2005

Spatial analysis of landscapes: concepts and statistics

Published version

Helene H. Wagner

Marie-Josée Fortin


Copyright by the Ecological Society of America.

HOW TO CITE TSpace ITEMS

Always cite the published version, so the author(s) will receive recognition through services that track citation counts, e.g. Scopus. If you need to cite the page number of the TSpace version (original manuscript or accepted manuscript) because you cannot access the published version, then cite the TSpace version in addition to the published version using the permanent URI (handle) found on the record page.
SPATIAL ANALYSIS OF LANDSCAPES: CONCEPTS AND STATISTICS

HELENE H. WAGNER1,3 AND MARIE-JOSÉE FORTIN2

1WSL Swiss Federal Research Institute, 8903 Birmensdorf Switzerland
2Department of Zoology, University of Toronto, Ontario M5S 3G5 Canada

Abstract. Species patchiness implies that nearby observations of species abundance tend to be similar or that individual conspecific organisms are more closely spaced than by random chance. This can be caused either by the positive spatial autocorrelation among the locations of individual organisms due to ecological spatial processes (e.g., species dispersal, competition for space and resources) or by spatial dependence due to (positive or negative) species responses to underlying environmental conditions. Both forms of spatial structure pose problems for statistical analysis, as spatial autocorrelation in the residuals violates the assumption of independent observations, while environmental heterogeneity restricts the comparability of replicates. In this paper, we discuss how spatial structure due to ecological spatial processes and spatial dependence affects spatial statistics, landscape metrics, and statistical modeling of the species–environment correlation. For instance, while spatial statistics can quantify spatial pattern due to an endogeneous spatial process, these methods are severely affected by landscape environmental heterogeneity. Therefore, statistical models of species response to the environment not only need to accommodate spatial structure, but need to distinguish between components due to exogeneous and endogeneous processes rather than discarding all spatial variance. To discriminate between different components of spatial structure, we suggest using (multivariate) spatial analysis of residuals or delineating the spatial realms of a stationary spatial process using boundary detection algorithms. We end by identifying conceptual and statistical challenges that need to be addressed for adequate spatial analysis of landscapes.

Key words: autocorrelation; landscape metrics; multivariate analysis; multiscale ordination; spatial analysis; spatial regression; stationarity.

INTRODUCTION

Ecology has seen a paradigm shift from the assumption of homogeneity to the recognition of heterogeneity as a key for understanding the complexity of nature (Wiens 1989). The explicit consideration of spatial structure and spatiotemporal interaction of processes in ecological research is the main contribution of landscape ecology to this paradigm shift. Acknowledging the importance of spatial pattern and scale has changed the way ecological studies are designed and analyzed, and has provided new insights about ecological processes (Allen and Hoekstra 1992). Most ecological processes are inherently spatial as they operate between neighboring units (Levin 1992). Processes are also constrained by environmental conditions varying in space and time and by the local interaction with other processes, resulting in interwoven patterns at multiple spatial and temporal scales.

As the primary concern of ecology is the identification and understanding of ecological processes, complicating factors such as spatial heterogeneity were at first excluded from the conceptual framework of analysis (McIntosh 1991). By doing so, ecological studies could assume homogeneity, permitting the incorporation of environmental variation as a treatment or to control for known relationships using covariates. For instance, the effects of different levels of an environmental factor can be tested in an experimental setting using ANOVA-type analyses or analyzed along existing gradients using regression-type analyses. Hence by assuming that the study area is locally homogeneous with respect to that factor in space and time, each experimental plot, or sampling unit, is attributed to a single factor level and the neighborhood context of the plot or sampling unit does not matter (Fig. 1A). In contrast, landscape ecology assumes that the neighborhood context affects the ecological processes within a plot and the interaction between plots (Fig. 1B). A further complication is that environmental heterogeneity may occur at any spatial scale, and site conditions may vary in time.

The patchiness of species, and other ecological response variables, forms another type of spatial heterogeneity that ecologists need to consider (Fig. 2). Patchiness, created by ecological spatial processes such as competitive interactions or dispersal, violates the assumption of parametric tests that the residual errors are independent (Legendre 1993: Fig. 1A). Indeed, patchiness induces autocorrelation in the error structure of
an ANOVA or regression-type model (Fig. 1B), which reduces the degrees of freedom of the associated statistical tests (Dale and Fortin 2002).

The growing acceptance of the heterogeneous nature of ecological systems requires adapting ecological theory and methods to accommodate for “heterogeneity.” There is, however, little consensus on the exact meaning of the term (Kolasa and Rollo 1991, Li and Reynolds 1995). Here, we define spatial heterogeneity as the spatially structured variability of a property of interest, which may be a categorical or quantitative, explanatory or dependent variable.

When dealing with heterogeneity, one needs to consider some fundamental questions about the causes, types, and ecological consequences of heterogeneity. Approaches to answer these questions evolved in different contexts, ranging from population genetics to species diversity and ecosystem processes. Methods were borrowed from various fields, including geography, geology, spatial econometrics, physics, plant community ecology, and complex systems theory. While the different approaches can be contrasted by spatial data representation (Gustafson 1998, Dale et al. 2002, Perry et al. 2002), objective (Liebhold and Gurevitch 2002, Ver Hoef 2002), or disciplinary background (Liebhold and Gurevitch 2002), they often face similar challenges in attempting to quantify heterogeneity.

