Spatial statistics, spatial regression, and graph theory in ecology

Marie-Josée Fortin a,*, Patrick M.A. James b,c, Alistair MacKenzie a,d, Stephanie J. Melles a,e, Bronwyn Rayfield a,f

a Department of Ecology & Evolutionary Biology, University of Toronto, 25 Harbord St., Toronto (Ontario), Canada M5S 3G5
b Faculty of Forestry, University of Toronto, Toronto, ON. M5S 3B3, Canada
c Département de sciences biologiques, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, QC H3C 3J7, Canada
d Pinery Provincial Park, Ontario Parks, 9526 Lakeshore Road, RR2, Grand Bend, ON, N0M 1T0, Canada
e Department of Biology, Trent University, 2140 East Bank Drive, Peterborough, ON K9J 4B8, Canada
f Department of Biology, McGill University, 1205 Docteur Penfield, Montréal, Québec H3A 1B1, Canada

A R T I C L E   I N F O
Article history:
Received 5 December 2011
Accepted 27 February 2012
Available online 10 March 2012

Keywords:
Spatial heterogeneity
Spatial scales
Fragmentation
Connectivity
Wavelet
Species distribution

A B S T R A C T

A critical part of ecological studies is to quantify how landscape spatial heterogeneity affects species’ distributions. With advancements in remote sensing technology and GIS, we now live in a data-rich era allowing us to investigate species–environment relationships in heterogeneous landscapes at multiple spatial scales. However, the degree and type of spatial heterogeneity changes depending on the spatial scale at which species–environment relationships are analysed. Here we present the current spatial analytic methods used in ecological studies to quantify ecological spatial heterogeneity. To determine the key spatial scales at which underlying ecological processes act upon species, we recommend use of spectral decomposition techniques such as wavelet analysis or Moran’s eigenvector maps. Following this, a suite of spatial regression methods can be used to quantify the relative influence of environmental factors on species’ distributions. Finally, spatial graph metrics can be employed to quantify the effects of spatial heterogeneity on landscape connectivity across or within species’ ranges and can be used as additional predictors in spatial regression models. We emphasize how spatial statistics, spatial regression, and spatial graph theory can be used to provide insights into how land-
1. Introduction

Species at all ecological levels of organization (individual, population, community) are affected by the spatial heterogeneity of their respective landscapes (Levin, 1992; Agrawal et al., 2007; Currie, 2007; Hanski and Mononen, 2011). Spatial heterogeneity can be defined as the diversity and configuration of ecological components within a given area and is modified by natural and anthropogenic disturbances, global change (climate and land-use change; Ewers and Didham, 2006), and the interactions among them. Effects of these disturbances on species distributions can range from local population extirpation to global species extinction. For example, landscape fragmentation and the development of road networks can erode species’ habitat quality and reduce individual survival rates (Fahrig, 2003; Fischer and Lindenmayer, 2007). Land-use change often results in habitat removal and increased barriers to movement that can reduce gene flow (which in turn increases inbreeding; Cushman et al., 2006) and promotes species invasions (Real and Biek, 2007). Examination of species distribution patterns through time using modern methods of spatial analysis can help us to better understand the effects of ecological and environmental pressures.

To quantify the relative importance of natural and anthropogenic pressures at appropriate spatial and temporal scales, both spatial (Cressie, 1993; Haining, 2003; Fortin and Dale, 2005) and spatio-temporal analyses are required (Cressie and Wikle, 2011). Hence, to quantify species’ spatial distributions and relate them to environmental heterogeneity a series of statistical and mathematical techniques are needed. These techniques can be used to address specific questions such as: (1) what is the appropriate spatial scale of study, (2) what is the spatial distribution of ecological entities, (3) what are the species’ spatial responses to landscape spatial heterogeneity, and (4) what levels of connectivity exit among and within species’ distributions? Specifically, we stress here how spatial statistics, spatial regression, and spatial graph theory can be used to provide insight into the effects of landscape heterogeneity on species’ distributions to better inform conservation management and planning.

