summarize the information contained in the community table, **Y**, by producing a small number of orthogonal gradients of compositional change, often referred to as "factors" or "ordination axes," along which species or sites may be ordered to study their relative positions (Gauch 1982, Legendre and Legendre 2012). It allows one to separate structural information associated with selected factors (axes) from random noise. Among the plethora of available methods (see Legendre and Legendre 2012:388) we consider only the commonly used family of eigenvector-based approaches (or eigenanalyses) that include principal component analysis (PCA) and correspondence analysis (CA) as special cases.

Historically, ecologists have first used indirect approaches for interpreting the structures of species assemblages (structural information extracted by the eigenanalysis of **Y**) in relation to environmental variability: site scores along the ordination axes, which are composite indices of species abundances contained in **Y**, were compared a posteriori to environmental variables ("indirect comparison," "indirect gradient analysis"). Progressively, new techniques were developed to constrain the ordination according to the table **E** of explanatory environmental variables ("direct comparison," "direct gradient analysis"; see Legendre and Legendre [2012: Section 10.2] for a synthesis). Technically, direct gradient analysis can be viewed as an extension of multiple regression, which has a single response variable, to the case of a multi-species response table: **Y** is then partitioned according to **E**, into a table of fitted values, $\mathbf{F} = f(\mathbf{E})$, and a table of residuals, $\mathbf{R} = \mathbf{Y} - \mathbf{F}$ (Fig. 1). Constrained ordination (or canonical analysis) concentrates on the eigenanalysis of the fitted community table, **F**, allowing the direct analysis of the variation in species abundances explained by the environmental variability. Standard approaches include canonical correspondence analysis (CCA) and redundancy analysis (RDA). Partial canonical analysis is an extension of these constrained ordination methods to the case where explanatory variables include covariables whose effects on the species response variables should be controlled. For instance, one can focus on the residual community table **R** to analyze the variation in species abundances unrelated to what has been modeled by the environment table **E** (partial PCA, or partial residual analysis [PRA]).

In the past, the application of constrained and unconstrained ordination methods has been mainly motivated by the desire to assess and describe ecological structures. Nowadays, the availability of fast computers and the development of permutation procedures permit the use of these methods for evaluating ecological hypotheses in an inferential framework (Manly 1997). More specifically, incorporating space as an explicit component in these analyses represents a further step toward the objective of uncovering unmeasured or unmeasurable ecological processes (McIntire and Fajardo 2009).

SPATIAL STRUCTURE, DEPENDENCE, AND AUTOCORRELATION

The spatial distribution of organisms may vary from aggregated (clustered) through a random pattern to the regular (uniform) case. Individuals are randomly positioned if the location of one individual is independent of the positions of all others. Aggregation occurs when individuals tend to be close together, resulting in a high variance of density estimates across space (over-dispersion). On the other hand, regularity occurs when individuals tend to avoid each other (e.g., spatial distributions of territorial organisms) so that density estimates are less variable (under-dispersion). According to the type of spatial distribution, appropriate sampling should be designed to ensure proper estimates of abundances and their distributions (Andrew and Mapstone 1987). It follows that the spatial structures manifest themselves by the relationship (or lack of independence) between abundance values observed at neighboring sites in space. In many instances, individuals are aggregated so that sampling sites that are closer together tend to display abundance values that are more similar than sites that are further apart, resulting in positive spatial dependence. Regular distribution of individuals may produce the opposite effect, inducing negative spatial dependence. Note that different ecological processes could act simultaneously, inducing both negative and positive dependencies in a data set (Dray 2011).

At the community level, spatial dependence may be induced by the functional dependence of the response variables (species) on some explanatory variables (e.g.,

environmental) that share some parts of their spatial structures (induced spatial dependence [Legendre and Legendre 2012, Fortin and Dale 2005]). This phenomenon has long been described in ecology as the environmental control model, which is the foundation of niche theory. If all important spatially structured explanatory variables are included in the analysis, the community model $Y = f(E) + R = F + R$, correctly accounts for the spatial structure and matrix R contains independent and identically distributed (i.i.d.) residuals. On the other hand, if the function is misspecified, for example through the omission of salient explanatory variables with spatial patterning such as a large-scale trend, or through inadequate functional representation, one may incorrectly interpret the spatial patterning of the residuals as autocorrelation.

Autocorrelation refers to a type of spatial dependence that may appear in species distributions as the result of population or community dynamics; it is also designated as true autocorrelation, inherent autocorrelation, or autogenic autocorrelation (Fortin and Dale 2005), or interaction model (meaning interaction among the sites) by Cliff and Ord (1981:141). From a statistical standpoint, autocorrelation is the spatial structure found in the error component of a community–environment model, once the effect of all important spatially structured environmental variables has been accounted for (i.e., included in the model). In practice, it is fairly difficult (if not impossible) to know whether all important environmental drivers have been included (and with correct functional forms) in the analysis of a particular data set. Hence, if autocorrelation remains in residuals, it could be driven by any number of processes including unmeasured (small-scale) environmental variables, or population/community dynamics. Nevertheless, theoretically, a complete model can be written as

$$Y = f(E) + R = F + R = F + T + U \text{ (i.e., } R = T + U)$$

where R is broken down into the spatial autocorrelation in the residuals, T, and a random error component, U. However, these two residual components are impossible to partition unless expectations from specific autocorrelated models (such as dispersal models for instance) can be obtained in the same metric as R, which no longer contains abundances, but signed deviations of the observed abundances from their fitted values as predicted by the environmental variables. As a consequence, the presence of a spatial structure in residuals, which makes the species residuals in R not i.i.d., has long been seen as a statistical nuisance for the environmental control model (Legendre 1993, Griffith and Peres-Neto 2006), often generating inflated type I error rates in the environmental model and spurious (apparent) species–environment concordance.

Indeed, apparent species–environment concordance can be generated when species distributions and environmental factors are both independently spatially structured. In this case, modeling tools (regression,

canonical analysis) will have a tendency to indicate significant habitat affinities even if measured environmental factors are not actually significant drivers of species occurrences or co-occurrences (Dormann et al. 2007, Peres-Neto and Legendre 2010). In cases of spatial dependence, the effective number of degrees of freedom in the sample is smaller than the one estimated from the number of observations; in other words, observations are not independent and thus cannot be freely permuted at random to create the reference (null) distribution of the test statistic. As a consequence, statistical tests (parametric or non-parametric, such as permutation tests) generate narrow confidence limits, making the test too liberal and generating inflated type I error levels (i.e., the null hypothesis is too easy to reject [Bivand 1980, Legendre et al. 2002]). Whether parameter estimates may be affected by spatial dependence depends on the model, the type of statistic and the degree and sign, positive or negative, of the spatial dependence. Theoretical and simulation work is still needed to clarify the situations in which parameter estimates are also affected (but see Peres-Neto and Legendre [2010] for some discussion).

An important point that is often overlooked is that spatial dependence could cause inferential problems even when only some variables within both response (Y) and predictor sets (E) (i.e., some species and some environmental variables) are spatially structured (Legendre et al. 2002). Therefore, the first analytical step should always be to test for the presence of spatial dependence in the residual table, R, in order to assess whether there are missing spatially structured covariates in the model. Univariate statistics such as Moran's I (1950) or multivariate alternatives (e.g., trace of the variogram matrix, see the section *Spatial ordination*) can be used to evaluate the significance of spatial patterns in residuals using different procedures, including Monte-Carlo permutation tests (Cliff and Ord 1981).

If a spatial structure is detected in residuals, several techniques are available to take it into account in the case of the univariate environmental control model (i.e., when analyzing a single response variable [Dormann et al. 2007, Beale et al. 2010]) and can be considered for testing summary statistics of community data as response variables, such as species richness or species scores along a single ordination axis (Diniz-Filho and Bini 2005). Rangel et al. (2007) chose another strategy by modeling species richness as sums of individual species responses predicted from univariate spatial models. There are however few alternatives for dealing with multiple species responses in multivariate autoregressive models (but see Jin et al. [2005]; M13 in Table 1). Alternatively, restricted permutations may also help by considering the appropriate number of "exchangeable units" (and consequently degrees of freedom) in multi-way sampling designs, particularly when sites are embedded within ecological classes (Anderson and ter Braak 2003, Couteron and Pélissier 2004, Heegaard and

TABLE 1. Description of various properties of methods for the analysis of spatial ecological data.

			Inputs	
Labels	Methods	Bibliographic references	Species–environment link	Treatment of species information
M1	Mantel test	Mantel (1967), Legendre and Troussellier (1988)	possible (if partial Mantel)	distance matrix
M2	autocorrelation measure of sites scores after simple ordination		no	ordination method
M3	autocorrelation measure of sites scores after canonical ordination		yes	ordination method
M4	multivariate spatial analysis based on Moran's I (MULTISPATI)	Dray et al. (2008)	no	ordination method
M5	partial canonical ordination	Borcard et al. (2004)	yes	ordination method
M6	variation partitioning	Borcard et al. (1992), Borcard and Legendre (1994)	yes	ordination method
M7	spatial filtering	Griffith (2000)	yes	raw data table
M8	multiscale ordination (MSO)	Wagner (2003, 2004), Couteron and Ollier (2005)	yes	ordination method
M9	multiscale pattern analysis (MSPA)	Jombart et al. (2009)	possible	raw data table
M10	linear model of coregionalization (LMC)	Bellier et al. (2007), Warckernagel (2003)	yes	raw data table
M11	constrained clustering	Gordon (1996)	possible	raw data table
M12	boundary detection	Jacquez et al. (2008)	possible	raw data table
M13	multivariate autoregressive model	Jin et al. (2005)	yes	raw data table

Notes: The ability to consider both species (**Y**) and environmental (**E**) information, the pre-treatment of the species abundance table, and the type of spatial component included in the analysis is detailed. The different outputs produced by the method (ability to describe the multiscale properties, tools to summarize the spatial patterns, and associated significance tests) are presented.