This paper brings together some common analytical threads related to the spatial analysis of ecological data at the landscape level, while pointing to unresolved conceptual and statistical challenges. We start with summarizing the causes, types, and ecological consequences of spatial heterogeneity, focusing on relevant aspects for the design and analysis of an ecological study. We then discuss how and to what degree three different approaches (namely spatial statistics, landscape metrics and statistical modeling), deal with these aspects of spatial heterogeneity. Specifically, we highlight how these three spatial approaches can provide new insights about landscape spatial pattern and to what degree these methods can disentangle the patterns due to species response to a spatially structured environment and those due to ecological spatial processes. Finally, we point to promising new approaches at meeting the challenges of spatial analysis in a heterogeneous environment, including the statistical assessment of changes in space and time, the quantification of local landscape structure, and the merging of discrete and continuous landscape models.

**Causes of heterogeneity**

Any spatial process operating between neighboring units can cause spatial heterogeneity. Fig. 2A shows a simulated random distribution of a species in a homogeneous environment, while Fig. 2B illustrates the patchy distribution produced by a simple spatial process starting from the pattern in Fig. 2A. Spatial analysis aims to assess the process generating these non-random patterns. As this process is stochastic, Fig. 2B represents only one of many possible outcomes of the same process given the initial conditions in Fig. 2A (Fortin et al. 2003). In practice, however, we often have only one observed pattern representing a single realization of the process of interest, which makes inference about this process difficult.

Inference from a pattern on the underlying process is further hindered by variation in the process in space or time as well as by the presence of additional, confounding processes. Fig. 2C shows the random distribution of the simulated species in Fig. 2A but constrained by a linear environmental gradient, and Fig. 2D reflects the confounded pattern of patchiness and an environmental gradient. In fact, most of the observed patterns result from more than one processes that possibly interact with each other, such as biotic (e.g., ecological spatial processes) and abiotic (e.g., environmental factors) drivers (Fig. 3).

**Types of heterogeneity**

The heterogeneity of a categorical variable is best described by a mosaic of patches. This includes the special case of binary data, where only one factor level is of interest (e.g., patches of suitable habitat) and any other levels are collapsed into one (e.g., matrix of non-habitat). The basic properties of a mosaic are composition and configuration: composition describes the number and relative frequency of the factor levels (e.g., habitat types), whereas configuration refers to the spa-
Fig. 2. Simulated species distribution in a grid of $40 \times 40$ cells under different combinations of a homogeneous environment (A, B) or a linear environmental gradient (C, D) with random (A, C) or patchy (B, D) distribution of the species.

Spatial arrangement of the patches defined by the factor levels (Gustafson 1998).

For a quantitative variable, the distinction between composition and configuration is not as straightforward. Composition refers to the density distribution function of the variable, whereas configuration is usually described in terms of the spatial covariance structure of the variable. The latter summarizes the strength, range, and directionality (anisotropy) of the spatial autocorrelation. The intensity of spatial autocorrelation is related to the degree of self-similarity of the values of a variable at nearby locations that can be expressed in terms of fractal dimension (Palmer 1992, McGarigal and Cushman 2005).

The type of heterogeneity depends on the nature of the variable rather than how it is sampled, analyzed, or displayed. For example, in geographic information systems (GIS), categorical data typically are represented by vector data (polygon maps), whereas quantitative data are treated as raster data (grid surfaces). However, drawing a line between two values of a quantitative variable such as biomass is artificial. Similarly, displaying a qualitative variable as a mosaic-like grid surface, by resampling a categorical map of patches at regular intervals, does not make the abrupt transition between two patches any smoother. This example illustrates that ecological data may not always fit easily in either GIS vector or raster data type, and the choice may have implications for our ability to detect patterns and insights about the underlying processes that generated them (Cova and Goodchild 2002, Cushman and McGarigal 2004).

Ecological consequences of spatial heterogeneity

The pattern created by one process may affect another process and its resulting pattern (Levin 1992). In a homogeneous environment, for instance, spatial population dynamics can create heterogeneity in the abundance of a species. Nearby locations that are linked by dispersal tend to have interdependent population dynamics, leading to autocorrelation in species abundance (Fig. 3). The land-use mosaic imposes additional constraints on the local population dynamics, introducing spatial structure in species abundances due to the spatial distribution of site conditions and disturbance. In a spatially structured environment, where nearby locations tend to have similar site conditions, the pattern induced by species response to spatially structured environmental factors may be mistaken for spatial autocorrelation due to a spatial ecological process. Hence, the environmental heterogeneity creates exogeneous spatial dependence in the species ecological process, while the spatial interaction in the population dynamics is an endogeneous spatial ecological process. Not only may
Fig. 3. Spatial effects in ecological data. Species are spatially structured for several reasons: (1) ecological processes are inherently spatial as they operate between neighboring individuals, thus creating autocorrelation; (2) species respond to variation in environmental factors, which are themselves spatially structured, thus inducing spatial dependence in species distributions; and (3) species respond to the environment at a specific scale, they may respond to the same factor differently at different scales, and the response may be nonlinear. Thus, the exogenous spatial structure may be more complex than the spatial structure of the environment.

the observed spatial pattern of abundance include both types of underlying processes (Fig. 3), but these processes may interact in a linear or non-linear way. For instance, the land-use mosaic may constrain the dispersal of organisms if some land-use types are more difficult to traverse than others. Thus, the probability that two habitat patches separated by a given distance are connected by dispersal depends on the land-use in-between (D’Eon et al. 2002). It is clear from this example that species patchiness and the spatial structure induced by environmental heterogeneity depend on the perspective of a specific organism, as habitat requirements, life history attributes, and dispersal abilities will vary between species. Hence, species may respond to environmental heterogeneity in a non-linear manner (e.g., minimum threshold, or patch size requirements), it may require a specific combination of factor levels within its home range (Fahrig 2002), or it may respond to the temporal variability of environmental factors. Overlaps and interactions of different processes pose a formidable challenge to ecological research that explicitly investigates the spatial response of a species to landscape heterogeneity, as is the case in metapopulation studies.

