ECOGRAPHY 25: 578–600, 2002

Illustrations and guidelines for selecting statistical methods for quantifying spatial pattern in ecological data

J. N. Perry, A. M. Liebhold, M. S. Rosenberg, J. Dungan, M. Miriti, A. Jakomulska and S. Citron-Pousty

Perry, J. N., Liebhold, A. M., Rosenberg, M. S., Dungan, J., Miriti, M., Jakomulska, A. and Citron-Pousty, S. 2002. Illustrations and guidelines for selecting statistical methods for quantifying spatial pattern in ecological data. – Ecography 25: 578–600.

This paper aims to provide guidance to ecologists with limited experience in spatial analysis to help in their choice of techniques. It uses examples to compare methods of spatial analysis for ecological field data. A taxonomy of different data types is presented, including point- and area-referenced data, with and without attributes. Spatially and non-spatially explicit data are distinguished. The effects of sampling and other transformations that convert one data type to another are discussed; the possible loss of spatial information is considered.

Techniques for analyzing spatial pattern, developed in plant ecology, animal ecology, landscape ecology, geostatistics and applied statistics are reviewed briefly and their overlap in methodology and philosophy noted. The techniques are categorized according to their output and the inferences that may be drawn from them, in a discursive style without formulae. Methods are compared for four case studies with field data covering a range of types. These are: 1) percentage cover of three shrubs along a line transect; 2) locations and volume of a desert plant in a 1 ha area; 3) a remotely-sensed spectral index and elevation from 10^5 km^2 of a mountainous region; and 4) land cover from three rangeland types within 800 km^2 of a coastal region. Initial approaches utilize mapping, frequency distributions and variance-mean indices. Analysis techniques we compare include: local quadrat variance, block quadrat variance, correlograms, variograms, angular correlation, directional variograms, wavelets, SADIE, nearest neighbour methods, Ripley's $\hat{L}(t)$, and various landscape ecology metrics.

Our advice to ecologists is to use simple visualization techniques for initial analysis, and subsequently to select methods that are appropriate for the data type and that answer their specific questions of interest. It is usually prudent to employ several different techniques.

J. N. Perry (joe.perry@bbsrc.ac.uk), PIE Division, Rothamsted Experimental Station, Harpenden, Hertfordshire, U.K. AL5 2JQ. – A. M. Leibhold, Northeastern Research Station, USDA Forest Service, Morgantown, WV 26505, USA. – M. S. Rosenberg, Dept of Biology, Arizona State Univ., Tempe, AZ 85287-1501, USA. – J. Dungan, NASA Ames Research Center, Moffett Field, CA 94035-1000, USA. – M. Miriti, Dept of Ecology and Evolution, State Univ. of New York, Stony Brook, NY 11794-5245, USA. – A. Jakomulska, Remote Sensing of Environment Laboratory, Faculty of Geography and Regional Studies, Univ. of Warsaw, Warsaw PL-00-92, Poland. – S. Citron-Pousty, Dept of Ecology and Evolutionary Biology, Univ. of Connecticut, Storrs, CT 06269, USA.

Current consensus within ecology that "space is the final frontier" (Liebhold et al. 1993) is expressed by intense interest in issues of spatial scale (Wiens 1989, Dungan et al. 2002), metapopulation dynamics (Hanski and Gilpin 1997), spatio-temporal dynamics (Hassell et al. 1991), spatially-explicit modelling (Silvertown et al. 1992) and spatial synchrony (Bjørnstad et al. 1999). Temporal ecological processes have been well studied

Accepted 14 January 2002

ISSN 0906-7590

since the work of Lotka and Volterra (May 1976) was applied to insect population dynamics some sixty years ago. By contrast, progress in spatial analysis was somewhat hampered by the lack of computing resources required, and intensive work began perhaps two or three decades later.

When ecologists seek spatial pattern, evidence of spatial non-randomness, they find it is the rule rather than the exception. The natural world is indeed a patchy place (Dale 1999), as we might expect, because randomness implies the absence of behaviour, and so is unlikely a priori on evolutionary grounds (Taylor et al. 1978). Ecologists study spatial pattern to infer the existence of underlying processes, such as movement or responses to environmental heterogeneity. Spatial structure may indicate intraspecific and interspecific interactions such as competition, predation, and reproduction. Observed heterogeneity may also be driven by resource availability. Care is required in inferring causation, since many different processes may generate the same spatial pattern. Spatial pattern has implications for both the theoretical issues outlined above and applied problems such as the management of threatened species. Knowledge of spatial structure can assist in the adjustment of statistical tests and the improvement of sampling design (Legendre et al. 2002).

Although spatial analysis reveals spatial pattern, it is usually an empirical approach based on observational data, and is rarely model-based. As such, it is the forerunner and facilitator of the formation of specific testable hypotheses that may later be tackled experimentally, in manipulative studies. Only then can this process generate new ecological theory. The level at which data are analyzed depends critically on the amount of detail in the information already gathered, concerning the biology and ecology of the species studied. Models can only be formed when there is considerable detailed information about life-stages, their dynamics, demography, movement, behaviour, etc. The approach outlined here is more generic and less specific; it is targeted at situations when data are collected from initial studies, relatively early in a project. When more is known, models of spatial pattern (Keitt et al. 2002) may be used to support inference or for estimation and prediction in unsampled locations.

Methods for analyzing spatial pattern have been developed in a wide variety of disciplines. Seminal papers include Bliss (1941) in animal ecology, Watt (1947) in plant communities, Skellam (1952) in statistics, Matern (1986) in forestry and Rossi et al. (1992) in geostatistics. This wide variety has led to some overlap of methodology and philosophy (Getis 1991, Dale et al. 2002). As ecologists approach a spatial problem for the first time, they are often overwhelmed by the multiplicity of available techniques, uncertain about which set of methods to use, perplexed by the consequences of their choice of methods, and mystified as to the interpretation of seemingly conflicting results.

The purpose of this paper is to provide a framework for taking such decisions concerning analytical approaches and interpretations. This we do by comparing and contrasting many commonly-used techniques, using four case-studies with field data. We start by describing the types of spatial data that may be encountered. We then briefly describe a variety of techniques that can be used to analyze data of each form, and attempt a simple characterization in terms of their output. All of the methods we will illustrate have been described elsewhere in detail. We next use the four real data sets to illustrate the results that are obtained from each analysis, focusing on how the methods differ with regard to appropriateness for data type and information provided. There have been few previous studies that have compared and contrasted different techniques, other than the paper of Legendre and Fortin (1989), text books, and informal surveys such as that reported at: ⟨http://www.stat.wisc.edu/statistics/survey/survey.html⟩. This paper differs from those due to the greater scope of both the types of data and statistical procedures illustrated.