† See Table 2.

Vandvik 2004). Another approach is to filter out (remove) the effects of spatial dependence by detrending ("spatial filtering" sensu Griffith [2000]; M7 in Table 1). For instance, partial canonical analysis can be used to partial out the effect of spatial predictors (see *Specifying the spatial component*), so that tests for statistical significance of other sets of predictors of interest (e.g., environment) can be performed from a new community table corrected for spatial dependence (see Griffith and Peres-Neto 2006, Tiefelsdorf and Griffith 2007, Peres-Neto and Legendre 2010).

These different methods usually "correct" or account for the properties of the data to allow proper statistical inference in the presence of spatial dependence. In this context, the existence of spatial structures is considered as a "nuisance" that should be removed or at least taken into account (see Dormann et al. [2007] for a review of methods under this point of view). We want to highlight an alternative viewpoint focusing on the biological information contained in spatial patterns that are the signature of processes (e.g., environmental filtering, limited dispersal, historical biogeography) that shape species distributions. Hence, we suggest that the information contained in these spatial structures should be exploited rather than removed in order to improve the analysis of ecological structures. In this context, the spatial component should be introduced explicitly and considered as a proxy/surrogate of unmeasured process-es (McIntire and Fajardo 2009). Ecological inference can then be performed from the multiscale analysis of spatial structures in original data tables (**Y**), but also in each of the fitted (**F**) or residuals (**R**) tables from an environmental (or other) model.

SPECIFYING THE SPATIAL COMPONENT

Whatever the viewpoint adopted (nuisance or surrogate), there is a clear need to add a "spatialize" step to the usual "assemble" and "predict" steps of community–environment modeling. Spatial information can be introduced implicitly using mapping techniques in a two-step procedure: once community structures have been summarized by any constrained or unconstrained ordination technique, the sites scores along selected ordination axes are represented in geographical space as a way to identify spatial patterns. Goodall (1954) introduced this indirect approach using contour lines of PCA scores to map the main spatial structures of vegetation data and link them to environmental variability. Since this early work, there has been an increasing interest to consider both spatial and multivariate aspects simultaneously by introducing the spatial component (depicted as table **S** in Fig. 1) *explicitly* into community analysis (Table 2; see also Dray et al. 2006).

Note that an important issue to keep in mind is that in any field study, the sampling design imposes an artificial spatial structure on the data through decisions about the

TABLE 1. Extended.

Inputs	Outputs		
Type of spatial component†	Explicit multiscale detection	Description of the spatial structures	Significance tests of spatial structures
S1	no	a statistic	yes (but very low power)
S3 (but a posteriori)	no	a statistic, mapping	yes
S3 (but a posteriori)	no	a statistic, mapping	yes
S3	no	biplot, mapping	no
S2, S4, S5, S6, S7	yes (if submodels)	biplot, mapping	yes
S2, S4, S5, S6, S7	yes (if submodels)	Venn diagram	yes
S2, S4, S5, S6, S7	no	not relevant	not relevant
S3 (multiple)	yes	variogram	yes
S5	yes	biplot	no
S3 (multiple)	yes	mapping	no
S3	no	mapping of patches	no
S3	no	mapping of boundaries	yes
S3	no	autoregressive parameters	yes

extent of the study, the number of samples, and the size, shape, and spatial arrangement of the sampling units (e.g., Andrew and Mapstone 1987, Fortin and Dale 2005). Moreover, different alternatives that have been proposed to describe the pairwise relationships among sampling locations will affect the type of spatial information included or constraining the multivariate analysis.

Geographical distance matrix

In the past, ecologists used to represent spatial links (i.e., proximities) by functions of geographic coordinates to construct a matrix of inter-site geographic distances (S1 in Table 2; Legendre and Fortin 1989). Then, spatial structures were identified by approaches such as the Mantel (1967) test (M1 in Table 1), which has been recommended in the past to identify spatial structures by testing the correlation between two distance matrices, where one matrix represents the spatial configuration of sites and the other summarizes pairwise ecological similarities (e.g., Bray-Curtis index) among sample locations. While Mantel's test summarizes spatial structures by a global measure, the Mantel correlogram (Oden and Sokal 1986, Borcard and Legendre 2012) partitions the analysis into a series of distance classes that allows changes in the intensity of spatial patterns at different distances (scales) to be identified. Recently, this approach has gained much attention in the form of distance-decay similarity approaches to analyze how beta diversity varies across space (e.g., Soininen et al. 2007). Ferrier (2002) extended this distance-based approach and proposed generalized dissimilarity modeling (GDM), a nonlinear method to model ecological dissimilarities as a function of geographical and environmental distances. The significance of explanatory distance matrices (including space) can be tested by comparing the deviances of different models using permutation procedures (Ferrier et al. 2007). However,

TABLE 2. Summary of the different tools available to consider the spatial component and their ability to explicitly describe multiscale patterns.

| Labels	Methods	Family	Explicit multiscale decomposition
S1	geographical distance matrix	distance based on geographic coordinates	no
S2	polynomial	template based on geographic coordinates	broad scales
S3	spatial weighting matrix (SWM)	SWM	no
S4	principal coordinates of neighbor matrices (PCNM)	SWM-derived spatial template	broad and medium scales
S5	Moran's eigenvector maps (MEM)	SWM-derived spatial template	yes
S6	asymmetric eigenvector maps (AEM)	spatial template derived from directed graphs	yes
S7	Fourier, wavelets	predefined spatial template	yes
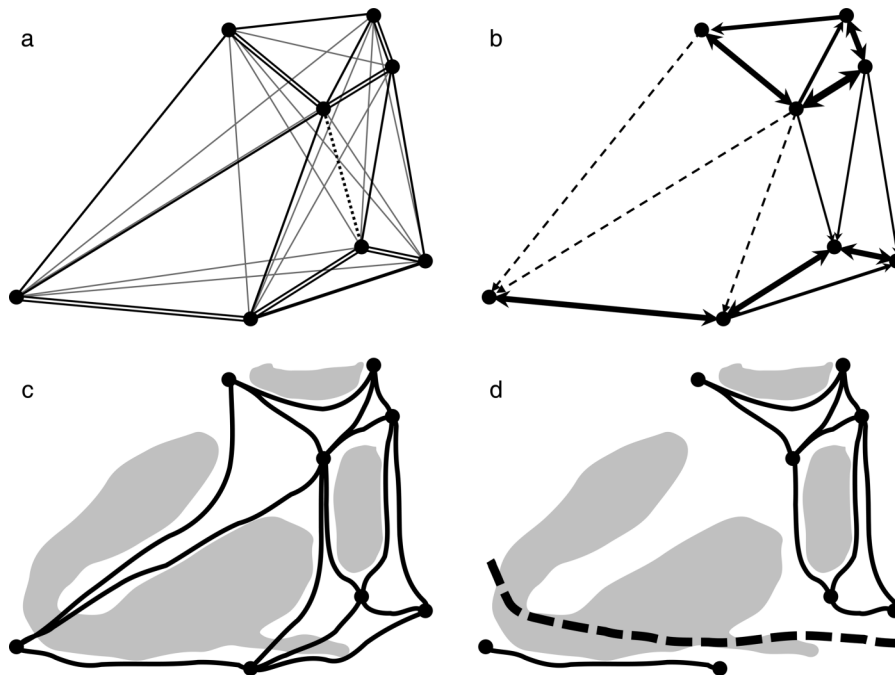
FIG. 2. Spatial link types. (a) Complete graph of direct links between sampling points (all lines, black and gray). The double lines are the nearest-neighbor links between points. The dotted line and the nearest-neighbor links create a minimum spanning tree. All black lines (simple, double, dotted) create a Delaunay triangulation graph. The links can also be seen as the Euclidean distances between sampling points. (b) Unidirectional and bidirectional links as indicated by the arrows; the links have different weights as indicated by the different thicknesses of the lines. (c) Least-cost links accounting for land cover types (gray patches) between sampling locations that impede species movement; animals avoid as much as possible crossing these patches. (d) As in panel (c), but with a barrier (dashed line) that prevents movement (i.e., links) between some sampling locations.

no explicit measure of the intensity of the spatial pattern is provided by this GDM approach. Legendre and Fortin (2010) have shown that the sum of squares partitioned in Mantel tests and regression on distance matrices differs from the sum of squares partitioned in linear correlation, regression, and canonical analysis. Specifically, the R^2 statistic from regression on distance matrices is unrelated to that which would be obtained from the use of canonical methods, and cannot be interpreted as an explained proportion of the variance of the response data. Moreover, Legendre et al. (2005) showed that the Mantel test has very low power in contrast to other spatial frameworks to detect spatial structures and thus should not be used for that purpose; the efficiency of related methods such as GDM has not been evaluated yet. Mantel-based approaches have also been conceptually criticized as a general approach to beta diversity analysis (Laliberté 2008, Legendre et al. 2008, Pélissier et al. 2008). Moreover, because they aggregate the original data in the computation of sites-by-sites distance matrices, these Mantel-based methods do not allow one to obtain information about particular species and environmental variables. Nonetheless, Mantel-based approaches remain used by ecologists probably because geographic distances are a very intuitive way to consider space, even though they are based on the assumption that ecological processes are strictly distance dependent and that this dependence is constant over the whole study area.