Spatial Approaches to Landscape Analysis

There are important practical considerations for the spatial analysis of landscapes (as summarized in Table 1) that should be incorporated into students’ ecological curricula. Here, we discuss the advantages and limitations of three analytic approaches to the analysis of spatial heterogeneity: spatial statistics, landscape metrics and spatial regression modeling. These approaches differ in their assumption on the number of underlying processes and in their general objective (Table 2). To better appreciate these considerations, we will first revisit the main philosophical principles assumed by Fisherian (parametric) statistical methods used in ecology.

Nonspatial (Fisherian) statistics

In controlled experiments, nonspatial statistics have been extremely powerful to quantify and test ecological relationships. Unfortunately, when applied to heterogeneous systems, most parametric statistics and multivariate statistics (e.g., ordination) are usually applied in inappropriate ways (e.g., Legendre 1993, Dale and Fortin 2002). Correlation analysis, for example, quantifies the association between two response variables, such as the abundance of two species. The method assumes independence of the residual errors which is usually achieved by using a random sample from a homogeneous environment as depicted in Fig. 2A. In the presence of patchy data, a random sampling design (Fortin et al. 1989) and completely randomized experimental design (Legendre et al. 2004) do not guarantee that the residual errors are independent. In order to account for patchiness (Fig. 2B), Dutilleul (1993) presented a corrected $t$ test for pairwise correlation coefficients, which adjusts the degrees of freedom proportionally to the degree of spatial autocorrelation present in each variable. Alternatively, spatially constrained (restricted) randomization tests have been proposed for testing interspecific interactions while accounting for species-specific patchiness (cf. Roxburgh and Matsuki 1999). Note that methods accounting for species patchiness may still be invalid due to environmental heterogeneity (Fig. 2C and 2D) if the correlation changes with site conditions (Legendre et al. 2002, 2004).
Stationarity assumptions concern the model of the species–environment correlation is likely to have a full degree of freedom such that the significance of parametric statistics (e.g., correlation, regression, ANOVA) is not assessed with the appropriate degree of freedom. Based on data alone, it is not possible to distinguish between exogenous deterministic structure (spatial dependence) and endogenous spatial autocorrelation (ecological spatial process).

The species–environment correlation is likely to change with scale (Levin 1992). Stationarity assumptions concern the model of the underlying process and allow inference from the observed pattern to the entire study area. Note that an observed pattern is a single realization of that process (Fortin et al. 2003).

Stationarity rarely prevails in real landscapes; the data may show a trend (change in mean) or local variability (change in variance) (Fortin et al. 2003).

ANOVA or regression models may be used for relating population density to one or several environmental factors, thus assuming independent observations from a heterogeneous environment. Although such models explicitly include variability in at least one environmental factor, they are susceptible to spatial effects (Fig. 2C). Spatial autocorrelation in the residuals may render statistical tests too liberal (Cliff and Ord 1981). Individual observations may not bring a full degree of freedom such that the significance of parametric statistics (e.g., correlation, regression, ANOVA) is not assessed with the appropriate degree of freedom.

Based on data alone, it is not possible to distinguish between exogenous deterministic structure (spatial dependence) and endogenous spatial autocorrelation (ecological spatial process).

The species–environment correlation is likely to change with scale (Levin 1992). Stationarity assumptions concern the model of the underlying process and allow inference from the observed pattern to the entire study area. Note that an observed pattern is a single realization of that process (Fortin et al. 2003).

Stationarity rarely prevails in real landscapes; the data may show a trend (change in mean) or local variability (change in variance) (Fortin et al. 2003).

If the scale of patchiness is known, it can be used to enforce an appropriate minimum distance between observations for a systematic or a random sample (Dungan et al. 2002).

Tests adjusting the degree of freedom according to the degree of spatial autocorrelation in the data should be used (Dale and Fortin 2002). When analyzing directional relationships (regression, ANOVA), this is only necessary if there is autocorrelation in the residuals, whereas autocorrelation in the raw data may not be a problem.

Based on data alone, it is not possible to distinguish between exogenous deterministic structure (spatial dependence) and endogenous spatial autocorrelation (ecological spatial process).

The species–environment correlation is likely to change with scale (Levin 1992). Stationarity assumptions concern the model of the underlying process and allow inference from the observed pattern to the entire study area. Note that an observed pattern is a single realization of that process (Fortin et al. 2003).

Stationarity rarely prevails in real landscapes; the data may show a trend (change in mean) or local variability (change in variance) (Fortin et al. 2003).

If the scale of patchiness is known, it can be used to enforce an appropriate minimum distance between observations for a systematic or a random sample (Dungan et al. 2002).

Tests adjusting the degree of freedom according to the degree of spatial autocorrelation in the data should be used (Dale and Fortin 2002). When analyzing directional relationships (regression, ANOVA), this is only necessary if there is autocorrelation in the residuals, whereas autocorrelation in the raw data may not be a problem.

Hypothesis testing and experimental design are needed to disentangle these two possibilities (Legendre et al. 2004).

A multiscale study design is needed unless the scale of response is known (Fortin et al. 1989; Cushman and McGarigal 2004). The presence of stationarity could be either assumed when the behavior of the underlying process is known or checked by estimating local mean and variance using a moving window approach.