Types of spatial data

We summarize spatial data types according to the standard geographic system that defines points, lines, areas and volumes (Burrough and McDonnell 1998), but we restrict discussion to the more ubiquitous point- and area-referenced data. There is a natural hierarchy of information contained within data types; some may be derived from others higher in this hierarchy, through transformation or sampling.

Dimensionality of study arena

Ecological data are often collected along a one-dimensional linear transect (Fig. 1a, below), defined by the set of coordinates {**L**} that comprise it (bold type denotes a set of values represented by a vector). The arena more usually studied is two-dimensional, often rectangular, although irregular shapes (e.g. Fig. 1a, top) are common. Definition of the boundary of a two-dimensional arena requires considerably more complexity than for a transect; the descriptor set {**A**} must comprise at least three pairs of (x, y) coordinates. Rarely, the area studied may explicitly exclude some regions within its boundary (e.g. Korie et al. 2000), perhaps because some habitat within the area cannot be utilized by the species studied.

Subdivision of the study arena

The study arena may be subdivided spatially with no intrinsic loss of information into contiguous regions (e.g. Fig. 1b) such as physiographic provinces (Bailey and Ropes 1998), or defined by some qualitative factor, such as habitat type, or quantitative variable, such as tree density in four classes.

If it is not possible to record data as a complete census of the study arena, then samples may be taken over smaller proportions of it. For example, an area may be sampled by several quadrats. When subareas are defined, then the quadrats are usually placed so as to represent them. Alternatively, the quadrats may be placed at random, or in some pre-determined arrangement, often to form a rectangular grid (Fig. 1c). Occasionally, samples are taken at points. Sampling results in a loss of information compared to a complete census, and caution should be taken to ensure that the effects of this are minimized (see Dungan et al. 2002). Statistical methods must be used to relate the sample to the larger population and to make inferences.

A sampling plan may be characterized by a combination of variables termed the extent, sample unit size (known in geostatistics as the "support") and lag. The extent describes the dimensions of the study arena, and its area, A. The sample unit size is the area of the quadrat (dw in the example in Fig. 1c). The lag refers to the distance between each quadrat in a grid (l_d and l_w, in the x and y directions, respectively, in Fig. 1c). For contiguous subareas with no sampling, the lag is zero. Changes to extent, support and lag may influence the inferences drawn from analysis (Dungan et al. 2002).

Data types

We distinguish three prevalent spatial data types, defined by the topology of the entity to which the recorded information refers. These are 1) point-referenced (Fig. 1d), 2) area-referenced (Fig. 2), and 3) non-spatially referenced (attribute-only).

Point-referenced data are common in plant ecology; forms derived from it are widespread in terrestrial ecology. The simplest point-referenced data are a complete census of the individuals recorded along a transect (Fig. 1d, below). Each individual is considered identical to all others and the only information recorded is its location. This type is denoted as (**x**). An example would be the locations of a particular weed species along a field margin. For a two-dimensional arena, the point-referenced data denoted as (**x**, **y**) describes the censussed locations of all individuals within it, with reference to two coordinate dimensions (e.g. Fig. 1d, top).

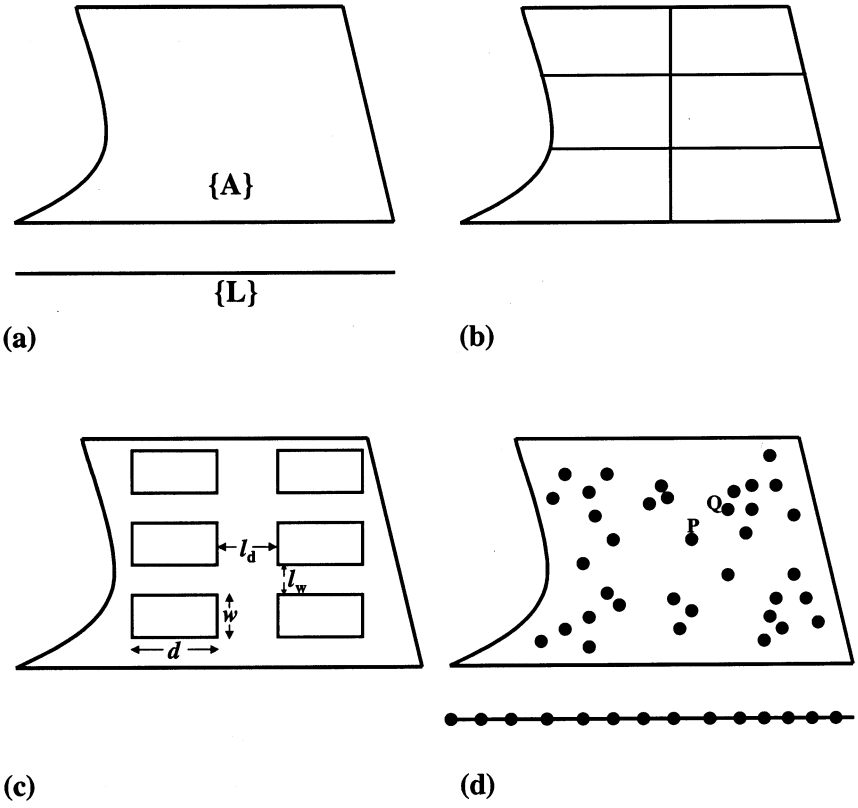

Fig. 1. (a, below) Example of one-dimensional transect with location described by the set **L**, comprising two x-coordinates; (a, top) example of two-dimensional study arena, with location described by the set of (x, y) boundary coordinates **A**, and with area denoted by A. (b) The entire study arena is divided into six contiguous subareas of unequal size and shape. (c) Six parts of the study arena are sampled with a rectangular grid of non-contiguous quadrats, each one representative of that subarea, shown in (b), in which it is located. Each quadrat is itself rectangular, with dimensions d and w, and separated from its neighbour by distances l_d and l_w, in the x and y directions, respectively. (d) Example of point-referenced data in the form of censussed, mapped individuals in the study arena (top) and transect (below) of (a).