2. Spatial analysis at the landscape level

Landscapes are being altered by humans at an unprecedented rate in many different parts of the world. The resulting landscapes are increasingly spatially heterogeneous and too often highly fragmented due to high road densities and pervasive habitat loss. Such landscape change represent significant challenges to our ability to identify clear cause and effect relationships between spatial patterns and the biological processes that determine species occurrence. Two specific questions that are highly relevant to characterizing species–environment relationships are: (1) which underlying factors generated observed spatial patterns, and (2) at which spatial scales do these factors operate (Fortin and Dale, 2005)? It is important to bear in mind, however, that the type of spatial heterogeneity identified is scale dependent (Fig. 1; Dungan et al., 2002; Fortin and Dale, 2005): at small extents and fine spatial resolutions (local scale), as well as at large extents and coarse resolutions (continent scale), spatial patterns may appear to be homogeneous; whereas at intermediate spatial extents and resolutions (landscape, region), spatial heterogeneity emerges. Decisions related to sampling and analysis of data influence the spatial pattern detected (Dungan et al., 2002). Furthermore, one needs to differentiate between the finest resolution of the data (often referred to as the grain) as well as the unit size and shape at which data are sampled. For example if one is interested in the spatial pattern of tree densities, the grain corresponds to the individual trees, whereas the size of the sampling unit used to determine the density could be either 10 × 10 m or 20 × 20 m. In geostatistics the sampling unit is referred to as the support of the analysis (Cressie, 1993; Dungan et al., 2002).

When only one process acts over the study area of interest, it is assumed that spatial stationarity prevails (Cressie, 1993; Fortin and Dale, 2005). Spatial stationarity implies that the process underlying
Table 1
Generalities emerging from the spatial analysis of different spatial components.

<table>
<thead>
<tr>
<th>Spatial component</th>
<th>Methods</th>
<th>Generalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>x-y coordinates</td>
<td>Euclidean distances, point pattern analysis</td>
<td>Influence (magnitude, spatial scale) of values at various geographical distances</td>
</tr>
<tr>
<td>Spatial dependence</td>
<td>Global or local spatial autocorrelation coefficients, geostatistics</td>
<td>Influence (magnitude, spatial scale, directionality) of values at various geographical distances</td>
</tr>
<tr>
<td>spatial autocorrelation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial relationship</td>
<td>Conditional autoregressive model, simultaneous autoregressive model, geographical weighted regression</td>
<td>Influence of neighbouring values (abiotic/predictors; biotic/response variable)</td>
</tr>
<tr>
<td>Spatial legacy</td>
<td>Local spatial autocorrelation coefficients, spatial regression, spatio-temporal analysis</td>
<td>Influence of past spatial pattern on current processes and current pattern; influence of values at geographical distances (abiotic/predictors; biotic/response variable)</td>
</tr>
<tr>
<td>spatial contingency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial perception</td>
<td>Multiple-scalar analysis, spatial graph analysis, network analysis</td>
<td>Influence of intervening landscape features on animal movement/dispersal ability</td>
</tr>
<tr>
<td>Multiple spatial scales</td>
<td>Moron’s eigenvector maps, wavelet hierarchical decomposition, spatial graph analysis</td>
<td>Additive spatial scales influence current spatial pattern</td>
</tr>
<tr>
<td>Meta-organizational scales</td>
<td>Incidence function, spatial regression, gravity model</td>
<td>Influence of intervening landscape features (links) on patches (nodes)</td>
</tr>
<tr>
<td>(meta-population, meta-community)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a spatially structured pattern is constant across the entire region of interest in terms of mean, variance and directionality (Fortin and Dale, 2005); it is important to bear in mind that stationarity is a property of the process generating the pattern and not of the pattern itself. In such circumstances, there are several spatial statistics (e.g., Moran’s I, Geary’s c, semi-variance γ) that can be used to quantify and test for the significance of observed spatial structure in the response variable (Dale et al., 2002). Yet these spatial statistics measure only the statistical property of the spatial pattern in terms of the degree of spatial autocorrelation and they cannot determine whether the pattern is due to exogenous factors (such as environmental conditions), endogenous factors (such as ecological processes), or both (Wagner and Fortin, 2005). To determine the source(s), an alternative hypothesis testing approach should be used, based on appropriate and hypothesis-specific sampling designs (McIntire and Fajardo, 2009) and analysis using variance partitioning approaches (Peres-Neto and Legendre, 2010). In most landscape studies, several processes operate and interact over the study area of interest, and the resulting spatial pattern represents a combination of several pattern-generating processes in which stationarity is not likely the rule. Although it is impossible to directly test whether or not a process is spatially stationary, one can explore this assumption by calculating the mean and variance within a moving window over the study area (Corstanje et al., 2008; Haskard and Lark, 2009). If spatial variability in the mean (or variance) is identified, one could first segment the study extent into spatially homogeneous sub-regions using spatial clustering or boundary detection techniques (Fortin and Dale, 2005; Philibert et al., 2008) so that each sub-region can be analysed separately (Fig. 1; Table 1). This can however considerably reduce the number of samples per sub-region, thus affecting the power to detect trends, and further complicating subsequent analyses.