Spatial weighting matrix

Substantial improvement has stemmed from the concept of a spatial weighting matrix (SWM) used as a starting point to construct parsimonious representations of space (S3 in Table 2). In its broader sense, a SWM is usually a square symmetric matrix (sites-by-sites) that contains non-negative values expressing the strengths of the potential exchanges between the spatial units; conventionally, diagonal values are set to zero. In its simplest form, a SWM is a binary matrix, with ones for pairs of sites considered as neighbors and zeros otherwise. These binary connectivity matrices are directly related to matrix representations of graphs (Fig. 2a) constructed using distance criteria or tools derived from graph theory to depict connectivity (Fall et al. 2007, Dale and Fortin 2010); they may also describe spatial discontinuities or boundaries (Fig. 2d; Fortin and Dale 2005, Jacquez et al. 2008). Binary spatial links may appear too restrictive to represent complex inter-site relationships, however. SWMs that explicitly weigh the spatial relationships among sampling locations may be more appropriate in that case (Fig. 2b and c). For instance, the weights can be derived from functions of spatial distances (Dray et al. 2006), least-cost links

between sampling locations (Fig. 2c; Fall et al. 2007) or any other proxies/measures of the easiness of transmission of matter (e.g., organisms via dispersal), energy, nutrients, or information (e.g., social information among individuals) among sampling sites. In geography, the use of SWM has been popularized by Cliff and Ord (1973:12) who advocated its "great flexibility" and has become the standard way of describing spatial configurations of sampling sites.

Univariate measures of spatial correlation such as the Moran (1950) and Geary (1954) statistics evaluate the strength of spatial structures using a representation of spatial links in the form of a SWM. Different procedures, including Monte-Carlo permutations tests, can be used to test the significance of spatial structures using these indices (Cliff and Ord 1973). If spatial structures have been detected, they can be characterized by structure functions to study the effect of spatial scale on spatial correlation. Correlograms and (semi)-variograms plot spatial correlation indices (respectively, Moran's I and semi-variance, which is intimately linked to Geary's index) against distance classes among sites. Distance classes can be computed as distances along the links of a neighborhood graph corresponding to a SWM (instead of straight-line geographic distances). Several ecological studies have demonstrated the efficiency of variograms and correlograms to identify characteristics of spatial structures such as their periodicity, directionality, or patch size (Radeloff et al. 2000, Fortin and Dale 2005). In the context of community ecology, these tools can be useful for evaluating the strength of the spatial structures displayed along the ordination axes extracted from the analysis of the initial ($\mathbf{Y}$), predicted ($\mathbf{F}$), or residual ($\mathbf{R}$) tables (M2 and M3 in Table 1; Pyke et al. 2001, Baraloto and Couteron 2010).

Spatial templates

Another approach to specify the spatial component consists in building a spatial model from table $\mathbf{S}$ to produce spatial templates against which the structure of community tables can be analyzed. Perhaps the most common approach is the generation of trend surfaces from polynomial functions of the geographic coordinates (S2 in Table 2; Gittins 1985). More recently, Dray et al. (2006) showed that the diagonalization of a centered SWM is a generic way to generate "eigenvector maps," which are efficient representations of spatial relationships (Appendix). They coined the general name "Moran's eigenvector maps" (MEM; S5 in Table 2) for a concept that embraces as special cases distance-based eigenvector maps (for distance-defined SWMs), among which those provided by principal coordinates of neighbor matrices (PCNM; the pioneering technique of Borcard and Legendre 2002; S4 in Table 2), and Griffith's eigenfunctions (1996, 2000) based on a matrix of binary (or topological) links. The eigenvectors of any centered SWM are orthogonal and have a straightforward interpretation as spatial correlation templates that

can be ranked according to their Moran's (1950) index (Appendix).

Since symmetric relationships among sampling locations are not always adequate to reflect processes that are directional or asymmetric, Blanchet et al. (2008a) developed asymmetric eigenvector maps (AEM; S6 in Table 2), a form of spatial eigenfunctions that allows one to model directional spatial patterns such as the ones along river networks or oceanic currents (see also Mahecha and Schmidtlein [2008] for an approach using anisotropic spatial filters; and Salomon et al. [2010] for how asymmetric dispersal can impose patterns in species coexistence). However, AEMs cannot be reduced to the MEM framework.

SWM and its associated eigendecomposition (MEM) provide very efficient and flexible ways of specifying the spatial component in an analysis. In the remainder of this paper, we focus on several recently developed methods that explicitly introduce an SWM- or MEM-based spatial constraint into multivariate community analysis. Griffith and Peres-Neto (2006) proposed the general expression of eigenfunction-based spatial analysis, or spatial eigenfunction analysis, for the latter approach. Spatially constrained ordination and clustering methods (see *Spatially constrained ordination and clustering*) usually introduce the spatial component by extending univariate autocorrelation measures (SWM-based) or by using spatial eigenfunctions (MEM-based) as predictors/covariables in a modeling framework of species abundances. MEM can be also used as an orthonormal basis (Ollier et al. 2006) on which the species variances are decomposed to obtain a signature of their spatial distributions, allowing the study of community compositional variation at multiple scales.

SPATIALLY CONSTRAINED ORDINATION AND CLUSTERING

Spatial ordination

Informally speaking, spatial ordination embraces any ordination procedure that explicitly integrates a spatial component $\mathbf{S}$ (Fig. 1). A way to account for space in ordination methods consists in considering $\mathbf{S}$ as a set of spatial predictors (or spatial templates) representing a multiscale decomposition of space on which are regressed the tables $\mathbf{Y}$, $\mathbf{F}$, or $\mathbf{R}$, depending on the ecological hypotheses that are being tested. This type of analysis was initiated with constrained ordination (either CCA or RDA) using an explanatory table of spatial predictors, such as a polynomial trend-surface function of the geographic coordinates of sites (Gittins 1985, Borcard et al. 1992). In canonical trend surface analysis, site scores optimize the spatial component of the variation of community composition and they can be mapped to depict the spatial patterns of beta diversity. Correlations can be computed between these sites scores and the environmental variables to evaluate if the main spatial structures are related or not to environmental variability. Then, significant environmental variables can be introduced as covariables in a partial canonical

analysis to control for their effect and focus on unexplained spatial patterns that could reflect other ecological processes such as biotic relationships (Borcard and Legendre 1994, Pélissier et al. 2002, Gimaret-Carpentier et al. 2003). Several limitations to the use of polynomial functions have been reported in the literature such as their ability to only account for smooth broad-scale spatial patterns, or the collinearity found among the spatial predictors unless orthogonal polynomials are used (Dray et al. 2006).

The MEM framework solves most of the problems posed by polynomials and offers a powerful alternative to construct spatial predictors that can be used in spatial ordination. However, the large number (number of sites − 1) of spatial predictors, called eigenfunctions, produced by the MEM approach renders the multiple regression stage of canonical methods unstable and meaningless. To alleviate this problem, one can divide the MEM eigenfunctions into two subsets displaying positive and negative spatial correlation, corresponding to those with positive and negative eigenvalues, respectively, and analyze the subset that is relevant for the study (broad or fine spatial scales). As a complement, one can use forward selection to select a subset of spatial predictors to be incorporated in the statistical model (see Blanchet et al. 2008b); in that case, only a part of the spatial information in the SWM is retained in the analysis as significant eigenfunctions. Selection of MEM eigenfunctions is a crucial issue and Bini et al. (2009) showed that the use of different criteria could strongly influence the interpretation of environmental effects in the case of a univariate response variable. Detailed discussions on the selection of spatial predictors can be found in Peres-Neto and Legendre (2010) in the case of statistical inference in the presence of spatial dependence; further work is required to confirm these results when MEM are used as a proxy for unmeasured processes.

Several methods that incorporate the spatial constraint directly in the form of the SWM have been proposed. Compared to methods based on spatial explanatory variables, they have the advantage of avoiding a preliminary step of selection of predictors. These multivariate procedures are based on the diagonalization of a matrix describing the spatial relationships between the variables. The construction of that matrix is similar to the computation of a variance–covariance matrix except that the diagonal elements are univariate statistics of spatial autocorrelation instead of variances while the off-diagonal elements contain bivariate autocorrelation statistics instead of covariances. This leads to the variogram matrix (sensu Wagner 2003) or to the spatial correlation matrix (Wartenberg 1985) depending on whether the Geary (1954) or Moran (1950) indices were used to estimate spatial autocorrelation. In the ecological literature, Thioulouse et al. (1995) proposed a general framework that reconciles these two approaches, while Dray et al. (2002) presented a Geary-based

method to link data sets from two different spatial samples in the same geographic area (e.g., if environmental descriptors and species have been sampled in the same region but not exactly at the same sites). To date, the most integrated, yet flexible example of spatial ordination is probably the MULTISPATI method (multivariate spatial analysis based on Moran's I; Dray et al. [2008]; see also Jombart et al. [2008]; M4 in Table 1), which generalizes Wartenberg's (1985) analysis to CA and potentially to any ordination method. The application to vegetation data clearly showed that the method is able to reveal spatial patterns of floristic composition, which were not apparent in the mapping of CA scores alone (Dray et al. 2008). Contrary to methods based on spatial predictors, this approach preserves all the spatial information contained in the SWM, but unfortunately it does not yet allow the consideration of a table of environmental explanatory variables.

A variety of options have been proposed that relate to the definition of the SWM, the form of the spatial constraints (SWM or its eigenfunctions) and the choice of a particular ordination method (e.g., PCA, CA); the latter implying different ways to quantify beta diversity. The construction of the SWM is recognized as a critical step in these analyses. Spatial connections between sites may be postulated a priori, using pre-existing knowledge about the question one wants to investigate through spatial ordination (e.g., dispersal routes and rates).