When the data show a spatial trend it should be removed only if it has an ecological interpretation, as the observed pattern may exhibit trend-like structure by chance. In the case of local variability, environmental heterogeneity needs to be measured and accounted for when quantifying spatial pattern (Wagner 2003).

Spatial statistics

Even though spatial statistics were developed in different fields (geography, ecology, economics, mining), many methods were developed as an adaptation of time series analysis to spatial problems. However, while time is a single dimension and temporal effects are unidirectional, geographic space has at least two dimensions, and spatial processes may operate in any direction and may not necessarily have the same intensity in all directions (i.e., anisotropic processes). The spatial statistical approaches most commonly used by ecologists differ in their practical objectives. Geostatistical methods focus on the estimation of the spatial covariance structure of a spatially structured variable (e.g., variogram modeling) in order to use the spatial parameters to interpolate values at unobserved locations (e.g., kriging). Spatial statistics, on the other hand, aim at testing for the presence of a spatial process in order to model this process or to account for spatial autocorrelation when assessing the relationship between spatially structured variables (Cliff and Ord 1981, Fortin et al. 2001, Liebhold and Gurevitch 2002).

Spatial statistics that test for spatial autocorrelation (e.g., Moran’s I, Geary’s c) assume stationarity, meaning that the underlying process should have at least roughly the same parameter values (mean and variance) for the entire study area (Fig. 2B). These global spatial statistics (Boots 2002) further assume that the spatial covariance structure of the variable (i.e., the values of

### Table 1. Six “points of wisdom” to keep in mind for the spatial analysis of landscapes.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Practical implication</th>
</tr>
</thead>
<tbody>
<tr>
<td>A random sample does not guarantee independent observations, but is designed to avoid bias (Fortin et al. 1989).</td>
<td>If the scale of patchiness is known, it can be used to enforce an appropriate minimum distance between observations for a systematic or a random sample (Dungan et al. 2002).</td>
</tr>
<tr>
<td>Spatial autocorrelation in the residuals may render statistical tests too liberal (Cliff and Ord 1981). Individual observations may not bring a full degree of freedom such that the significance of parametric statistics (e.g., correlation, regression, ANOVA) is not assessed with the appropriate degree of freedom.</td>
<td>Tests adjusting the degree of freedom according to the degree of spatial autocorrelation in the data should be used (Dale and Fortin 2002). When analyzing directional relationships (regression, ANOVA), this is only necessary if there is autocorrelation in the residuals, whereas autocorrelation in the raw data may not be a problem.</td>
</tr>
<tr>
<td>Based on data alone, it is not possible to distinguish between exogenous deterministic structure (spatial dependence) and endogenous spatial autocorrelation (ecological spatial process).</td>
<td>Hypothesis testing and experimental design are needed to disentangle these two possibilities (Legendre et al. 2004).</td>
</tr>
<tr>
<td>The species–environment correlation is likely to change with scale (Levin 1992). Stationarity assumptions concern the model of the underlying process and allow inference from the observed pattern to the entire study area. Note that an observed pattern is a single realization of that process (Fortin et al. 2003).</td>
<td>A multiscale study design is needed unless the scale of response is known (Fortin et al. 1989; Cushman and McGarigal 2004). The presence of stationarity could be either assumed when the behavior of the underlying process is known or checked by estimating local mean and variance using a moving window approach.</td>
</tr>
<tr>
<td>Stationarity rarely prevails in real landscapes; the data may show a trend (change in mean) or local variability (change in variance) (Fortin et al. 2003).</td>
<td>When the data show a spatial trend it should be removed only if it has an ecological interpretation, as the observed pattern may exhibit trend-like structure by chance. In the case of local variability, environmental heterogeneity needs to be measured and accounted for when quantifying spatial pattern (Wagner 2003).</td>
</tr>
</tbody>
</table>
spatial autocorrelation at different spatial distances or lags) is similar over the entire study area. Nonstationary processes imply that the mean, variance, or spatial covariance structure of a variable vary across a study area which may pose severe problems to spatial statistics. In spatially heterogeneous landscapes (Fig. 2D), nonstationarity is likely to occur, so that the test may become too liberal in rejecting the null hypothesis of no autocorrelation.

Detrending is often used for addressing problems of non-stationarity (Haining 1997). For instance, a large-scale trend is removed prior to spatial analysis by fitting a linear or polynomial trend surface as a function of the geographic coordinates of the sampling units. Detrending removes the mean but does not affect the variance, which is often related to the mean and may still depend on the location. Non-parametric methods may be more robust in moderate cases of non-stationarity (Bjørnstad and Falck 2001), and join-count statistics have been extended to accommodate nonstationarity of the mean (Kabos and Csillag 2002). For instance, population density is likely to be related to environmental factors. When these factors are spatially structured, the assumption of stationarity may be met by performing spatial statistics on the residuals of an environmental response model (ANOVA, regression). However, there may still be problems due to nonconstant variance, or the spatial process itself may depend on the environmental factors.