Wavelet analysis represents another tool that can be used to delimit homogeneous areas (Csillag and Kabos, 2002) or regions of relative stationarity in complex spatial or temporal contexts. Until recently, one of the main problems associated with using wavelet analysis in such a way was the lack of tests of statistical significance for assessing whether boundaries are real or simply detected by chance. Currently, restricted randomization procedures that account for the spatial structure of the data can be used to test the significance of wavelet boundaries. For example, expected wavelet spectra based on simulated bootstrapped autoregressive (e.g., autoregressive model of lag 2) and geostatistical
Fig. 1. Spatial analysis methods and corresponding spatial scales (grain and extent): local, landscape, region, continent. Several aspects need to be considered: grain size, spectral frequency of the data (high amplitude and high frequency at the local scale and low amplitude and low frequency at the continent scale); support (sampling unit size); relevant processes at each spatial extent; relevant measures at each spatial extent; and appropriate spatial analytical methods.

models (Cressie, 1993) were successfully applied to test the significance of boundaries in historical spruce budworm (Choristoneura fumiferana; SBW) defoliation in Ontario (James et al., 2010), and to identify scale-specific zones of budworm susceptibility in both north-western Ontario and northern Minnesota (James et al., 2011b).

Wavelets can also be used to describe periodicity in spatial pattern. One of the main benefits of wavelet analysis is that wavelets can summarize both global and local scale-specific attributes of data, whereas many other methods of scale-specific analysis (e.g., variography) can only identify at what scale the data are structured globally. One can determine the overall distribution of scales in the data as well as where (either in time or space) those different scales are most strongly represented, on the basis of wavelet variance (Dale and Mah, 1998). Wavelet analysis has also been used extensively to determine the spatial patterns and scales of soil properties in one and two dimensions (e.g., Lark and Webster, 2001; Hassler et al., 2011).
The significance of these scales can be then assessed though comparison of empirical data to simulated null distributions (e.g., from autoregressive models or geostatistical models). Using filtering and back-transformation techniques (Mallat, 1999), it is then possible to remove the effects of non-significant or non-interesting scales. Such “de-noising” in the scale domain can be useful for examining specific scales of interest without the potentially confounding effects of patterns at different scales (i.e., noise). Finally, as the results of a wavelet analysis include a set of potentially orthogonal coefficients for each scale of the decomposition, one can examine regression models at each separate scale (e.g., Keitt and Urban, 2005). Within a multiple-regression framework, one can then examine whether a response variable is influenced by different sets of factors at different scales. Although appealing, this approach does require that response and predictor variables are available at the same extent and support, and is generally best suited for analysis of remotely sensed spatial data in raster format.