Some spatial ordination techniques can also be used to measure and test the importance of spatial structures. Constrained ordination using an explanatory table of spatial predictors (e.g., MEM) provides a way to evaluate and test the significance of spatial structures. The intensity of spatial patterns can be estimated as the part of variation in community data that is explained by spatial predictors and tested by permutation procedures (Peres-Neto et al. 2006). Environmental descriptors can also be introduced in this framework, leading to the variation partitioning in canonical analysis pioneered by Borcard et al. (1992) and Borcard and Legendre (1994) and modified by Peres-Neto et al. (2006) to estimate and test the relative importance of spatial structures, environmental variables or both on the variation in species compositions (M6 in Table 1). Some spatial multivariate statistics have also been defined from the variogram matrix or the spatial correlation matrix for significance testing. For instance, Wagner (2003) proposed the trace and the total sum of elements of the variogram matrix to characterize multivariate spatial patterns. These quantities can be tested by permutation procedures. Deriving statistics from the spatial correlation matrix directly is much harder, given that it can contain positive and negative values corresponding to positive and negative spatial correlation. Hence, statistics based on the sum of elements of that matrix have no meaning (i.e., positive and negative autocorrelated patterns cancel each other out) and cannot be used. Alternative statistics to test either positive or negative
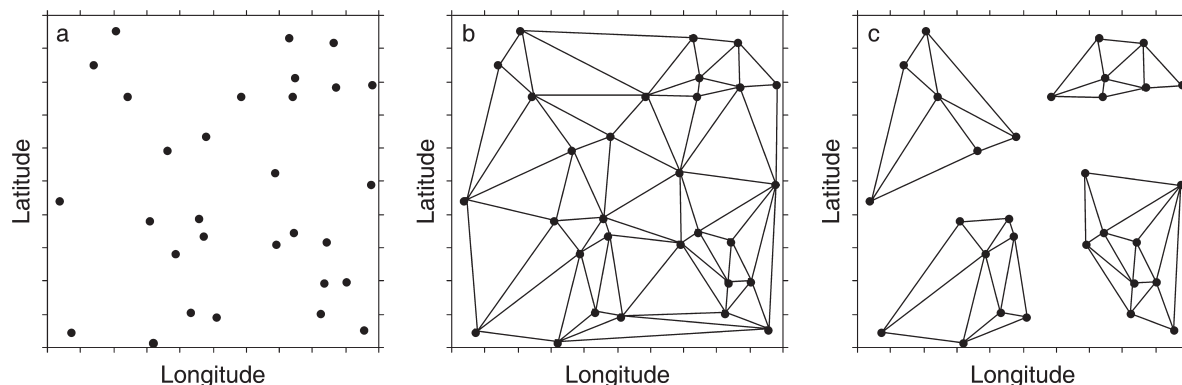
FIG. 3.   Principle of clustering with spatial contiguity constraint. (a) Draw the sites on a map. (b) Link the sites by a connection network, here a Delaunay triangulation, which can be represented by an adjacency matrix. (c) Cluster the (univariate or multivariate) community composition response data by agglomerative clustering of *K*-means partitioning, using the link edges as constraints: only adjacent sites or groups of sites can be clustered by the algorithm operating on the response data.

multivariate autocorrelation have been proposed using multiple regression on two sets of MEMs (Jombart et al. 2008).

Spatial clustering

Exploratory analysis using spatial ordination techniques may reveal that the data could be clustered across space into groups with either narrow (sharp) or wide boundaries between them (Fortin and Drapeau 1995). A first but indirect way of obtaining a spatial partition is to apply a clustering method on the site scores produced by a spatial ordination method. A more elegant and efficient way to delimit spatially homogeneous groups is to incorporate spatial constraints in the clustering method (Legendre and Fortin 1989, Gordon 1996; M11 in Table 1). Spatial adjacency or contiguity of sampling sites can be represented by a binary SWM, which is then used to restrict the clustering algorithm to only cluster sampling sites that are contiguous in space, thus creating patches (Legendre and Fortin 1989). Although two or more patches of similar species composition would form a single entity in an unrestricted clustering, the spatial restriction of contiguity forces them to remain separated. Examples of application of spatially constrained clustering to floristic variation are found in Fortin and Drapeau (1995) and Tuomisto et al. (2003). Likewise, *K*-means partitioning algorithms may be constrained by a contiguity matrix (Legendre and Fortin 1989). Similarly, multivariate classification and regression trees can produce spatially constrained clusters if the geographic coordinates of sampling sites are used as the constraints in the analysis (Bachraty et al. 2009). By determining spatially homogeneous patches, spatial clustering delineates boundaries between adjacent spatial clusters (Fig. 3; Fortin and Drapeau 1995).

Sometimes the spatial structure is neither patchy nor in the form of gradients, but a combination of the two. In such cases, instead of trying to delineate patches, one may look for transition zones showing abrupt changes in

community composition, or boundaries. Boundary detection methods (e.g., edge detection, segmentation, Wombling algorithms; see Fortin and Dale [2005] for a review; M12 in Table 1) do not produce patches like constrained clustering techniques but patches can arise as a combination of different boundary types (width, shape), intensities or scales (Csillag et al. 2001, Philibert et al. 2008). Hence, boundary detection may be regarded as a more flexible approach than constrained clustering. The pioneering work of Womble (1951) on genetic boundary techniques to detect zones of rapid spatial change provided the conceptual framework to develop a series of boundary detection methods for quantitative or qualitative observational data (Oden et al. 1993, Fortin 1994, Jacquez et al. 2008) with emphasis on ecological data. Alternatively, Monmonier's algorithm (Monmonier 1973) detects boundaries by finding the path exhibiting the largest differences (found in a distance matrix) between neighboring objects.

After deriving spatial patches either by spatial clustering or boundary detection, one can investigate if these patches are the result of self-organization (ecological processes within and among species), environmental influence, or both. To approach this question, patches in community structure may be compared to patches derived from environmental variables (Jacquez et al. 2008).

DECOMPOSING SPATIAL PATTERNS OF SPECIES ASSEMBLAGES AT MULTIPLE SCALES

Scale is generally defined on the basis of the main features of the sampling design such as the extent of the study area or the size and spacing of the sampling units (Wiens 1989). MEMs are orthogonal maps that provide a decomposition of the spatial relationships among the sampling sites based on a given SWM. Hence, spatial eigenfunctions reflect the spatial distribution of sites and are usually interpreted in terms of separate scales as a spectral decomposition (Appendix; Borcard and Legen-

dre 2002). Results about the most influential eigenfunctions in the analyses presented in the previous section thus straightforwardly translate into scale analyses. Another point of view relates to the distance beyond which observations appear as fairly independent: this distance can be estimated by the range of the semivariogram (Bellier et al. 2007). From these premises, different methods have been proposed to study the multiscale characteristics of community data. Fully worked examples of several of the methods described in the following section are presented in Chapter 7 of Borcard et al. (2011).

Spatial predictors in submodels

One can identify the MEMs that contribute significantly to the explanation of the species response data using canonical ordination methods. These MEMs can be assembled into a small number of submodels, e.g., broad scale, intermediate scale, and fine scale. The predicted values, generated for each submodel, can be then reanalyzed by canonical analysis against environmental variables in order to identify the environmental variables linked to the species distributions at the scale represented by each submodel (Borcard et al. 2004; M5 in Table 1).

An alternative is multivariate variation partitioning (Borcard et al. 1992), which allows researchers to partition the variation of a species response data table among two or several explanatory tables. These data tables may contain environmental variables as well as spatial predictors (all MEMs, or those corresponding to one of the submodels, e.g., the broad-scale variation, or to several of them). This modeling allows one to determine how much of the species variation is spatially structured, and within that, how much variation can be related to the influence of the measured environmental variables (see Legendre et al. 2009, Peres-Neto and Legendre 2010).

Analysis of a scalogram

One can perform a complete and additive decomposition of the variability of a single response variable or a multivariate data table onto the MEMs basis. This principle is similar to the spectral decomposition based on Fourier transforms using sine and cosine functions (S7 in Table 2; Renshaw and Ford 1984, Munoz et al. 2007) or on wavelet orthogonal bases (Keitt and Fischer 2006). However a major point is that unlike MEMs, these methods are restricted to regular sampling designs (grid). A scalogram can be constructed showing how well each MEM eigenfunction explains the variability of the response data (Legendre and Borcard 2006). Different statistics can be computed and tested using permutation to summarize some properties of the scalogram and thus identify the main scales of variation (see Ollier et al. [2006] in a phylogenetic context). Jombart et al. (2009) computed scalograms for a set of individual species and assembled them into a table that

was then subjected to a particular PCA to produce a summary of the multiscale covariation and identify the main scales at which the species distributions were structured (multiscale pattern analysis, MSPA; M9 in Table 1). By analogy with Fourier analysis, Munoz (2009) suggested that a smoothing procedure (Munoz et al. 2007) helped to improve process inference and allowed one to grasp independent signatures of habitat structure (environmental control) and metapopulation dynamics (an aspect of biotic control).

Empirical variography

Multiscale ordination (MSO; M8 in Table 1) is another way to reveal and analyze the multiscale structure of the spatial distributions of organisms. It consists of computing a series of variance-covariance matrices corresponding to different scales (empirical variogram matrix [Wagner 2003]). The multiple scales are represented by multiple aggregations of the original SWM into higher-order SWMs (Couteron and Ollier 2005). Wagner (2003, 2004) concentrated on distance-based SWMs and offered a wide range of tools to interpret and analyze variogram matrices. As a generalization of the empirical (semi)-variogram, a multivariate variogram depicts how the total variance of the community data changes as a function of distance (Wagner 2004). The method has been generalized to a wide range of ordination methods; this allows one to directly interpret variogram values as portions of beta diversity (Couteron and Ollier 2005). This generalization also includes canonical ordination methods, allowing one to compare the spatial structures embodied by the initial ($\mathbf{Y}$), predicted ($\mathbf{F}$), and residual ($\mathbf{R}$) tables.