**Landscape metrics and related measures**

The recent development of GIS provided ecologists with a technical framework for landscape-scale analyses (Greenberg et al. 2002). GIS include tools that characterize and quantify the properties of data (area, perimeter, proportion). To these basic tools, some spatial statistics have been added to analyze spatial patterns. Spatial analyses of landscapes have also been facilitated by the availability of remotely sensed images, from which land cover is derived into classes. In ecology, the landscape structure of such categorical data is usually quantified in terms of landscape composition (i.e., proportions of habitat patches) or landscape configuration (i.e., spatial arrangement of patches) using landscape metrics (O’Neill et al. 1988, Gustafson 1998; FRAGSTATS, available online).4

4 (http://www.umass.edu/landeco/research/fragstats/fragstats.html)

Landscape metrics are often used as predictors of ecological processes, such as dispersal, which results in the observable distribution of organisms across a landscape, but this approach suffers from several problems (Bélisle et al. 2001). (1) While no single index can capture landscape structure, many landscape metrics are strongly correlated (Gustafson 1998). Several authors have attempted, either empirically (McGarigal and McComb 1995, Rüitters et al. 1995) or theoretically (Li and Reynolds 1995), to identify the intrinsic dimensions (uncorrelated components) of landscape structure, but this search has not yet resulted in a generally applicable minimum set of landscape metrics (Gustafson 1998, Fortin et al. 2003). (2) Landscape metrics are highly sensitive to scale, i.e., the assessment of the structure of a landscape may change with the grain (resolution) and extent (area covered) of the map on which they are calculated (Cain et al. 1997, Turner et al. 2001, Wu et al. 2002). (3) An organism may respond to a landscape characteristic in a nonlinear way, such as requiring a specific minimum patch size or displaying threshold behavior in dispersal. In such cases, landscape metrics need to be rescaled in terms of organism characteristics. Alternatively, the nonlinear behavior could be modeled with a nonlinear regression model (e.g., by choosing an appropriate link function using generalized linear models; Guisan and Zimmerman 2000). (4) Landscape metrics assume the mapped property to be nominal or binary. In general, they do not consider ranks or other measures of gradual differences between factor levels, such as different levels of habitat suitability (Verbeylen et al. 2003). (5) Landscape metrics quantify the pattern of a categorical map and may be strongly affected by classification errors (e.g., if the map was derived from remote-sensing data) or other forms of uncertainty introduced during the mapping process. Reliable results can only be achieved by assessing the mapping uncertainty and its propagation in subsequent analysis (Hess 1994, Mowrer 1999).

The statistical properties of landscape metrics cannot be defined as they depend on the landscape composition, which can vary in the presence of spatial heterogeneity. Hence there are no standard tests for differences between two observed patterns, or rather their generating processes (Fortin et al. 2003). While each observed pattern corresponds to a single outcome of a stochastic process, inference about the process requires

---

**Table 2.** Main spatial approaches to analyze landscapes.

<table>
<thead>
<tr>
<th>No. processes</th>
<th>Spatial pattern analyses</th>
<th>Spatial modeling analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>Global spatial statistics (continuous variable)</td>
<td>Spatial regression analysis</td>
</tr>
<tr>
<td></td>
<td>Landscape metrics (categorical variable)</td>
<td>Spatial regression analysis</td>
</tr>
<tr>
<td>Several</td>
<td>Local spatial statistics</td>
<td>Partial canonical analysis</td>
</tr>
<tr>
<td></td>
<td>Local landscape metrics</td>
<td>Residual analysis</td>
</tr>
</tbody>
</table>

*Note: Approaches are grouped by the number of generating processes and on whether the analysis focuses on the description or the modeling of spatial pattern.*
knowledge of the distribution of patterns it may produce. Stochastic models can be used to derive such distributions by simulation (e.g., neutral landscape models), but these models typically assume stationarity. The interpretation of landscape pattern indices needs to be based on stochastic models that handle landscape heterogeneity and where spatial parameters are estimated from observed data (Fortin et al. 2003).

Quantifying landscape structure is rarely the ultimate goal for ecologists, but it is an important requisite for understanding how landscape structure affects ecological processes. However, it is not that easy to determine causality between process and pattern, as the correlation between landscape metrics and ecological processes is often inconsistent (Tischendorf 2001). In fact, there is no a priori causal ordering in space as there is in time, and there are no statistical techniques that will unambiguously uncover species-landscape relationships in the absence of informed ecological understanding that poses the hypothetical relationships which the statistics then test (Henebry and Merchant 2001). In a hypothesis-testing framework, graph theory in conjunction with a resource selection model, offers a promising approach to study species-environment relationships at the landscape level. This approach combines the topological spatial arrangement of landscape elements (patches) and species responses to patch types in terms of habitat preference (Urban and Keitt 2001, Manseau et al. 2002).

Statistical modeling

Modeling aims at quantifying the species–environment relationships by specifying the underlying processes (dynamic modeling) or by predicting the observed patterns of the organisms from the spatial distribution of environmental factors (statistical modeling). A new realm of spatially explicit models exist to model ecological processes (Dieckmann et al. 2000) as well as disturbances and their stochasticity (Mladenoff and Baker 1999). Here, however, we focus on statistical modeling in a regression context, highlighting three rather different approaches: (1) spatial regression models where a spatial term is added to a regression; (2) partialling-out methods (e.g., ordination techniques) where the spatial component is factored out while estimating species–environment relationships; and (3) residual analysis following a multiscale ordination that identifies and characterizes spatial components due to unobserved environmental factors or ecological spatial processes.

Spatial regression modeling.—This approach is most actively being developed in geography and spatial econometrics, although it is increasingly used in ecology (e.g., autologistic model; Lichstein et al. 2002, Fortin et al. 2003, Burgman et al. 2005). Here, we discuss three issues raised by Anselin (2002) in econometrics that are equally relevant for ecological applications. First, spatial modeling may be based on either of two data models, and the decision between the lattice and the random field models has far reaching implications. A metapopulation is a good example of a situation where the lattice model is appropriate. This model implies that each data point represents a discrete local population and that within the extent of the study, all local populations are included. The primary goal is extrapolation, or inference from the observed metapopulation (n = 1) to other metapopulations beyond the study area. Spatial analysis is based on the network, or topology, of local populations and requires that the neighbors for each population are defined and assigned appropriate weights. The specification of neighborhood and weights is essentially arbitrary, yet it may have a great influence on the results.