There are cases where the resulting spatial pattern of several processes is additive (e.g., the sum of a large trend, intermediate gradients, local patchiness and some random noise) which further complicates efforts to identify clear cause and effect relationships. Additive spatial patterns can be analysed by treating space as a predictor (Dray et al., 2006). This requires the determination of the most appropriate spatial scale(s) for analysing the data. Identifying these spatial scale(s) can be achieved using a hierarchical decomposition approach based on Moran’s eigenvector maps (MEM; Dray et al., 2006; Griffith and Peres-Neto, 2006) or nested kriging (Wackernagel, 2003; Bellier et al., 2007) for sampled data, and a wavelet multi-scale decomposition for grid data (Keitt and Urban, 2005; James et al., 2011b). In either case, decomposition methods generate spatial predictors at multiple scales, and it is therefore necessary to determine which predictors to retain in subsequent analysis. Ideally, selection of spatial predictors operating at particular scales should be based on a priori hypotheses about the key processes involved (McIntire and Fajardo, 2009). An alternative solution is to use statistical approaches to select the most appropriate scales (Jombart et al., 2009) such as significance testing of wavelet decompositions using null models (Torrence and Compo, 1998; James et al., 2011b). Once key spatial predictors (and scales) are selected they can be used to model the relationship between species’ distributions and environmental variables.

3. Spatial regression

If species–environment relationships can be fully specified using a combination of spatial and environmental predictors that represent known ecological processes, then it is not necessary to worry about spatially correlated residuals (Dormann et al., 2007; Beale et al., 2010). However, it is rarely possible in ecology to have complete knowledge about the processes that cause spatial structure, particularly when biotic processes such as competition, dispersal, and year to year population fluctuations are involved. A wide variety of spatial regression methods can be used to account for the presence of significant spatial autocorrelation in the data (either in the response or explanatory variables; Haining, 2003; Fortin and Dale, 2005; Dormann et al., 2007; Beale et al., 2010; Bini et al., 2009). Spatial regression models can be classified into three main categories depending on how spatial effects are modelled (Dormann et al., 2007; Beale et al., 2010): (1) space included in covariate predictors, (2) space included in the error term, and (3) spatial effects in the response or explanatory variables are replaced by transforming the original data.

“Space” can be modelled using covariates in a variety of ways. For example, the x–y coordinates of sampling locations (or some polynomial function of x–y coordinates) can be used as predictors. Spatial weights can be derived from spatial filtering methods such as using Moran’s eigenvector maps (Dray et al., 2006), or using a spatial connectivity/weight matrix of neighbours as in autoregressive models. The effect of sampling location (e.g., x–y coordinates) or having a study extent that covers several regions (e.g., several landscapes, ecosystems) can also be modelled by adding a random effect in a generalized linear mixed model (GLMM) or generalized additive mixed model (GAMM) framework. According to Beale et al. (2010), the simultaneous autoregressive and generalized additive mixed models show an overall good performance in their ability to model spatially structured data without too much bias.

Another way to model the spatial structure of data in a spatial regression framework is to include “space” in the error term. For example, regression kriging uses generalized least squares for parameter
estimation of environmental predictors modelled as fixed linear effects and “space” as a spatially autocorrelated error term (Diggle et al., 1998; Hengl et al., 2004, 2007; Fortin and Melles, 2009; Diggle and Ribeiro, 2010). Regression kriging can model different kinds of regressions (e.g., multiple linear, in addition to Poisson, logistic regression methods) while modelling the residual as a stationary random spatial function. However, with grouped or count data, variance in the dependent variable depends on number of observations per group; so care must be taken to ensure that the statistical properties of residuals from more complex regression models meet the assumptions of kriging (i.e., stationarity; Fortin and Melles, 2009).

Simultaneous autoregressive models use a neighbourhood matrix to account for spatial and temporal lags in the error term (Haining, 2003; Fortin and Dale, 2005; Dormann et al., 2007; Beale et al., 2010). Autoregressive models (simultaneous autoregressive, conditional autoregressive, moving average, etc.) assume however that spatial structure is isotropic and stationary.