Multivariate geostatistics

Variography can also be done by fitting a nested variogram model to the observed experimental variogram. This model considers an observed phenomenon as the sum of several independent subphenomena acting at different characteristic scales (Wackernagel 2003, Bellier et al. 2007). The resulting model is a weighted sum of elementary variogram models with different parameters (range, sill). Different fitting procedures can be used such as least-squares, weighted least-squares, or maximum likelihood (Cressie 1993). Model selection procedures such as the Akaike Information Criterion can help to choose among nested variogram models (Webster and McBratney 1989). The scale components of a nested variogram model, which directly relate to the ranges of the individual variogram models, can then be extracted and mapped by filter kriging, which explicitly decomposes the hierarchical layering of the spatial structures identified in the data (Wackernagel 2003). Spatial components extracted by filter kriging are conceptually analogous to spatial eigenfunctions because filter kriging techniques bear some analogy with spectral analysis methods (Wackernagel 2003). The nested variogram approach can also be used to model variograms and

cross-variograms simultaneously through a linear model of coregionalization (LMC [Wackernagel 2003]; M10 in Table 1). LMC models the variance–covariance matrix of several spatially structured variables at multiple scales, hence depicting how relationships between species and environment change across scales (Bellier et al. 2007).

Bellier et al. (2007) compared PCNM (a distance-based form of MEM) and nested variogram models and obtained comparable results for identifying relevant scales of species distributions, though patterns at finer scales identified by the PCNM submodels appeared smoother than those obtained by filter kriging. They suggested that this could be due to the fact that nested variogram models combine the sampling scheme and the data in one step, whereas the PCNM method first uses the sampling scheme to construct eigenfunctions and then selects the eigenfunctions that are the most highly related to the response data. In MEM-based methods, the definition of submodels using a potentially large number of synthetic spatial predictors is an arbitrary step and methodological developments are still required to refine an appropriate and rigorous approach for submodel selection. On the other hand, nested vario-gram modeling requires the fitting of numerous vario-gram and cross-variogram parameters and makes strong assumptions about stationarity for both the response and predictor variables (i.e., the processes that structure the environment). In multiscale ordination and scalo-grams, there is no estimation of additional parameters and the whole spatial information is considered, avoiding the subjective selection of spatial predictors.

WORKED EXAMPLE: FLORISTIC COMPOSITION ALONG THE PANAMA CANAL

Condit et al. (2002) used a model of distance decay in similarity derived from the spatially explicit neutral model of Chave and Leigh (2002) to illustrate the fact that the floristic dissimilarity between forests plots along the Panama Canal (PCW data from Pyke et al. 2001) decreased with distance as a result of limited dispersal and habitat effects. They showed that the observed pattern was as predicted by the neutral model in the intermediate range of distances (0.2–50 km). This, however, cannot fully explain the steep decline of similarity observed at shorter distances (<0.1 km). They further advocated the role of local habitat heterogeneity (canopy light gaps) to explain the departure from neutrality at the local scale. Their approach consisted in examining the discrepancy between the observed dissimilarities and those expected from a theoretical model of similarity decay fitted to the observed similarity vs. distance function, based on the approxi-mation of a general dispersal parameter common to all species. Hypotheses about the niche processes acting at local scales were suggested a posteriori to explain departures from the fitted model.

We used here a complementary approach, initially applied by Pyke et al. (2001) to study the PCW data set, which consists in examining how the spatial pattern of beta diversity changes when considering the initial species abundance table ($\mathbf{Y}$), its approximation by environmental variables ($\mathbf{F}$), and its residual counterpart when environmental variables are factored out ($\mathbf{R}$) as proposed by McIntire and Fajardo (2009). While Pyke et al. (2001) used a posteriori semi-variogram modeling of ordination scores, instead, here, we used the MEM framework to estimate and test the multiscale compo-nents of spatial patterns in $\mathbf{Y}$, $\mathbf{F}$, and $\mathbf{R}$.

We considered an initial floristic table $\mathbf{Y}$ containing the abundances of 778 species in 50 forest plots (supplemental Table 1 of Condit et al. [2002], from which we omitted the 50 1-ha subplots from Barro Colorado Island) and a table $\mathbf{E}$ containing four explanatory environmental variables (annual precipita-tion, elevation, age, and geology given in Appendix B of Chave et al. [2004]). We applied a chi-square transfor-mation (Legendre and Gallagher 2001) on table $\mathbf{Y}$ and used a principal component analysis (PCA) to identify the main patterns in community data. We used the chi-square transformation to put emphasis on rare species, which represent the majority of the 778 species in this data set (526 species have less than 10 individuals and 581 occur in five sites or fewer). Data and R scripts are available in the Supplement and allow one to reproduce the different analyses presented here. We also provide the code to perform the analyses using the Hellinger transformation that gives more weight to abundant species, thus producing contrasting results.

Redundancy analysis (RDA) was used to identify the main structures explained by the measured environmen-tal variables (analysis of $\mathbf{F}$) while partial residual analysis (PRA) allowed us to remove the effects of measured environmental variation (analysis of $\mathbf{R}$). The spatial component $\mathbf{S}$ was considered using MEMs based on a Gabriel graph (Legendre and Legendre 2012:836–837). For each table (i.e., $\mathbf{Y}$, $\mathbf{F}$, and $\mathbf{R}$), scalograms were computed by projecting the sites scores on the first two axes of the different analyses (PCA, RDA, and PRA, respectively) onto the spatial basis formed by the 49 MEMs, which produced a partitioning of the respective variances according to spatial scales ranked from the broadest to the finest. In order to avoid aliasing effects (i.e., undesired sampling artefacts at fine scales; see Platt and Denman [1975]), scalograms are presented in a smoothed version with seven spatial components formed by groups of seven successive MEMs (Munoz 2009). In the absence of spatial structure, the individual R^2 values (measuring the amount of variation explained by a given scale) that form a scalogram are expected to be uniformly distributed (Ollier et al. 2006). We used a permutation procedure (with 999 repetitions) to test if the maximum observed R^2 (R2Max, corresponding to the smoothed MEM at which the ecological pattern is mainly structured) is significantly larger than values
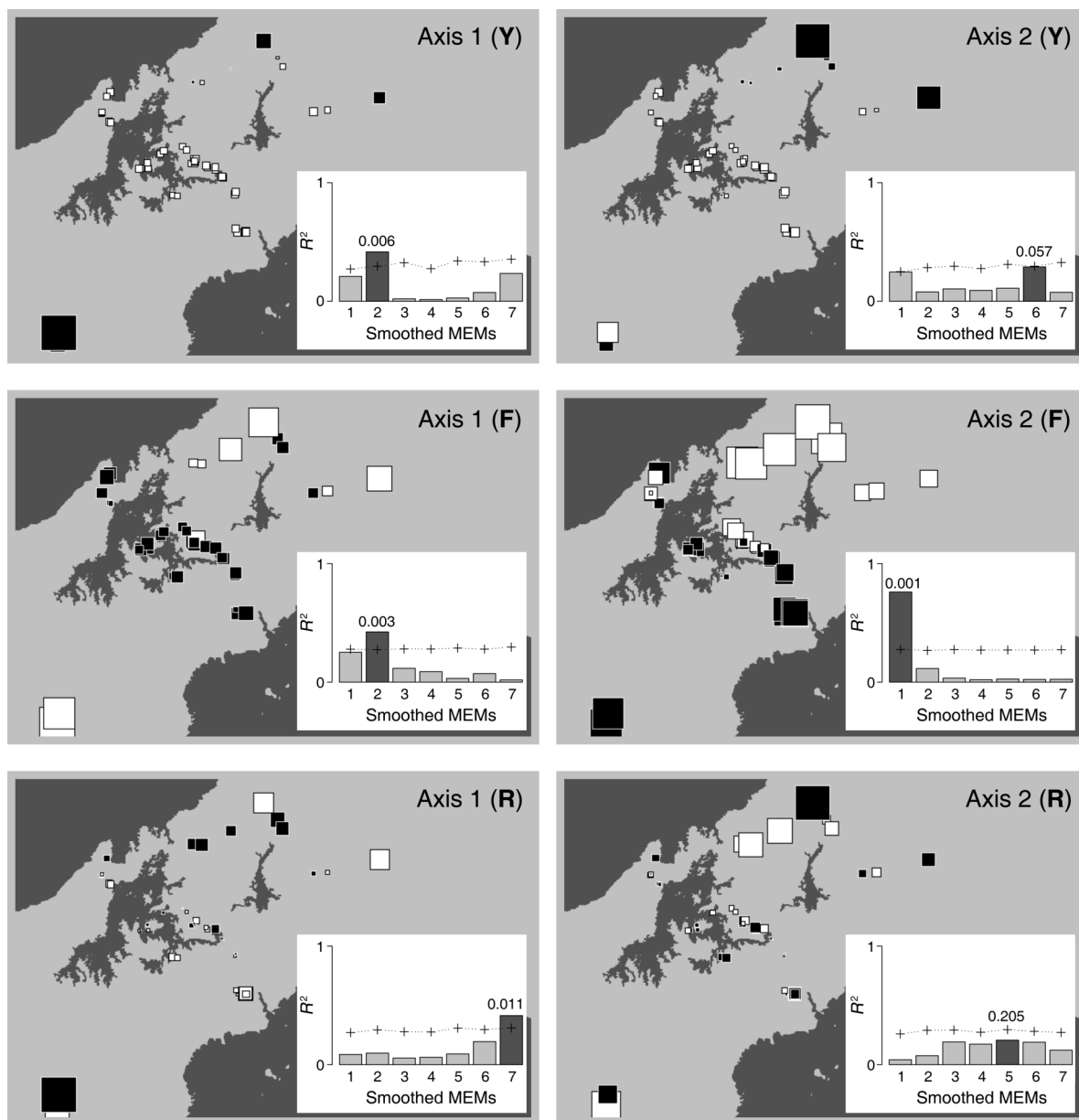
FIG. 4. Maps along the Panama Canal of the site scores on the first and second axes of the analysis of the original table (principal component analysis of **Y**), the approximated table **F** (redundancy analysis with **E** as predictors) and the residual table **R** (partial principal component analysis with **E** as covariables). For each score, a smoothed scalogram (the 49 Moran's eigenvector maps [MEMs] are assembled in seven groups) indicates the portion of variance (R^2) explained by each spatial scale. For each scalogram, the scale corresponding to the highest R^2 (in dark gray) is tested using 999 permutations of the observed values (P values are given). The 95% confidence limit is also represented by the line of plus signs.

obtained in the absence of a spatial pattern. This worked example is, to our knowledge, the first application of this method, which was originally developed for phylogenetic comparative studies (Ollier et al. 2006), in a spatial context.