A typical example of a random field is the plant species richness of nonadjacent sampling quadrats, where the observations represent a systematic or random sample of the surface of the study area. The primary goal is the prediction (interpolation) of values at unobserved locations within the study area. The spatial covariance structure (e.g., obtained by estimating a variogram model), is fitted directly as a function of the geographic distance between quadrats without specifying neighbors or weights. However, quadrat size and shape, which are arbitrarily defined as part of the sampling design, may have a great effect on the estimated covariance structure (i.e., modifiable areal unit problem [MAUP]; Openshaw 1984, Dungan et al. 2002).

Second, it is important to distinguish between theory-driven and data-driven specification of the spatial regression model: is there a theoretical foundation for a spatial process, or does the residual spatial structure reflect shortcomings of the data? Spatial processes may include situations where the behavior of an organism is affected by the neighbors’ decisions, either directly or indirectly through the shared use of a limited resource, or it may result from a spatial diffusion process. Alternatively, spatial structure in the data may be due to a missing explanatory factor that is spatially structured, a mismatch of the scales of the process and the data, or spatially interpolated explanatory variables (Bradshaw and Fortin 2000, Dungan et al. 2002). The different processes may create similar patterns difficult to discriminate without experimental design and hypothesis testing. Nevertheless, a model should reflect the assumptions about the process. For example, a hypothesized spatial interaction can be modeled by a spatial lag model, which includes an autoregressive term (Cressie 1993), where the response y; at location i is a function of the neighboring values yj. A neighborhood response of organisms to the environment can be modeled by a spatial cross-regressive term where y; is a function of the environmental factor xj at neighboring locations j. In data-driven model specification, the residual spatial structure is interpreted as noise and modeled by a spatially correlated error term where the error ε i at location i is a function of the neighboring errors ε j.
Third, a spatial error term can be fitted simultaneously for all data points (AR model) or conditionally for each data point given the known values of its neighbors (CAR model; Cliff and Ord 1981, Griffith 1988, Keitt et al. 2002). Autoregressive models are often used for modeling a binary response variable describing the observed presence or absence of a species. It is important to understand that logistic regression models the latent probability of occurrence, which cannot be observed directly but only through its realized outcome as presence or absence. Only the conditional model (CAR) can deal with a spatial latent variable. However, the conditional model cannot explain the spatial pattern, and prediction is essentially limited to missing observations with known presence/absence information for all of its neighbors.

Partialling out the spatial component.—Spatial autocorrelation in the residuals may make statistical tests too liberal and affect parameter estimates, so that the importance of an environmental factor may be over- or underestimated (Keitt et al. 2002, Lichstein et al. 2002). Dutilleul’s (1993) modified t test adjusts the degree of freedom according to the degree of spatial autocorrelation in the data. Broad-scale spatial structure in the predictor combined with local spatial autocorrelation in the response may, however, reduce the power of Dutilleul’s modified t test (Legendre et al. 2002). The effect of any ecological factor (relevant or not) may be overestimated if it shows a similar spatial pattern as the observed response because both depend on the same, unmeasured environmental factor (Legendre and Legendre 1998, Lichstein et al. 2002).

In order to avoid such problems of false correlation, partialling-out methods can remove trends or large-scale spatial structure in the data before estimating regression parameters or performing constrained ordination. This can be achieved by fitting a polynomial trend surface (Borcard et al. 1992) or more complex and flexible models of spatial structure derived from the relative spatial locations of the sampling units (Borcard and Legendre 2002). However, spatial dependence may not indicate spurious correlation (Lichstein et al. 2002), nonspatial correlation does not guarantee causation, and the directionality and asymmetry of causal relationships must be explicitly assessed. Imagine a simple gradient with a linear increase of moisture along a transect. The plant species composition can be explained equally well by moisture as by transect position. After partialling out the spatial component, moisture has no explanatory power, although it is the moisture that the plants respond to. It is clear from this example that the spatial-dependence component is part of the species–environment correlation and should not be removed for parameter estimation without careful consideration. If the residuals are spatially correlated, however, this implies the presence of an unknown process, which may be accounted for by adding an autoregressive term or a spatial error term in the regression analysis (Haining 1997, Keitt et al. 2002, Lichstein et al. 2002).

Residual analysis with multiscale ordination.—Residual analysis may help to discriminate between spatial autocorrelation due to an ecological spatial process and spatial dependence induced by environmental response, and it may indicate specification errors such as the omission of an important factor or a mismatch of scales of observation and response. Ordinary regression analysis and constrained ordination methods partition the total variance in the uni- or multivariate response into two components, the explained and the residual variance (Fig. 4A). Regression residuals are commonly checked for (1) evidence of heteroscedasticity, where the variance depends on the mean (2) systematic deviation from the normal distribution, (3) influential observations that may have a large impact on parameter estimates, and, increasingly, (4) spatial autocorrelation. However, there is no equivalent for multivariate analysis, where the large number of response variables may make the above methods impractical. Partial constrained ordination can be used to further partition both the explained and the residual variance into a spatially structured and a nonspatial part (Fig. 4B), so that the relative importance of purely environmental (a) and purely spatial effects (c) can be compared and their degree of overlap (b) be assessed (Borcard et al. 1992).