Spatial regression techniques are used because these methods account for spatially correlated data, and resulting parameter estimates will be more accurate and precise as long as estimates of spatial structure (e.g., in the variance–covariance matrix) can be accurately modelled. Prior knowledge or hypotheses about the underlying causes of autocorrelation, however, are not required to accurately model the spatial pattern in errors. Moreover, most forms of spatial regression provide a single set of regression parameters for the entire region, which are calculated by taking residual autocorrelation in the data into account.

Geographically weighted regression (GWR) is another type of spatial regression which is different from other spatial regression models because parameter estimates are allowed to vary with location. Fixed or adaptive kernels can be used to weight more adjacent locations than distant ones in parameter estimation (Fotheringham et al., 2002). The advantages of GWR are that: (1) it accounts for a relationship that varies with location, (2) it helps to visualize different sub-regions in the data, and (3) the residuals tend to be less spatially autocorrelated. One of the major drawbacks of GWR is, however, that there are as many regression models as there are sampling locations, which limits any type of generalization or prediction. Yet GWR can be used as an efficient explorative tool to determine whether or not the study area should be re-analysed by sub-regions.

4. Spatial graph theory

As human populations and their activities increase, landscapes are becoming increasingly fragmented, such that habitat patches are isolated within a mosaic of patches of various qualities (various land-cover types, patch sizes and landscape features) which may restrict species movement. In conservation (Beier et al., 2011) and reserve network selection studies (Jacobi and Jonsson, 2011), there is a lot of interest in maintaining landscape connectivity to allow individuals to move through fragmented landscapes (Fig. 2) because connectivity can be crucial to population and species persistence (Corlatti et al., 2009). Structural landscape connectivity (Fig. 2) can be assessed using spatial graph algorithms (Fall et al., 2007; Dale and Fortin, 2010) by considering the structural configuration of high quality habitat patches (nodes) and the Euclidean distances between them (links). Measures of structural connectivity based on these spatial graphs are limited to the landscape itself and may not reflect species’ movement abilities in heterogeneous and fragmented landscapes. It is therefore important to create a functional spatial graph that considers both the structure and quality of the intervening landscape between habitat patches that may impede species’ movements (Fig. 2). In such functional spatial graphs, links are weighted according to matrix quality (least-cost or resistance links; Rayfield et al., 2011) and can also have directionality (Treml et al., 2008; Dale and Fortin, 2010).

The efficacy of spatial graph analyses depends on the ascription of links and nodes (Urban and Keitt, 2001), the assignment of their functional properties (Fall et al., 2007), and the selection of graph connectivity measures (Laita et al., 2011; Rayfield et al., 2011). The identification of links between nodes can be based on the topological geometry of the nodes using many different methods including ones based on nearest neighbours, minimum spanning trees, Gabriel graphs, and Delaunay graphs (Dale and Fortin, 2010). In a spatially explicit context of a fragmented landscape, nodes are patches that have a size and shape, and links between them are spatial pathways that reflect the degree of impediment of the various land-cover types (e.g., water, urban, farmland, forest). In such
circumstances, the Delaunay tessellation can be modified to have edge-to-edge links between patches forming a minimum planar graph, MPG (Fall et al., 2007).