The environmental variables explained a significant proportion of the variation of the initial floristic table, **Y** ($R^2 = 0.31$, $P = 0.041$ based on 999 permutations). The estimated table **F** exhibited two prominent axes,

representing, respectively, 28.3% and 21.8% of the total variance in **F**, and correlating mainly with elevation ($r = -0.94$ and -0.26, respectively) and rainfall ($r = -0.35$ and -0.80, respectively). Fig. 4 shows maps and associated scalograms of the main ordination axes. The scalograms for the first two axes of the initial floristic table, **Y**, have similar shapes with variance accumulation in both broad- and fine-scale components (Fig. 4, top). The first axis exhibited a broad-scale

nonrandom spatial pattern (R2Max = 0.42, P = 0.006) while the second axis had an important but nonsignificant fine-scale component (R2Max = 0.29, P = 0.057). The first two axes of table **F** (R2Max = 0.42, P = 0.003 and R2Max = 0.76, P = 0.001 for axis 1 and 2, respectively) showed significantly skewed distributions of the spatial variance toward the broad-scale components (Fig. 4, middle). This result is not surprising given that all available environmental variables included in **X** varied essentially at large spatial scales. On the other hand, the first two axes of the residual table **R** (Fig. 4, bottom) showed a significant accumulation of the spatial variance in the fine-scale components for the first axis (R2Max = 0.41, P = 0.011) and a nonsignificant structure at fine and medium scales for axis 2 (R2Max = 0.20, P = 0.205). This means that a significant fine-scaled spatial pattern remained in the data after the large-scale effects attributable to the measured environmental gradients (mainly a combination of elevation and rainfall) were partialed out.

Finally, we also performed variation partitioning of **Y** (Borcard et al. 1992) considering the environmental, broad-scale, and fine-scale components. Following Blanchet et al. (2008*b*), a forward selection procedure was applied to the MEM spatial predictors (those associated with nonsignificant Moran's indices were removed a priori), and 13 MEMs explaining 42.9% of the total variation of **Y** were selected. These MEMs were then divided in two groups corresponding to broad scales (9 MEMs associated with a positive Moran's statistic) and fine scales (4 MEMs associated with a negative autocorrelation). Variation partitioning (Fig. 5) identified a significant pure broad-scale spatial fraction (adjusted R^2 = 0.11, P = 0.005), a significant pure fine-scale spatial fraction (adjusted R^2 = 0.10, P = 0.005), a slightly significant pure environmental fraction (adjusted R^2 = 0.06, P = 0.049) and a fraction corresponding to broad-scale structured environment (adjusted R^2 = 0.06). These results indicate prominent effects of large-scale environmental drivers on the spatial structure of communities. The detection of additional significant broad-scale spatial patterns in residuals suggests that there are other important large-scale drivers (be they environmental, historical or biotic) in this system. Fine-scale structures could be generated by dispersal processes (Condit et al. 2002, although probably limited by the omission of the 50 1-ha plots from Barro Colorado Island), biotic interactions, or micro-site effects of unmeasured environmental factors.

Concluding Remarks

Beyond the standard nuisance viewpoint, an alternative and more promising perspective is that describing spatial structures in data can help us challenge our models and improve our understanding of species and community distributions (Legendre 1993). Spatially structured residuals usually indicate either that the model may be misspecified in the sense that important
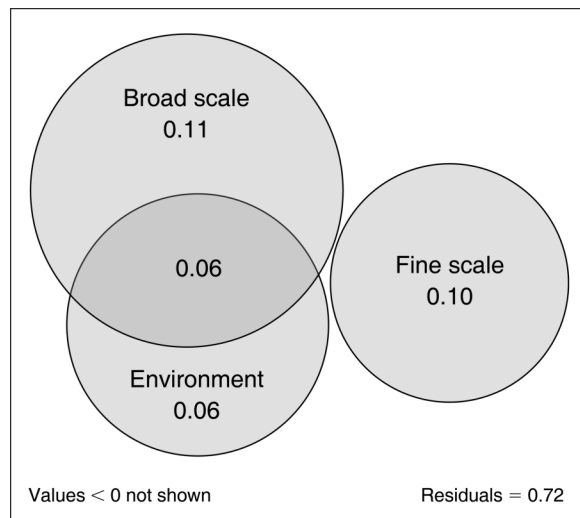

Fig. 5. Variation partitioning results of the Panama Canal data among an environmental component (lower left circle), a broad-scale spatial component (upper left circle), and a fine-scale MEM spatial component (right circle). The empty fractions in the plot have small negative adjusted R^2 values.

predictors may be missing from the model or that other processes are important besides the effects of the measured environmental factors (Griffith 1992, Fortin and Dale 2005, Wagner and Fortin 2005, McIntire and Fajardo 2009). Habitat models relating habitat characteristics to community structure are expected to answer at least two questions: (1) How well is the distribution of a set of species explained by a set of predictor variables? and (2) Which predictors are irrelevant or redundant in the sense of failing to strengthen the explanation of patterns after other predictors have been taken into account? The first question relates to the predictive power of the model that can be used, for example, in conservation management, for questions such as estimating habitat suitability, forecasting the effects of habitat change due to human interference, establishing potential locations for species reintroduction, or predicting how community structure may be affected by the invasion of exotic species. The second question is important for heuristic development such as determining the likelihood of competing hypotheses to explain particular patterns in community structure (Peres-Neto 2004). Given that a number of ecological processes are spatially structured, one way to account for some of the unrecorded or unavailable information is to use spatial predictors in our models as proxies for missing, but otherwise important predictors, e.g., missing environmental variables, biotic interactions, dispersal, or even long-lasting historical effects such as biogeographical events (Leibold et al. 2010). The methods presented in this review and summarized in Table 1 offer different alternatives to explicitly introduce space into community analysis; an R package

implementing them is under development (*available online*).[17] We consider that these tools will help ecologists implement appropriate methods to model their data in a more precise spatially explicit framework for generating and testing specific relevant hypotheses regarding patterns and potential processes.

Another point that is rarely considered is that omitting relevant explanatory variables or representing them by an incorrect functional form leads to model misspecification, which can translate into spatial non-stationarity (Fotheringham et al. 2002). Within this framework, the use of local statistics (Anselin 1995) can be relevant to better understand the nature of the misspecification and to identify the spatial structure of variables that have been omitted from the model and should be introduced to improve its accuracy. It is clear that this perspective has been underexplored and its development could be useful in ecological studies. Recently, this approach has gained attention by the application of geographically weighted regression (Fotheringham et al. 2002), which allows the study of the spatial non-stationarity of coefficients estimates (i.e., local changes in the relationships between the response and explanatory variables in linear univariate models). However, this method has been questioned in the spatial statistics literature (e.g., Finley 2011) and it seems difficult to extend it to the case of multiple species responses. A promising alternative might be to capture this non-stationarity by introducing interaction terms between environmental variables and spatial eigenfunctions in species–environment models (Griffith 2008).

As specific spatiotemporal signatures are expected from population and community dynamics, an important future objective should be to express predictions of mechanistic models, either analytic or simulation-based, in a way that can be directly investigated by spatially explicit multivariate methods. Observed data could then be used for model testing, as in the worked example, and inference about model parameters (Beeravolu et al. 2009). Using such an approach, Munoz et al. (2007) applied Fourier analysis to gridded species occurrence maps by reference to a classical metapopulation model and uncovered distinguishable signatures of population dynamics and habitat structuring. Condit et al. (2002) considered the predictions of a neutral model regarding the decay of compositional similarity with distance and concluded that departures from expectations are probably explained by environmental determinants. Note that the focal model in Condit et al. (2002) was the neutral random dispersal of species with distance, rather than a model of species responses to environmental predictors. Yet, there is increasing evidence that compositional turnover, i.e., beta diversity, is shaped by interactions between geographic connections and environmental heterogeneity. An exciting challenge is

therefore to use the diversity of spatialized multivariate techniques to identify the various modalities of such interactions. Although recent studies based on artificial data generated by mechanistic models seem to indicate that variation partitioning is inefficient to distinguish signatures of neutral and niche processes (Smith and Lundholm 2010), we are convinced that a thorough analysis of the multiscale components of multivariate spatial structure, spatial dependence and spatial correlation through the partitioning of the community data (**Y**) into tables of fitted (**F**) and residuals values (**R**), will allow researchers to identify the important potential environmental factors that influence the variation of beta diversity. It could help to design further empirical studies by identifying unexplained spatial patterns that could be due to the omission of important environmental variables or to other types of ecological processes. Additional information on species (e.g., phylogeny or traits) could be integrated in this analytical framework to explore new questions and refine ecological hypotheses (Cavender-Bares et al. 2004, Ives and Helmus 2010, Leibold et al. 2010, Peres-Neto et al. 2012). In the future, our understanding of ecological processes and associated community structures would benefit from a general framework integrating the development of ecological theories, mechanistic models and statistical methods. This would help to define testable hypotheses concerning the observed patterns and would allow for a rigorous evaluation of statistical methods that could be used in empirical case studies.

Literature Cited

Agrawal, A. A., et al. 2007. Filling key gaps in population and community ecology. Frontiers in Ecology and the Environment 5:145–152.