**FIG. 4.** (A) Variance components of regression analysis or constrained ordination, (B) partial regression or ordination including space as a predictor, and (C) direct multiscale ordination. The components are (a) purely environmental effects, explained, not spatially structured variance; (b) overlap of spatial and environmental effects, spatially structured explained variance; (c) purely spatial effects, explained, spatially structured variance; and (d) unexplained variance that is not spatially structured. Components a and b appear in reversed order in (C) because a represents the nugget variance of the variogram of explained variance.
Spatial structure in the residuals may be due, however, to a spatial process or an unaccounted spatially structured environmental factor, and we have little power to tell these apart. Urban et al. (2002) suggested the development of partial Mantel correlograms to further investigate the spatial structure of the variance components. Direct multiscale ordination (MSO) with RDA or CCA (Wagner 2004) provides this information by estimating the total variance, the explained and the residual variance as well as the eigenvalues of ordination axes for a series of distance classes (Fig. 4C). The distance-dependent variance components, which are estimated from all pairs of observations that fall into a given distance class, are plotted against distance, resulting in a set of empirical variograms that effectively partition ordination results by distance. Fig. 5 provides an example of direct MSO for the simulated species distribution in Fig. 2D, which mimics the patchy distribution of a species along a simple environmental gradient. The global RDA results showed a total variance of 0.25, an explained variance of 0.1 and a residual variance of 0.15. Spatial partitioning by MSO revealed the spatial structure of the different variance components. The variogram of the total variance (circles) exhibited a continuous increase of variance with distance. After accounting for the environmental gradient, the variogram of the residual variance (squares) showed an initial increase before reaching a constant level. The spatial structure at larger distances was contained in the variance explained by the environmental gradient (triangles). The results of MSO can be used for checking modeling assumptions:

1) The variogram of the residual variance (Fig. 4C, thin line) provides an estimate of the scale of patchiness and may indicate problems with non-stationarity. For a stationary process, patchiness causes reduced variance at short distances, whereas at larger distances beyond the range, the variance reaches a constant level (sill). The range indicates the distance beyond which observations are spatially independent and may be used as a minimum distance in subsequent sampling (Fortin et al. 1989, Dungan et al. 2002, Legendre et al. 2004). In the presence of a sill, a Mantel test can be used to test each distance class for significant spatial autocorrelation (Wagner 2003, 2004). A continuous increase of the variance with distance, however, is often associated with spatial trend (Fig. 2C and D) and may indicate the presence of an unaccounted environmental factor that is spatially structured. If this is the case, the trend-like structure, which often exhibits directional (anisotropic) behavior, is likely contained in the first non-canonical axis. This can be checked by investigating the variogram of the respective eigenvalue. Plotting the axis scores in geographic space may help to identify the missing factor.

2) A systematic difference between the variogram of the total variance (Fig. 4C, bold line) and the sum of the variograms of the explained and residual variances (Fig. 4C, dashed line) may indicate problems with scale-dependence in the species–environment correlation. The global significance of an observed deviation can be tested using a point-wise confidence envelope for the variogram of total variance (Wagner 2004).

Conceptual and Statistical Challenges

Assessing changes in space and time

A major step forward in landscape ecology will lead from the “snap-shot mode” quantification of landscape structure to the “movie mode” assessment of changes in landscape structure in space and time. Testing the hypothesis that the generating process differs between landscapes or between time steps will have to rely on modeling of stochastic processes. Remmel and Csislag (2003) proposed a general framework for comparing two categorical maps, which might also be adapted to quantitative data: first, the composition and configuration of each map needs to be estimated accounting for their interdependence (Fortin et al. 2003). Replicate landscapes are simulated based only on these parameters (cf. Hargrove et al. 2002, Fortin et al. 2003) and the landscape metrics are computed for each realization to generate a confidence interval at some specified level (conditional simulation). The two patterns are considered significantly different if their confidence intervals don’t overlap. While this procedure is relatively straightforward assuming a stationary process, the ex-
tension to nonstationary processes poses a formidable challenge both conceptually and computationally (Remmel and Csillag 2003).

Quantifying local landscape structure

It is likely that different species respond to their environment at different scales and that these scales are related to the movement ranges of organisms (e.g., D’Eon et al. 2002, Holland et al. 2004). This implies that instead of analyzing global landscape patterns, one should quantify the local landscape structure across space as it may be experienced by the organism of interest (Potvin et al. 2001, McGarigal and Cushman 2005). Local versions exist for many spatial statistics (Boots 2002, 2003), but have not yet been widely adopted by ecologists (Pearson 2002).

Holland et al. (2004) provided an algorithm for identifying the scale of maximum correlation between species abundance and landscape characteristics through resampling of spatially independent observations with increasing size of the window within which the landscape metrics are calculated. Thompson and McGarigal (2002) systematically varied both the grain and extent of the environmental predictors to assess the activity-dependent scale or multiple scales of environmental response by maximizing a correlation measure.

Landscape metrics can be calculated within a specified neighborhood around each cell using a moving window (Potvin et al. 2001; FRAGSTATS, see footnote 4). Such moving window analysis provides a distribution of values for each landscape metric obtained from all possible window positions. This implies that moving window analysis may be an alternative to conditional simulation for the statistical comparison of observed landscapes (Potvin et al. 2001), but this requires the assumption that the local landscapes are true replicates with an independent history but comparable conditions, so that the ecological processes are identical.

Currently, most GIS and other software performing moving window analysis are using geometric windows (e.g., circles or squares) of arbitrary size that do not reflect the spatial structure of the species or the environment (Brashaw and Fortin 2000). Research in geographical information sciences should address this issue in order to provide tools for detecting the patchiness or zone of influence of the data (e.g., by using local spatial statistics), and implementing flexible geographical (e.g., watershed) or behavioral (e.g., home range) windows that can be adapted to a specific situation.