From an ecological perspective, links can be established on the basis of individual animal movement data that define the true functional relationships between nodes. In the absence of observed field data, one can use estimates of animal dispersal ability as a threshold distance to determine which nodes are connected and the anticipated strength of that connection. The dispersal threshold distance may be combined with different graph models to test alternative hypotheses against observed patterns of species’ movement and distributions (Dale and Fortin, 2010). Links in an MPG can represent individual gap-crossing abilities (i.e., these are inter-patch links that only traverse the matrix). As such the MPG can serve as a movement hypothesis of stepping-stone behaviour (Haddad, 2000; Boscolo et al., 2008; Lees and Peres, 2009) whereby traversing the graph requires stepping-stone paths through topological neighbours (Urban et al., 2009). The dispersal distance threshold applied for the MPG has a direct ecological interpretation as the gap-crossing ability (Bélisle and Desrochers, 2002; Richard and Armstrong, 2010). In contrast, the thresholded complete graph directly connects all pairs of nodes that are less than the specified dispersal threshold distance apart (Saura and Torné, 2009). This graph encapsulates a distance-limited movement hypothesis that does not differentiate between gap-crossing links and links that traverse a mix of habitat and non-habitat. The dispersal threshold distance for links in a complete graph is less intuitive; estimates are often based on field studies that track the distance an individual moves over a given period of time (e.g., Pascual-Hortal and Saura, 2008), but spatial heterogeneity of the underlying landscape may limit their applicability and interpretability.

Assigning functional properties to links has largely been based on least-cost modelling which depends on resistance surfaces that describe the species-specific consequences of spatial heterogeneity for movement (Rayfield et al., 2010). Novel approaches based on genetic data (Garroway et al., 2011; Angelone et al., 2011; James et al., 2011a) and translocation experiments (Desrochers et al., 2011) are improving resistance estimates. The resulting spatial graphs will provide more meaningful assessments of landscape connectivity given careful selection of graph-theoretic connectivity metrics used to analyse them (Rayfield et al., 2011; Laita et al., 2011). To help researchers determine which graph-theoretic connectivity metrics to use, Rayfield et al. (2011) proposed a
classification of these metrics according to the structural level the graph analysed (element, first-order neighbour, component, and network) and the connectivity property measured (route-specific flux, route redundancy, route vulnerability, and connected habitat area). These graph metrics can therefore quantify ecologically relevant properties of the network at multiple scales (Corlatti et al., 2009). For example, the total flux through an ecological network can be measured locally between a pair of directly connected nodes along a given link or globally between indirectly connected nodes along a path of intermediate nodes. The multi-scale nature of graph-theoretic analyses makes them particularly relevant for ecological analyses that examine the feedback between fine-scale and broad-scale patterns and processes.

5. Concluding remarks

Multiple spatial analyses are routinely used in ecology to determine how “space” affects ecological systems, species distributions, and the movement of individuals. In general, ecologists aim to determine how space matters and to make predictions regarding the consequences of spatial heterogeneity and temporal variation for ecological systems and the organisms that comprise them using methods of spatial statistical analysis. Yet it is important that researchers bear in mind the limits of the spatial analytical tools that they are using, as stressed by Fortin and Dale (2009).

Nonetheless, advances in spatial ecology are more likely to be achieved by understanding all the various ways in which “space” can affect the spatial structuring of species distributions (Table 1). Greater emphasis on scale-specific analyses (Fig. 1) will permit researchers to develop and test specific hypotheses about the effects of space on ecological systems and processes. This emphasis will require greater understanding of the tools and methods available and how to apply them properly (Table 1 Fortin and Dale, 2005; McIntire and Fajardo, 2009).

Using geo-referenced data that are becoming increasingly available, ecologists are now addressing the pressing planetary challenges that climate and land-use change impose on ecological systems. Such studies require the spatio-temporal analysis of data that differ in terms of spatial extent, grain, and support, as well as temporal resolution. This is even more pressing as land-use planning decisions need to account for the spatial–temporal nature of ecological systems. Similarly for conservation planning to mitigate climate change, it is important to optimize reserve spatio-temporal connectivity to ensure species persistence. While recent analytical advancements provide us with the spatio-temporal statistics (Cressie and Wikle, 2011) needed to tackle these studies, researchers will need to learn the assumptions and limits of these new methods and continue to develop novel conceptual approaches that integrate methods, statistical tools, and fundamental biological features of the natural world.

Acknowledgements

We thank Pilar Hernandez for her thoughtful comments. This research was supported through an NSERC Discovery Grant to MJF. We would also like to acknowledge support provided by the Aquatic Research and Development Section of the Ontario Ministry of Natural Resources.

References