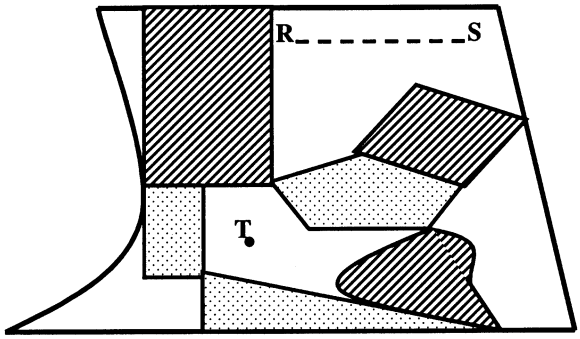

Fig. 2. Example of area-referenced geographic data, used typically in landscape ecology. The study arena of Fig. 1a has been partitioned into areas, each with one of three attributes, denoted by the shading. Areas may be polygonal or irregularly shaped with curved boundaries. Also represented are a linear feature, denoted RS, and a point, denoted T.

For any spatial data type, further information may be available for each individual, through the recording of an extra attribute(s), z; such data are denoted (**x**, **z**) or (**x**, **y**, **z**). Attributes may have different forms, of which the simplest is a categorical quantal variable (male or female, dead or alive, Fig. 3a). Another form of z attribute might be an ordered categorical qualitative variable, such as a life-stage. Alternatively, it could be a quantitative variable, such as the magnitude of an innoculum (Fig. 3b).

Many forms of data may be derived. One transformation frequently applied to two-dimensional point-referenced ecological data divides up the study arena according to the locations of individuals, into a meaningful tessellation of polygons, often named for Dirichlet, Thiessen or Voronoi (see Dale 1999: Fig. 1.4 for an example). A closely related technique is the Delaunay triangulation. These techniques leave the (**x**, **y**) coordinates of the data unchanged. They effectively create additional, derived, area-referenced data, similar to that described in section 2) below.

Point-referenced (**x**, **y**) data may be amalgamated into a single value, to represent a subarea. This results in partial loss of spatial information, since the original locations can no longer be recovered. The resulting derived data are still explicitly spatial and retain the form (**x**′, **y**′), but **x**′ and **y**′ must now be defined, for example as the centroids of the subareas, used to transform the (**x**, **y**) data of Fig. 1d to the (**x**′, **y**′, **z**′) data type in Fig. 3c. In this example, a value of z′ for each subarea has been derived from the count of the number of individuals within it. Comparing the two figures shows that almost all the information concerning clustering has been removed from the derived data. For transformations of (**x**, **y**, **z**) data, amalgamation of the **z** variable may also be achieved in very many ways, such as the mean magnitude of the z attribute, where the averaging is over the individuals located within each subarea.

Sampled data from quadrats is always derived, by definition, since all the data recorded from within the quadrat are somehow aggregated, to yield a single representative value. Usually, the derived (**x**′, **y**′) component representing the location of each quadrat is defined as its centroid, as in Fig. 3d, where the z' value for each of the quadrats in Fig. 1c has been derived from the count of the number of individuals contained within it. Observe how much variability is induced into

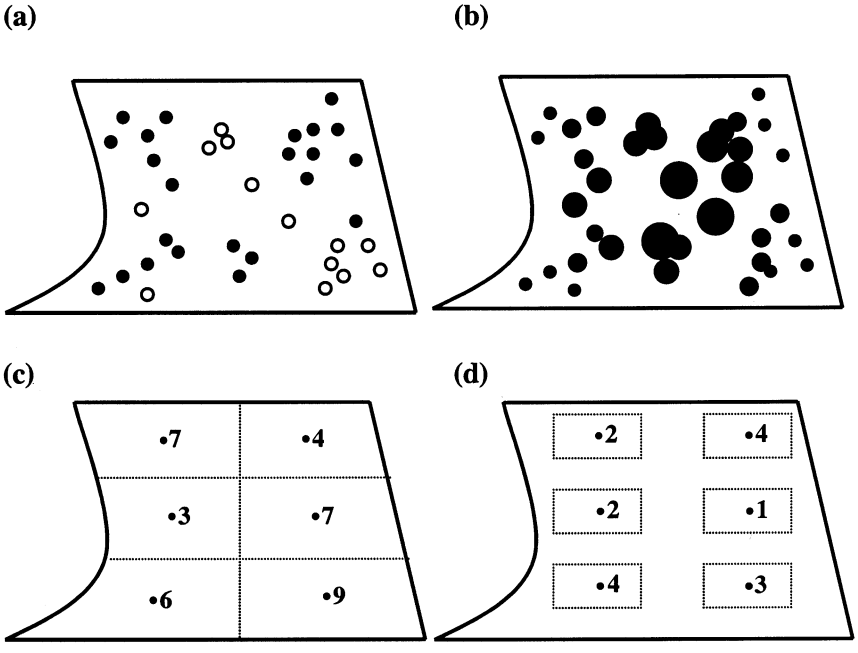

Fig. 3. Further examples of forms of point-referenced data. The censussed mapped individuals of Fig. 1d, (a) with additional quantal attribute; (b) with additional continuous attribute with magnitude indicated by size of symbol; (c) as counts within each of the six contiguous dashed subareas of Fig. 1b; (d) as counts sampled by the six dashed non-contiguous quadrats of Fig. 1c.

the counts of Fig. 3d by the sampling process, compared with the censussed values shown in Fig. 3c; whatever loss of information is involved in derived censussed data, the loss from sampled data will be greater. Also, note the difference between recording a z attribute(s) on exhaustively censussed point-referenced individuals, and taking a point sample of an attribute that is a continuously distributed variable over the study arena, such as elevation. Both may involve irregularly-spaced data of the form (**x**, **y**, **z**), but the latter has the properties of a sample, with the concomitant features of uncertainty and the intention to represent some larger unknown population.

Area-referenced data are common in landscape ecology and geography. All of the variations of point-referenced data discussed above apply equally to the area-referenced data exemplified in Fig. 2. This type is commonly represented either by a vector form, (**A**, **z**), where location coordinates defining each (possibly irregular) area are associated with attribute(s), or by a raster form, where locations are addressed by a grid of Cartesian coordinates and attributes pertain to a cell of fixed area at that location. Note that raster data can be stored as point-referenced data (**x**, **y**, **z**) but must include the crucial information of grid cell size; this representation is substantially equivalent to point-referenced data in subareas, where the loss of information is small and offset by a very large number of subareas.

In geography, a dichotomy is recognized between so-called ''object'' and ''field'' data models (Peuquet 1984, Gustafson 1998, Peuquet et al. 1999). The object model considers two-dimensional arenas populated by discrete entities whereas underlying the field model are variables assumed to vary continuously on a surface. Area-referenced data can be considered as either; within Geographic Information Systems software (Burrough and McDonnell 1998) (**A**, **z**) data are often represented as polygonal.