Anderson, M. J., and C. J. F. ter Braak. 2003. Permutation tests for multi-factorial analysis of variance. Journal of Statistical Computation and Simulation 73:85–113.

Andrew, N. L., and B. D. Mapstone. 1987. Sampling and the description of spatial pattern in marine ecology. Oceanography and Marine Biology Annual Review 25:39–90.

Anselin, L. 1995. Local indicators of spatial association. Geographical Analysis 27:93–115.

Bachraty, C., P. Legendre, and D. Desbruyères. 2009. Biogeographic relationships among deep-sea hydrothermal vent faunas at global scale. Deep-Sea Research I 56:1371–1378.

---

[17] https://r-forge.r-project.org/projects/sedar/

Baraloto, C., and P. Couteron. 2010. Fine-scale microhabitat heterogeneity in a French Guianan Forest. Biotropica 42:420–428.

Beale, C. M., J. J. Lennon, J. M. Yearsley, M. J. Brewer, and D. A. Elston. 2010. Regression analysis of spatial data. Ecology Letters 13:246–264.

Beeravolu, C. R., P. Couteron, R. Pélissier, and F. Munoz. 2009. Studying ecological communities from a neutral standpoint: a review of models' structure and parameter estimation. Ecological Modelling 220:2603–2610.

Bellier, E., P. Monestiez, J. P. Durbec, and J. N. Candau. 2007. Identifying spatial relationships at multiple scales: principal coordinates of neighbour matrices (PCNM) and geostatistical approaches. Ecography 30:385–399.

Bini, M. L., et al. 2009. Coefficient shifts in geographical ecology: an empirical evaluation of spatial and non-spatial regression. Ecography 32:193–204.

Bivand, R. 1980. A Monte Carlo study of correlation coefficient estimation with spatially autocorrelated observations. Quaestiones Geographicae 6:5–10.

Blanchet, F. G., P. Legendre, and D. Borcard. 2008a. Modelling directional spatial processes in ecological data. Ecological Modelling 215:325–336.

Blanchet, F. G., P. Legendre, and D. Borcard. 2008b. Forward selection of explanatory variables. Ecology 89:2623–2632.

Borcard, D., F. Gillet, and P. Legendre. 2011. Numerical ecology with R. Springer, New York, New York, USA.

Borcard, D., and P. Legendre. 1994. Environmental control and spatial structure in ecological communities: an example using oribatid mites (Acari, Oribatei). Environmental and Ecological Statistics 1:37–61.

Borcard, D., and P. Legendre. 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecological Modelling 153:51–68.

Borcard, D., and P. Legendre. 2012. Is the Mantel correlogram powerful enough to be useful in ecological analysis? A simulation study. Ecology 93:1473–1481.

Borcard, D., P. Legendre, C. Avois-Jacquet, and H. Tuomisto. 2004. Dissecting the spatial structure of ecological data at multiple scales. Ecology 85:1826–1832.

Borcard, D., P. Legendre, and P. Drapeau. 1992. Partialling out the spatial component of ecological variation. Ecology 73:1045–1055.

Cavender-Bares, J., D. D. Ackerly, D. A. Baum, and F. A. Bazzaz. 2004. Phylogenetic overdispersion in Floridian oak communities. American Naturalist 163:823–843.

Chave, J., R. Condit, S. Aguilar, A. Hernandez, S. Lao, and R. Perez. 2004. Error propagation and scaling for tropical forest biomass estimates. Philosophical Transactions of the Royal Society B 359:409–420.

Chave, J., and E. G. Leigh. 2002. A spatially explicit neutral model of β-diversity in tropical forests. Theoretical Population Biology 62:153–168.

Cliff, A. D., and J. K. Ord. 1973. Spatial autocorrelation. Pion, London, UK.

Cliff, A. D., and J. K. Ord. 1981. Spatial processes: models and applications. Pion, London, UK.

Condit, R., et al. 2002. Beta-diversity in tropical forest trees. Science 295:666–669.

Cottenie, K. 2005. Integrating environmental and spatial processes in ecological community dynamics. Ecology Letters 8:1175–1182.

Couteron, P., and S. S. Ollier. 2005. A generalized, variogram-based framework for multi-scale ordination. Ecology 86:828–834.

Couteron, P., and R. Pélissier. 2004. Additive apportioning of species diversity: towards more sophisticated models and analyses. Oikos 107:215–221.

Cressie, N. A. C. 1993. Statistics for spatial data. Wiley Interscience, New York, New York, USA.

Csillag, F., B. Boots, M.-J. Fortin, K. Lowell, and F. Potvin. 2001. Multiscale characterization of boundaries and landscape ecological patterns. Geomatica 55:509–522.

Currie, D. J. 2007. Disentangling the roles of environment and space in ecology. Journal of Biogeography 34:2009–2011.

Dale, M. R. T., and M.-J. Fortin. 2010. From graphs to spatial graphs. Annual Review of Ecology, Evolution, and Systematics 41:21–38.

Diniz-Filho, J. A. F., and L. M. Bini. 2005. Modelling geographical patterns in species richness using eigenvector-based spatial filters. Global Ecology and Biogeography 14:177–185.

Dormann, C. F., et al. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30:609–628.

Dray, S. 2011. A new perspective about Moran's coefficient: spatial autocorrelation as a linear regression problem. Geographical Analysis 43:127–141.

Dray, S., P. Legendre, and P. R. Peres-Neto. 2006. Spatial modeling: a comprehensive framework for principal coordinate analysis of neighbor matrices (PCNM). Ecological Modelling 196:483–493.

Dray, S., N. Pettorelli, and D. Chessel. 2002. Matching data sets from two different spatial samples. Journal of Vegetation Science 13:867–874.

Dray, S., S. Saïd, and F. Débias. 2008. Spatial ordination of vegetation data using a generalization of Wartenberg's multivariate spatial correlation. Journal of Vegetation Science 19:45–56.

Dungan, J. L., J. N. Perry, M. R. T. Dale, P. Legendre, S. Citron-Pousty, M.-J. Fortin, A. Jakomulska, M. Miriti, and M. S. Rosenberg. 2002. A balanced view of scale in spatial statistical analysis. Ecography 25:626–640.

Fall, A., M.-J. Fortin, and M. Manseau, and D. O'Brien. 2007. Spatial graphs: principles and applications for habitat connectivity. Ecosystems 10:448–461.

Ferrier, S. 2002. Mapping spatial pattern in biodiversity for regional conservation planning: where to from here? Systematic Biology 51:331–363.

Ferrier, S., and A. Guisan. 2006. Spatial modelling of biodiversity at the community level. Journal of Applied Ecology 43:393–404.

Ferrier, S., G. Manion, J. Elith, and K. Richardson. 2007. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. Diversity and Distributions 13:252–264.

Finley, A. O. 2011. Comparing spatially-varying coefficients models for analysis of ecological data with non-stationary and anisotropic residual dependence. Methods in Ecology and Evolution 2:143–154.

Fortin, M.-J. 1994. Edge detection algorithms for two-dimensional ecological data. Ecology 75:956–965.

Fortin, M.-J., and M. R. T. Dale. 2005. Spatial analysis. a guide for ecologists. Cambridge University Press, Cambridge, UK.

Fortin, M.-J., and P. Drapeau. 1995. Delineation of ecological boundaries: comparison of approaches and significance tests. Oikos 72:323–332.

Fotheringham, A. S., C. Brunsdon, and M. E. Charlton. 2002. Geographically weighted regression: the analysis of spatially varying relationships. Wiley, Chichester, UK.

Gauch, H. G. 1982. Multivariate analysis in community ecology. Cambridge University Press, Cambridge, UK.

Geary, R. C. 1954. The contiguity ratio and statistical mapping. Incorporated Statistician 5:115–145.

Gimaret-Carpentier, C., S. Dray, and J.-P. Pascal. 2003. Broad-scale biodiversity pattern of the endemic tree flora of the Western Ghats (India) using canonical correlation analysis of herbarium records. Ecography 26:429–444.

Gittins, R. 1985. Canonical analysis—a review with applications in ecology. Springer Verlag, Berlin, Germany.

Goodall, D. W. 1954. Objective methods for the classification of vegetation III. An essay on the use of factor analysis. Australian Journal of Botany 2:304–324.

Gordon, A. D. 1996. A survey of constrained classification. Computational Statistics and Data Analysis 21:17–29.

Griffith, D. A. 1992. What is spatial autocorrelation? Reflection on the past 25 years of spatial statistics. L'Espace Géographique 3:265–280.

Griffith, D. A. 1996. Spatial autocorrelation and eigenfunctions of the geographic weights matrix accompanying geo-referenced data. Canadian Geographer 40:351–367.

Griffith, D. A. 2000. A linear regression solution to the spatial autocorrelation problem. Journal of Geographical Systems 2:141–156.

Griffith, D. A. 2008. Spatial-filtering-based contributions to a critique of geographically weighted regression (GWR). Environment and Planning A 40:2751–2769.

Griffith, D. A., and P. R. Peres-Neto. 2006. Spatial modeling in ecology: the flexibility of eigenfunction spatial analyses. Ecology 87:2603–2613.

Heegaard, E., and V. Vandvik. 2004. Climate change affects the outcome of competitive interactions—an application of principal response curves. Oecologia 139:459–466.

Ives, A. R., and M. R. Helmus. 2010. Phylogenetic metrics of community similarity. American Naturalist 176:128–142.

Jacquez, G. M., A. Kaufmann, and P. Goovaerts. 2008. Boundaries, links and clusters: a new paradigm in spatial analysis? Environmental and Ecological Statistics 15:403–419.

Jin, X., B. P. Carlin, and S. Banerjee. 2005. Generalized hierarchical multivariate CAR models for areal data. Biometrics 61:950–961.