Merging of discrete and continuous landscape models

Landscape ecologists have been preoccupied with the patch–matrix model of discrete landscapes, which is highly compatible with the theory of island biogeography and with metapopulations (Turner et al. 2001). The gradient-based concept of landscape structure (McGarigal and Cushman 2005) is ideally suited for integrating landscape analysis with niche theory and the study of changes in ecological communities along environmental gradients, a core topic of community ecology. Multiscale ordination as discussed above is based on a formal integration of geostatistics with multivariate ordination methods, and its great potential for the empirical integration of spatial analysis and gradient analysis needs yet to be explored.

Gradient analysis in plant community ecology could profit from an explicit consideration of local heterogeneity and the organism-specific scale of response: organisms including plants are likely to respond not only to a local average of an environmental factor, but also to its variability in space and time at a scale related to the organisms size, mobility, and life span. This could be quantified by calculating the standard deviation or a local spatial statistic within a moving window of an appropriate size. However, compared to landscape metrics, these statistics for continuous variables provide rather crude measures of the spatial configuration of the environmental factor. As an equivalent to landscape metrics for continuous environmental data, McGarigal and Cushman (2005) proposed applying surface metrology metrics (Pike 2001), which are used for quantifying surface roughness in microscopy and molecular physics.

The patch–matrix and the gradient concepts of landscape structure represent two extremes of landscape structure, with most real landscapes falling somewhere in between. While the best choice will always depend on the research question, it will be increasingly important to incorporate internal heterogeneity and gradual differences between habitat types into landscape metrics as well as discontinuities into spatial statistics. Dorner et al. (2002) proposed modifications to landscape metrics so as to reflect topographic variability. When applying spatial statistics to landscapes with a discontinuous, mosaic-like structure, homogeneous areas dominated by the same stationary process can be delimited empirically using boundary detection algorithms (Fagan et al. 2003). Furthermore, ecological boundaries and ecotones determined from species data can be spatially related to environmental boundaries (Fortin et al. 2000), so that their spatial coincidence can be tested (Fortin et al. 1996) and their effects monitored.

Fuzzy set theory has been successfully applied to the problem of gradual transitions between ideal vegetation types (Roberts 1989): rather than drawing an arbitrary line for classification, a degree of membership to each type is attributed to each observation. Habitat maps could be represented in a similar way as multivariate surfaces of membership. Such an approach would not only accommodate internal heterogeneity within formerly discrete, assumedly homogeneous patches, but also retain information on mapping uncertainty, so that
its propagation through subsequent analyses could be assessed (Brown 1998, Bolliger and Mladenoff 2005).

Wavelet analysis provides a promising alternative for characterizing and partitioning landscapes in the presence of multiple, overlapping processes (i.e., not stationary), and this method can easily handle large data sets (i.e., continuous data) such as remote sensing data (Bradshaw and Spies 1992, Csillag and Kabos 2002; McGarigal and Cushman 2005). The integration of discrete and continuous landscape concepts may also profit from attempts in geography and GIS to combine discrete and continuous data models through the definition of fields of spatial objects (Cova and Goodchild 2002). Similar efforts are made towards representation of space-time data, another important shortcoming of GIS that is impeding the integration of spatial and temporal processes in ecology (Henebry and Merchant 2001, Peuquet 2001). Finally, spatiotemporal analysis of landscape dynamics could help to assess the importance of ecological memory or answer the question of how much randomness there is in real landscapes (Peterson 2002).

**Conclusion**

The basic problem of spatial analysis of landscapes is that several processes creating heterogeneity often operate at the same time. These processes may interact, so that the parameters of one process change with the heterogeneity resulting from other processes. This means that the observed pattern can rarely be attributed to a single, stationary process, as many methods of spatial analysis assume. Furthermore, most spatial processes in ecology are stochastic, so that many replicates are needed for an accurate quantification of the underlying process. However, replications are hard to obtain because the parameters of the process are likely to change through space or time due to environmental heterogeneity.

Local spatial statistics offer a way to accommodate spatial variation in pattern and even to obtain replicate landscapes at a finer scale. However, the size of such local landscapes needs to be determined in an ecologically meaningful and methodologically sound way. Statistical methods for testing hypotheses about non-stationary processes urgently need to be developed. As the hypothesis concerns the spatial process (which is not directly observable) rather than the empirical pattern, confidence intervals are best derived by conditional simulation. Local statistics and statistical tests that can accommodate nonstationarity are needed for the analysis of discrete patterns with landscape metrics as well as for the quantification of continuous surfaces with spatial statistics. However, both gradients and discontinuities are a reality in ecological systems, and we need to find ways of integrating discrete and continuous aspects of heterogeneity.

Facing these challenges may enable ecologists to go beyond quantifying patterns in order to finally address the interaction between environmental heterogeneity and the ecological processes causing species patchiness. This can only be achieved by distinguishing, both conceptually and empirically, between endogeneous autocorrelation due an ecological spatial process and exogeneous spatial dependence induced by environmental response.

**Acknowledgments**

This work was partly funded by the Swiss National Science Foundation within the NCCR Plant Survival (H. Wagner) and NSERC (M.-J. Fortin). We thank Jacqueline Bolli, Jesse Kalwij, Stephanie Melles, Bronwyn Rayfield, Christoph Scheidegger, Anurag Agrawal, Geoffrey Henebry, and Dean Urban for their valuable comments.

**Literature Cited**


information uncertainty in natural resources. Ann Arbor Press, Chelsea, Michigan, USA.