Sometimes, explicit spatial information might not exist, or if the data for analysis is recycled from previous results, the information may have been degraded and lack the detail of the original records. Some methods of data recording or amalgamation remove all explicit spatial information. For example, a common and powerful entity in spatial pattern analysis is termed a nearest neighbour (NN) distance (Donnelly 1978, Diggle 1983, Ripley 1988), defined for each censussed individual as the distance to its nearest neighbour. Hence, in the (**x**, **y**) data shown in Fig. 1d above, the individual labelled P has the individual labelled Q as its nearest neighbour and the distance PQ would be the NN distance associated with P. As can be seen, the relationship is not necessarily commutative, since the nearest neighbour of Q is not P. The set of NN distances for all individuals is not spatially-referenced, so is denoted by (**z**). However, it contains much implicit information concerning location, and may be used to test for spatial randomness. A further example is the frequency distribution of counts in subareas or samples (e.g. the set $\mathbf{z} = \{3, 4, 6, 7, 7, 9\}$ from Fig. 3c), stripped of its (**x**, **y**) coordinate information. From these six counts may be derived associated statistics such as the sample mean, sample variance, and quantities such as the variance to mean ratio, known as the index of dispersion (Fisher et al. 1922), I, which in this case is exactly 4. Note how I fails to capture any of the pattern in the original Fig. 1d.

Data taxonomy

Other proposals for data taxonomy have been made; some are not helpful because they are specific to disciplines tangential to mainstream ecology (e.g. Gustafson 1998). Cressie's (1991) taxonomy is not recommended because it makes insufficient distinction between the information available and the collection methods used to record it, particularly to reflect the effect of sampling. Also, it formally defines lattice data as encompassing a finite, countable set of spatial locations, while confusingly exemplifying it by reference to remote sensing and medical image data that are both spatially continuous in nature. Additionally, it would benefit from a revision to link the term ''geostatistical'' with models or methods for spatial analysis, rather than with types of data.

Techniques for spatial analysis

Characteristics of spatial pattern

Many terms are used in ecology to describe various aspects of generic non-randomness in spatial data. The terms ''aggregated'', ''patchy'', ''contagious'', ''clustered'' and ''clumped'' all refer to positive, or ''attractive'', associations between individuals in point-referenced data, such as those in Fig. 4a. The terms ''autocorrelated'', ''structured'' and ''spatial dependence'' indicate the tendency of nearby samples to have attribute values more similar than those farther apart. Conversely, the terms ''negatively autocorrelated'', ''inhibited'', ''uniform'', ''regular'' and ''even'' refer to negative or repulsive interactions between individuals, such as those in Fig. 4b. The term ''overdispersion'' has been used very confusingly in the past, to indicate excess variability or heterogeneity by statisticians and regularity of distribution by ecologists; for this reason we agree with Southwood's (1966, p. 25) recommendation that this and its antonym ''underdispersion'' be used sparingly, and then only with rigorous definitions.

The large number of methods derived for spatial analysis reflect the multitude of non-random characteristics that may be of interest to ecologists and the

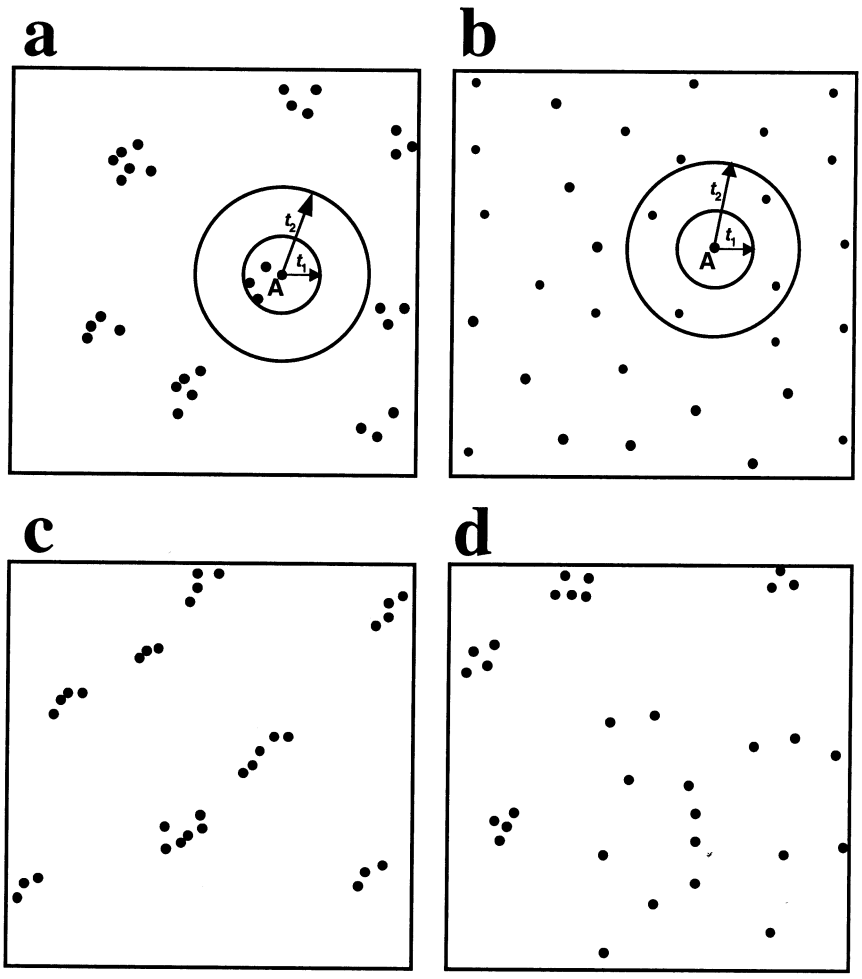

Fig. 4. Point locations of 32 individuals: (a) aggregated into 8 clusters, each with a random number of individuals, and centred on random locations; (b) arranged regularly; (c) that display spatial anisotropy; (d) that display non-stationary spatial structure, with half of the individuals aggregated into four clusters towards the top and left of the study arena and the other half, in the bottom-right of the area, arranged randomly. Note that in d the intensity of the process is roughly equal in the two parts of the area. In both a and b, one of the points has been labelled A and two circles, radii t_1 and t_2, are shown centred on A.

corresponding multiplicity of mechanisms that generate them. Most methods identify particular characteristics of pattern, ranging from generic statistical measures, such as coefficient of variation among samples, to more specialized methods that address explicit ecological questions, such as indices of landscape habitat patch shape. Occasionally, methods such as geostatistics have arisen in one discipline and proved so useful that they have been embraced by mainstream ecology (Rossi et al. 1992, Liebhold et al. 1993). We focus on twelve methods, exemplified in analyses of the case studies. Some methods, such as fractal analysis (Stoyan and Stoyan 1994) or spectral methods (Mugglestone and Renshaw 1996), are omitted due to constraints of space. Detailed descriptions of methods may be found in cited sources or in the companion paper of Dale et al. (2002).