Jombart, T., S. Devillard, A. B. Dufour, and D. Pontier. 2008. Revealing cryptic spatial patterns in genetic variability by a new multivariate method. Heredity 101:92–103.

Jombart, T., S. Dray, and A. B. Dufour. 2009. Finding essential scales of spatial variation in ecological data: a multivariate approach. Ecography 32:161–168.

Keitt, T. H., and J. Fischer. 2006. Detection of scale-specific community dynamics using wavelets. Ecology 87:2895–2904.

Laliberté, E. 2008. Analyzing or explaining beta diversity? Comment. Ecology 89:3232–3237.

Leibold, M. A., E. P. Economo, and P. Peres-Neto. 2010. Metacommunity phylogenetics: separating the roles of environmental filters and historical biogeography. Ecology Letters 13:1290–1299.

Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm. Ecology 74:1659–1673.

Legendre, P., and D. Borcard. 2006. Quelles sont les échelles spatiales importantes dans un écosystème? Pages 425–442 in J. J. Droesbeke, M. Lejeune, and G. Saporta, editors. Analyse statistique de données spatiales TECHNIP, Paris, France.

Legendre, P., D. Borcard, and P. R. Peres-Neto. 2005. Analyzing beta diversity: partitioning the spatial variation of community composition data. Ecological Monographs 75:435–450.

Legendre, P., D. Borcard, and P. R. Peres-Neto. 2008. Analyzing or explaining beta diversity? Comment. Ecology 89:3238–3244.

Legendre, P., M. R. T. Dale, M.-J. Fortin, J. Gurevitch, M. Hohn, and D. E. Myers. 2002. The consequences of spatial structure for the design and analysis of ecological field surveys. Ecography 25:601–615.

Legendre, P., and M.-J. Fortin. 1989. Spatial pattern and ecological analysis. Vegetatio 80:107–138.

Legendre, P., and M.-J. Fortin. 2010. Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. Molecular Ecology Resources 10:831–844.

Legendre, P., and E. D. Gallagher. 2001. Ecologically meaningful transformations for ordination of species data. Oecologia 129:271–280.

Legendre, P., and L. Legendre. 2012. Numerical ecology. Third English edition. Elsevier Science, Amsterdam, The Netherlands.

Legendre, P., X. Mi, H. Ren, K. Ma, M. Yu, I. F. Sun, and F. He. 2009. Partitioning beta diversity in a subtropical broad-leaved forest of China. Ecology 90:663–674.

Legendre, P., and M. Troussellier. 1988. Aquatic heterotrophic bacteria: modeling in the presence of spatial autocorrelation. Limnology and Oceanography 33:1055–1067.

Levin, S. A. 1992. The problem of pattern and scale in ecology. Ecology 73:1943–1967.

Lichstein, J. W. 2007. Multiple regression on distance matrices: a multivariate spatial analysis tool. Plant Ecology 188:117–131.

Mahecha, M. D., and S. Schmidtlein. 2008. Revealing biogeographical patterns by nonlinear ordinations and derived anisotropic spatial filters. Global Ecology and Biogeography 17:284–296.

Manly, B. F. J. 1997. Randomization, bootstrap and Monte Carlo methods in biology. Second edition. Chapman and Hall, London, UK.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Research 27:209–220.

McIntire, E. J. B., and A. Fajardo. 2009. Beyond description: the active and effective way to infer processes from spatial patterns. Ecology 90:46–56.

Menge, B. A., and A. M. Olson. 1990. Role of scale and environmental factors in regulation of community structure. Trends in Ecology and Evolution 5:52–57.

Monmonier, M. S. 1973. Maximum-difference barriers: an alternative numerical regionalization method. Geographical Analysis 5:245–261.

Moran, P. A. P. 1950. Notes on continuous stochastic phenomena. Biometrika 37:17–23.

Munoz, F. 2009. Distance-based eigenvector maps (DBEM) to analyse metapopulation structure with irregular sampling. Ecological Modelling 220:2683–2689.

Munoz, F., P. O. Cheptou, and F. Kjellberg. 2007. Spectral analysis of simulated species distribution maps provides insights into metapopulation dynamics. Ecological Modelling 205:314–322.

Oden, N. L., and R. R. Sokal. 1986. Directional autocorrelation: an extension of spatial correlograms to two dimensions. Systematic Zoology 35:608–617.

Oden, N., R. Sokal, M. Fortin, and H. Goebl. 1993. Categorical wombling: detecting regions of significant change in spatially located categorical variables. Geographical Analysis 25:315–336.

Ollier, S., P. Couteron, and D. Chessel. 2006. Orthonormal transform to decompose the variance of a life-history trait across a phylogenetic tree. Biometrics 62:471–477.

Pélissier, R., P. Couteron, and S. Dray. 2008. Analyzing or explaining beta diversity? Comment. Ecology 89:3227–3332.

Pélissier, R., P. Couteron, S. Dray, and D. Sabatier. 2003. Consistency between ordination techniques and diversity measurements: two alternative strategies for species occurrence data. Ecology 84:242–251.

Pélissier, R., S. Dray, and D. Sabatier. 2002. Within-plot relationships between tree species occurrences and hydrological soil constraints: an example in French Guiana investigated through canonical correlation analysis. Plant Ecology 162:143–156.

Peres-Neto, P. R. 2004. Patterns in the co-occurrence of fish species in streams: the role of site suitability, morphology and phylogeny versus species interactions. Oecologia 140:352–360.

Peres-Neto, P. R., and P. Legendre. 2010. Estimating and controlling for spatial structure in the study of ecological communities. Global Ecology and Biogeography 19:174–184.

Peres-Neto, P. R., P. Legendre, S. Dray, and D. Borcard. 2006. Variation partitioning of species data matrices: estimation and comparison of fractions. Ecology 87:2614–2625.

Peres-Neto, P. R., M. A. Leibold, and S. Dray. 2012. Effects of spatial contingency and environmental filtering on metacommunity phylogenetics. Ecology 93(Supplement):S14–S30.

Philibert, M. D., M. J. Fortin, and F. Csillag. 2008. Spatial structure effects on the detection of patches boundaries using local operators. Environmental and Ecological Statistics 15:S447–S467.

Platt, T., and K. L. Denman. 1975. Spectral analysis in ecology. Annual Review of Ecology and Systematics 6:189–210.

Pyke, C. R., R. Condit, S. Aguilar, and S. Lao. 2001. Floristic composition across a climatic gradient in a neotropical lowland forest. Journal of Vegetation Science 12:553–566.

Radeloff, V. C., T. F. Miller, H. S. He, and D. J. Mladenoff. 2000. Periodicity in spatial data and geostatistical models: autocorrelation between patches. Ecography 23:81–91.

Rangel, T. F. L. V. B., J. A. F. Diniz-Filho, and R. K. Colwell. 2007. Species richness and evolutionary niche dynamics: a spatial pattern-oriented simulation experiment. The American Naturalist 170:602–616.

Renshaw, E., and E. D. Ford. 1984. The description of spatial pattern using two-dimensional spectral analysis. Vegetatio 56:75–85.

Ricklefs, R. E. 1987. Community diversity: relative roles of local and regional processes. Science 235:167.

Salomon, Y., S. R. Connolly, and L. Bode. 2010. Effects of asymmetric dispersal on the coexistence of competing species. Ecology Letters 13:432–441.

Seiferling, I. S., R. Proulx, P. R. Peres-Neto, L. Fahrig, and C. Messier. *In press*. Measuring protected-area isolation and correlations of isolation with land-use intensity and protection status. Conservation Biology.

Smith, T. W., and J. T. Lundholm. 2010. Variation partitioning as a tool to distinguish between niche and neutral processes. Ecography 33:648–655.

Soininen, J., R. McDonald, and H. Hillebrand. 2007. The distance decay of similarity in ecological communities. Ecography 30:3–12.

Thioulouse, J., D. Chessel, and S. Champely. 1995. Multivariate analysis of spatial patterns: a unified approach to local and global structures. Environmental and Ecological Statistics 2:1–14.

Tiefelsdorf, M., and D. A. Griffith. 2007. Semi-parametric filtering of spatial autocorrelation: the eigenvector approach. Environment and Planning A 39:1193–1221.

Tuomisto, H., K. Ruokolainen, M. Aguilar, and A. Sarmiento. 2003. Floristic patterns along a 43-km long transect in an Amazonian rain forest. Journal of Ecology 91:743–756.

Vellend, M. 2010. Conceptual synthesis in community ecology. Quarterly Review of Biology 85:183–206.

Wackernagel, H. 2003. Multivariate geostatistics. Third edition. Springer-Verlag, Berlin, Germany.

Wagner, H. H. 2003. Spatial covariance in plant communities: integrating ordination, geostatistics, and variance testing. Ecology 84:1045–1057.

Wagner, H. H. 2004. Direct multi-scale ordination with canonical correspondence analysis. Ecology 85:342–351.

Wagner, H. H., and M. J. Fortin. 2005. Spatial analysis of landscapes: concepts and statistics. Ecology 86:1975–1987.

Wartenberg, D. 1985. Multivariate spatial correlation: a method for exploratory geographical analysis. Geographical Analysis 17:263–283.

Webster, R., and A. B. McBratney 1989. On the Akaike information criterion for choosing models for variograms of soil properties. Journal of Soil Science 40:493–496.

Whittaker, R. H. 1972. Evolution and measurement of species diversity. Taxon 21:213–251.

Wiens, J. A. 1989. Spatial scaling in ecology. Functional Ecology 3:385–397.

Womble, W. H. 1951. Differential systematics. Science 114:315–322.

SUPPLEMENTAL MATERIAL

Appendix

From spatial weighting matrix (SWM) to Moran's eigenvector maps (MEM) (*Ecological Archives* M082-011-A1).

Supplement

Data and R scripts to reproduce the different analyses of the tropical forest data set (*Ecological Archives* M082-011-S1).

REVIEWS