Most methods seek to compare some feature of the spatial process "here" (at some local reference point), with the same feature "elsewhere", i.e. in another location(s). The reference point could be a randomly chosen location or individual. The feature described might be a count of a defined set of other individuals; or an attribute, for example the volume of some organism; or a derived value, such as local density. The other location(s) may be a particular distance away in some given direction; or defined by a given interval centred on the reference point; or the next contiguous area in a sequence; or the nth nearest individual. The definition "local" may shift, systematically or randomly, to encompass different parts of the study arena, in turn.

Methods vary in the degree of information they impart. The simplest methods are descriptive displays of spatial information, consistent with Chatfield's (1985) principles of initial data analysis. We stress the utility of mapping and simple summary statistics as a first step in the analysis of spatial data (Korie et al. 1998), and advocate the work of Tufte (1997) and Carr (1999) as an innovative source of ideas for visual presentation. More sophisticated methods allow inference in the form of a test of the null hypothesis, sometimes more complex than that of complete spatial randomness. Yet further information is offered by methods that estimate quantities. Finally, there are methods that fit spatial models, if enough data are available (Keitt et al. 2002).

Many ecological data sets exhibit different spatial pattern when viewed at one spatial extent than at another, a "scale" effect. For example, the point locations in Fig. 4a are aggregated into clusters at small spatial extents, although the clusters themselves have a random number of individuals and random locations. Certain methods are designed specifically to study the relationship between the degree of pattern and its scale (Table 1), and see Dungan et al. (2002).

Another common spatial characteristic of ecological data is anisotropy, where the pattern itself changes with direction. In Fig. 4c the clusters are elongate in a specific direction and tend to be associated with each other more strongly in that direction. Compare Fig. 4a, where the pattern is random with respect to direction.

Whereas global methods summarize patterns over the full extent of the area studied, in recent years there has been interest to develop methods that identify local variation (Anselin 1995, Getis and Ord 1995, Ord and Getis 1995, Sokal et al. 1998a, b) in spatial characteristics, such as aggregation, within the sampling domain (Table 1). Local methods such as wavelets and SADIE can be used to map and detect locations within the study arena that may drive the overall pattern, or which are outliers. When the underlying process that generates pattern varies across different regions within the area studied, a stationary model will not be appropriate. In Fig. 4d, although the intensity of the point process is roughly equal in the two parts of the area, individuals are aggregated in one part and randomly distributed in another. Many methods are based, to a small or large degree, on stationarity assumptions.

Methods appropriate for different data types

*Point-referenced data, (**x**, **y**)*

Questions asked of such data normally concern the spatial pattern of the individuals: are they clustered (as in Figs 1d, top and 4a) or regularly spaced (as in Figs 1d, below and 4b), or do they occur at random locations? Ripley (1977, 1988) derived a method based on the indices, $\hat{K}(t)$ and $\hat{L}(t)$, scaled for intensity, that averages the number of individuals within a distance t of a randomly chosen individual. For an individual in an aggregated pattern, for example point A in Fig. 4a, the expected number of individuals within a circle of radius t_1 that approximates to the dimensions of a typical cluster will be relatively large compared to the same number of individuals distributed randomly within the study arena. As the radius of the circle is gradually increased, say to t_2, the area encompassed extends outside the cluster to which A belongs, but not yet into other clusters, so there are few extra individuals counted. This characteristic increase in the count over particular small intervals of t provides a signature to graphs of $\hat{L}(t)$ vs t, that may be formally tested by comparison with corresponding graphs generated by Monte Carlo simulation (Diggle 1983) for random patterns. Note that for a regular pattern, as in Fig. 4b, the reverse is the case: the count is fewer than expected for t_1 and greater for t_2. Methods for this data type must allow for edge effects (Ripley 1988, Haase 1995).

*Point-referenced data with attributes, (**x**, **y**, **z**)*

Typical questions for such data are: is the apparent segregation in Fig. 3a of the solid from the open symbols real, and vice versa? Is the infection measured by the quantitative degree of innoculum shown in Fig. 3b dispersing from a point source, and if so, can we confirm that the source is located close to the centre of the study arena? With such a continuous attribute it is often of interest to determine whether its values are more similar than expected to those nearby, or more formally whether there exists detectable spatial autocorrelation, and if so, what is its relationship with the distance between points? When there are several z attributes, we are usually interested to know whether the variables are positively spatially associated (Perry 1998b, Liebhold and Sharov 1998) with each other or negatively spatially associated (termed "dissociated"), once any induced (dis)similarity due to their individual spatial structure is allowed for (Clifford et al. 1989, Bocard et al. 1992, Dutilleul 1993). Alternatively, do they occur randomly with respect to one another?

Some methods that involve **z**-attributes are restricted to regularly spaced (**x**) or (**x**, **y**) data on a grid (Table 1). For example, "quadrat variance methods" (Dale 1999) which include "two-term local quadrat variance" (TTLQV) as well as 3TLQV and others, are used predominantly for one-dimensional regularly-spaced transect data. Second-order comparisons are made between blocks of a given length, say d, of contiguous quadrats, through computation of their variance. Agglomeration into blocks using increasing values of d, within a hierarchy, allows plots of variance against d. Since the blocks are contiguous, block size is identical to the distance between block centres. The idea is that a peak in variance on this graph will indicate maximum contrast between patch and gap, if there is a stationary process with constant cluster size, at a distance equivalent to approximate cluster size. This allows the detection, and in some cases the testing for significance of spatial pattern along the transect. Their analogues in two-dimensions include the block quadrat variance methods for ($\mathbf{x}'$, $\mathbf{y}'$, $\mathbf{z}'$) data where $\mathbf{z}'$ is a count, derived for insects by Bliss (1941), rediscovered for plants by Greig-Smith (1952) and later modifications such as 4TLQV (Dale 1999).

Very closely related to these is a large class of methods that focus on the dependence of autocorrelation on distance. Here, blocks of quadrats are replaced by single units or individual locations. The average variance between a pair of units is calculated for a

Table 1. Description of methods for analysis of spatial data.

Method	Type of data	Original use	Model based?	Allows hypothesis tests?	Info. available at multiple scales?	Info. available on anisotropy?	Info. available on local pattern?	1- or 2-dimensions?	Irregularly-spaced units allowed?
Ripley's K and L, etc.	point-referenced locations, (x, y)	plant ecology	no	yes	yes	possible	no	both	yes
Quadrat variance methods (TTLQV, etc.)	point-referenced locations, (x, z)	plant ecology	no	rarely	yes	no	no	1	no
Block quadrat variance methods (Greig-Smith, 4TLQV, etc.)	point-referenced locations with attributes, (x, y, z)	plant ecology	no	rarely	yes	possible	no	2	no
Correlograms (Moran's I, Geary's c, etc.)	point-referenced locations with attributes, (x, y, z)	geography	no	yes	yes	yes	no	both	yes
Geostatistics (variograms), PQV	(x, y, z)	Earth sciences	no	no	yes	yes	no	both	yes
Geostatistics (kriging)	(x, y, z)	Earth sciences	yes	yes	yes	yes	no	both	yes
Angular correlation	(x, y, z)	geography	no	yes	no	yes	no	2	yes
Wavelets	(x, y, z)	statistics	no	no	yes	no	yes	both	no
SADIE	(x, y, z)	insect ecology	no	yes	no	no	yes	both	yes
Landscape ecology metrics (edge density, shape indices, etc.)	area-referenced locations with attributes, (A, z)	landscape ecology	rarely	rarely	no	rarely	no	2	yes
Variance-mean methods (Morisita, Taylor, etc.)	non-spatially referenced data, attributes only, (z)	applied entomology	no	yes	no	no	no	both	yes
Nearest neighbour methods	(z)	plant ecology and forestry	no	yes	no	no	no	both	yes

particular distance, d, and the relationship between variance (or semi-variance) and d is graphed and studied, thereby quantifying structure at multiple spatial extents. Some of these methods, Moran's I (Moran 1950) and Geary's c, allow significance tests of complete spatial randomness (Cliff and Ord 1973, 1981, Sokal and Oden 1978, Oden 1984). Geostatistical methods in this class include the variogram, correlogram, and covariance function. We include paired quadrat variance (PQV) in this section because, although strictly a quadrat variance method, the size of its block never varies; indeed, despite their independent development, the variogram and PQV are mathematically identical and thus provide the same information (ver Hoef et al. 1993, Dale and Mah 1998, Dale et al. 2002). Geostatistical methods are not generally used for hypothesis testing, but their associated models of spatial dependence are used for spatial interpolation, modelling and simulation (Isaaks and Srivastava 1989, Rossi et al. 1992). The idea underlying all the methods is that spatial autocorrelation declines, and therefore variance increases, with increasing d, until some maximum variance, termed the "sill", is reached, at a value of d termed the "range". The range is an estimate of average patch and gap size. For very small d the variance, termed the "nugget", although a minimum, may still be non-zero; it is equated to measurement error plus the variance at smaller unit sizes. When **z** is a dummy (0, 1) variable, measuring presence, some methods in this section (such as indicator variograms) are effectively utilizing (**x**, **y**) data.

Many methods have been adapted for the study of anisotropy (Oden and Sokal 1986, Isaaks and Srivastava 1989, Falsetti and Sokal 1993) and local information (Anselin 1995, Getis and Ord 1995, Ord and Getis 1995, Sokal et al. 1998a, b). Specific methods (Table 1) to detect anisotropy include Rosenberg (2000) and an angular method using spatial correlation, derived by Simon (1997). The latter projects each (x, y, z) point onto an axis in a test direction, say θ, then calculates the correlation, r, between the positions of the projected points along this θ-axis and the z-attribute. The values of (θ, r) are then plotted in polar coordinates; a bulge in the resulting curve indicates the direction of greatest change in z.

Wavelet analysis may also be used to quantify spatial pattern over a variety of spatial extents, usually for one-dimensional data. It has much in common with the methods described above, albeit based on a different statistical approach (Bradshaw and Spies 1992, Dale and Mah 1998). The emphasis is on the "decomposition" of the data into repeating patterns that are compared with the wavelet function's shape over varying window-widths, which are equivalent to distance lags. In this there are similarities to the quadrat variance methods (Dale et al. 2002), but wavelets allow contrasts of a specific and more complex functional form, using shapes such as the "Mexican hat" (Dale 1999). A powerful feature of this method is the lack of any stationarity assumption.

SADIE is a class of methods designed to detect spatial pattern in the form of clusters, either of patches or gaps (Perry et al. 1999). The calculations (Dale et al. 2002) also involve comparisons of local density with those elsewhere, but made across the whole study arena simultaneously. Each sample unit is ascribed an index of clustering, and the overall degree of clustering into patches and gaps is assessed by a randomization test. A specific extension to spatial association (Winder et al. 2001) is made by comparing the clustering indices of two sets of data across the sample units. A local index of association, χ_k, may be derived at each sample unit, k, and these may be combined to give an overall value, X. The power of these methods comes from the ability to describe and map local variation of spatial pattern and association.

*Area-referenced data (**A**), (**A**, **z**)*

Data that are area-referenced are common in landscape ecology, particularly in polygonal form. The methodology for their analysis has not advanced as fast as other techniques described here. Only few involve inferential tests; they are largely descriptive. As for point data, one option is to sample by overlaying a grid and convert to (**x**, **y**, **z**) spatially-referenced attribute data (see above and Discussion). However, if the degree of fragmentation is of particular interest then methods such as those described in Turner and Gardner (1991) and McGarigal and Marks (1995) may be applied to polygonal data in either raster or vector form. Patch size and its coefficient of variation (CV) are examples of the numerous indices; there is no space here for an exhaustive list. The latter is a measure of landscape fragmentation and heterogeneity but with important limitations (McGarigal and Marks 1995). Edge density (Morgan and Gates 1982), defined as a measure of edge length standardized by enclosed area, is positively related to the degree of fragmentation of the habitat, but also to its shape. For regular shapes it is minimum for a circle, and maximal for a linear feature. Total edge length is the sum of all lengths of edges of patches of one landscape type. Mean shape index, the ratio of perimeter to area, is a measure of shape complexity; it approaches unity for perfect figures such as circles and increases with shape irregularity (McGarigal and Marks 1995). When sampling relatively small areas, the area-weighted shape index is considered more meaningful, because it gives greater weight to large polygons.

*Attribute-only data (**z**)*

NN methods (Ripley 1977, 1988, Diggle 1983) are used to analyze the distance from each of the individuals in a set of point-referenced data to its pth nearest neigh-

bour, where p is usually unity. The set of distances is not spatially explicit, even though it is derived from point-referenced locations. Hypothesis tests (Clark and Evans 1954) are usually based on the expected distribution of such distances for a random arrangement of individuals. Highly clustered patterns tend to have relatively small (e.g. Fig. 4a), and regular patterns (e.g. Fig. 4b) relatively large NN distances, respectively. They are similar to methods like Ripley's $\hat{L}(t)$ function in that they must be adjusted to allow for edge effects, but differ from them in that they utilize less spatial information and cannot identify pattern at multiple spatial extents.

Early studies of spatial pattern such as Bliss (1941) were based on summary statistics from frequency distributions (David and Moore 1954), and used little or no spatially-explicit data. Techniques relied on the fact that samples of randomly arranged individuals would yield counts that followed the Poisson distribution, but an observed Poisson distribution does not necessarily imply randomness (Hurlbert 1990). Indeed, the set of counts: {0, 0, 1, 1, 2, 2, 2, 2, 3, 3, 5}, conforms closely to a Poisson distribution, but if sampled in that order along a line transect shows an obvious linear trend departing strongly from randomness. Recent methods (Upton and Fingleton 1985, Perry and Woiwod 1992) include the index of dispersion I, Morista's index, Lloyd's (1967) index of mean crowding and Taylor's power law (Taylor et al. 1978), used extensively in animal ecology. Jumars et al. (1977) and Perry (1998a) have cautioned users to distinguish non-randomness in the form of statistical variance-heterogeneity from true spatial non-randomness. The only valid spatial inference possible to make from an observation of variance-heterogeneity is that there must be spatial pattern for some unknown support smaller than that used to derive the observed counts.

Origins of methods

While it can be shown that many of the methods described above are computationally similar (Dale et al. 2002), it is worth noting that many differences among the methods can be attributed to their historical development in different disciplines (Table 1). In applied entomology, limited computational resources before 1970 precluded the widespread development of methods that utilized specific spatial coordinates, although the SADIE method now exists to address this issue. Later, methods developed for plant ecology were applied to two-dimensional coordinate locations of individuals, but often lacked a tractable underlying statistical model. Geostatistical methods were developed for problems in the applied Earth sciences and emphasized estimation rather than the detection of spatial pattern by hypothesis testing. However, their development from an underlying generic statistical model has made these methods useful for purposes of spatial interpolation and simulation, and they are applied increasingly to ecological problems (Rossi et al. 1992, Liebhold et al. 1993).

Case studies

Four case studies are used to illustrate the application of the methods outlined above to the data types discussed above. These exemplify comparisons between methods, regarding their applicability to answer particular questions, their ability to provide different types of information and to allow contrasting inferences.

Case study 1: counts of three shrub species along a transect

The data (Dale and Zbigniewicz 1997) are censussed percentage cover of each of six boreal shrub species, of which only three are considered here. They are, in order of abundance, *Betula glandulosa*, *Salix glauca*, and *Picea glauca*, measured between 1 and 3 m height. Cover was measured, along a linear transect of 1001 contiguous quadrats, each 0.1 m^2, in shrub-dominated vegetation. Separate measurements were done for each species; the combined total for more than one species may be $> 100\%$. All three species were most abundant towards the left end of the transect and all were absent in $> 60\%$ of quadrats. Questions Dale and Zbigniewicz posed included: is there pattern of plants within species? What is the average patch- and gap-size? Is there spatial association between species? The (**x**, **y**, **z**) data are analyzed as collected: derived total cover per contiguous quadrat.

Betula glandulosa

The observed, raw data exhibited numerous, relatively small patches (Fig. 5). Initial inspection of presence-absence revealed 80 runs of quadrats that contained some *Betula glandulosa*, with mean length 0.48 m, separated by 80 runs that contained none, with a much greater mean length of 0.77 m. Indeed, for such sparse data, with $> 60\%$ zero cover, there must be large differences between patch and gap size.

The analysis begins with various approaches to estimate a dominant cluster size. Because of the large number of quadrats, the variogram (Fig. 6) was smooth and the lack of a nugget effect emphasized that most of the fine spatial structure was captured. A spherical model was appropriate and the range of ca 0.7 m implied a relatively small dominant cluster size.

The PQV plot, identical to the variogram except that it was plotted over a greater distance on the x-axis,

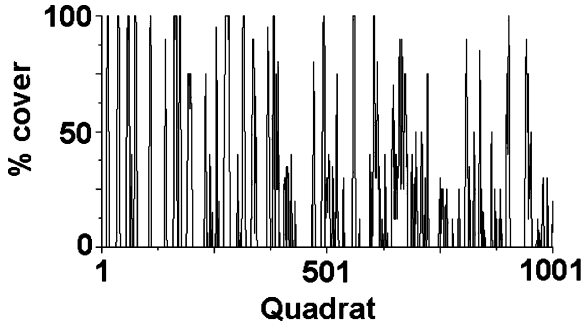

Fig. 5. Percent cover of *Betula glandulosa*. One quadrat is equivalent to 0.1 m.

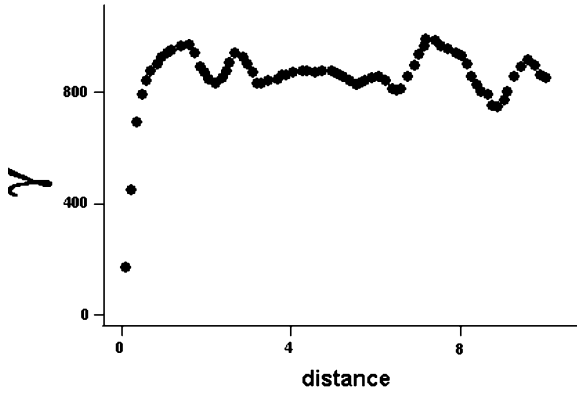

Fig. 6. Variogram of *Betula glandulosa* data in Fig. 5. Semivariance, gamma, is plotted against distance in metres.

added no useful information for these data; there was no evidence of damped periodicity beyond the estimated range. The TTLQV and 3TLQV plots were similar. The 3TLQV plot, using more data, had three peaks: one from ca 1.7 to 2.5 m, another from 4 to 7 m, and a third from 13 to 17 m. It was unclear to what features of the data the larger peaks corresponded.

The Mexican-Hat wavelet analysis (Fig. 7) showed by the intense yellow shading at small window-widths and by the marginal mean graph, that the greatest contrasts within windows occur at relatively small window widths of 10 quadrats, the estimated dominant cluster size. The spike of largest variance, occurring around quadrat 277, successfully identified the longest run of substantial cover, ten successive quadrats each of at least 90%. Another region in which the wavelet analysis identified large variance was the broad peak around quadrat 660; there, *B. glandulosa* was present in 34 successive quadrats, but with considerable variability in cover from quadrat to quadrat.

Moran's I was highly significant overall. At the shortest distances, I declined with increasing lag, until at a distance of 0.7 m spatial autocorrelation became negative; this was interpreted as an estimate of the dominant cluster size. At larger lag distances I had relatively small values. The strongly negative autocorrelations at lags 7.5 and 80 m were difficult to relate to the data and could not readily be interpreted.

The SADIE analysis showed very strong clustering. The mean patch and gap clustering indices were 3.12 ($p < 0.0005$) and -3.30 ($p < 0.0005$), respectively. Just over one-third of the total length of the transect had no detectable pattern (white areas, Fig. 8); these areas were spread fairly evenly. The size of the 73 patches ranged from 0.1 to 1.7 m (mean 0.40 m, SEM 0.046 m, median 0.2 m), and of the 106 gaps from 0.1 to 3.3 m (mean 0.30 m, SEM 0.031 m, median 0.2 m). The SADIE clusters relate to spatial pattern rather than abundance per se, and so the large proportion of the transect where there was no pattern broke up the runs of less-abundant quadrats into the relatively large number of gap clusters.

Of the methods studied, the variogram/PQV and Moran's I matched each other closely and gave sensible estimates of dominant cluster size, averaged of necessity over patches and gaps. The Mexican-Hat wavelet

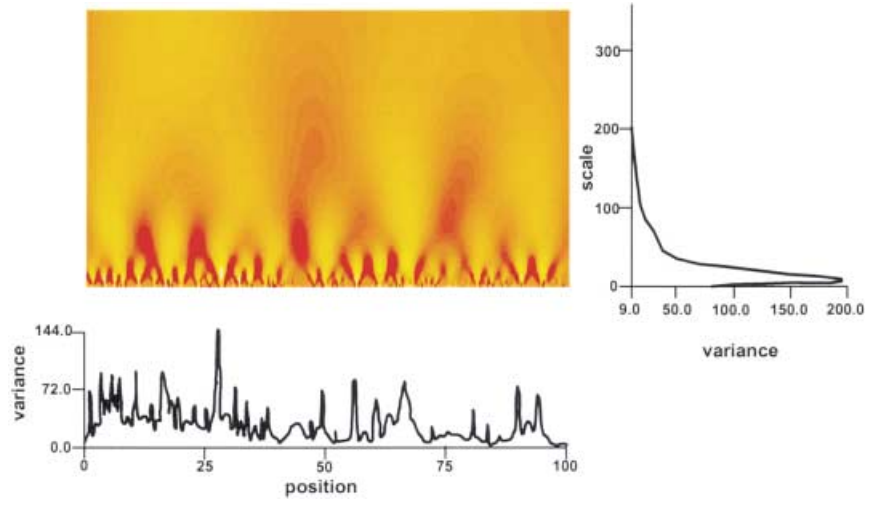

Fig. 7. Mexican-Hat wavelet analysis of *Betula glandulosa*. The variance surface is plotted vs position along transect in metres and scale of window-width, in units of quadrats (0.1 m); larger variances are shown in yellow and smaller variances shaded red. Also shown are graphs of the two marginal means over the surface. In the lower graph, variance is averaged over all window widths for each quadrat; to the side, variance is averaged over the entire transect for each window-width scale.

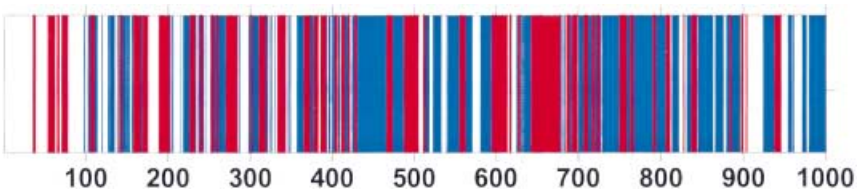

Fig. 8. SADIE analysis of *Betula glandulosa*. Red shading indicates patches with cluster index values > 1.5; blue shading indicates gaps with index values < -1.5; white areas represent lengths that are neither patches nor gaps.

provided a larger estimate, but gave very useful insights into other aspects of the data. TTLQV and 3TLQV performed poorly, and added no extra useful interpretation. The SADIE method provided separate and different estimates of patch and gap size, which were sensible.

Picea glauca

The observed *Picea glauca* data exhibited a small number of relatively small patches (Fig. 9a), having similar structure to the *Betula* data above, but of a much sparser nature. There were 20 small bursts of cover, ranging from 0.1 to 1 m with mean length 0.24 m. Less than 5% of quadrats were occupied and the runs containing no *Picea* ranged up to 18.1 m with mean length 4.54 m.

The PQV (Fig. 10, top) implied an average cluster size of 1.1 m, although for situations where the observed patch and gap size are as disparate as here this has little interpretive utility. Its interesting feature was a sudden decline in variance at around 10.5 m, unusual for a PQV/variogram, and caused by the sparseness of the data. This distance was highlighted because it represented roughly the minimum between any of the four patches containing at least one quadrat with $> 80\%$ cover shown in Fig. 9a. At shorter distances, variance was dominated by the comparison between those pairs of quadrats, one of which contained a member of one of these four patches and the other of which did not. The upward slope over long stretches of the plot indicated the clear trend in the data. The TTLQV and 3TLQV plots showed no sudden decline in variance; they indicated structure with an average cluster size ca 6 m and ca 3.8 m, respectively.

Marginal variance in the Mexican-Hat wavelet plot mirrored the positions of the four largest patches referred to above, but showed little further structure; marginal variance was maximal at 0.9 m.

Usually, Geary's c and Moran's I (Fig. 10, middle and below) are very similar except for inversion, but for these data they differed. Both indicated an dominant average cluster size of 0.9 m, but whereas I remained fairly flat thereafter, c continued to increase before crossing the x-axis again around 10.5 m, informatively indicating a second scale of pattern at the same distance as found by PQV (Fig. 10, top).

The SADIE analysis confirmed substantial spatial pattern, the mean clustering indices for patches being 3.25 ($p < 0.0013$) and for gaps -3.88 ($p < 0.0013$). A proportion of 0.29 of the total length of the transect had no detectable pattern, almost all within the left-hand half of the transect. The 19 patches were correctly identified, ranging in size from 0.1 to 0.5 m (mean 0.18 m, SEM 0.031 m, median 0.1 m), and the 43 gaps from 0.1 to 12.7 m (mean 1.6 m, SEM 0.36 m, median 0.8 m).

Salix glauca

Data for *Salix glauca* was similar in structure to both other species. There was considerable spatial pattern, with a degree of sparseness midway between the other two species (Fig. 9b). No feature of any analyses was fundamentally different from those discussed above, so none is reported, but the data are presented to inform

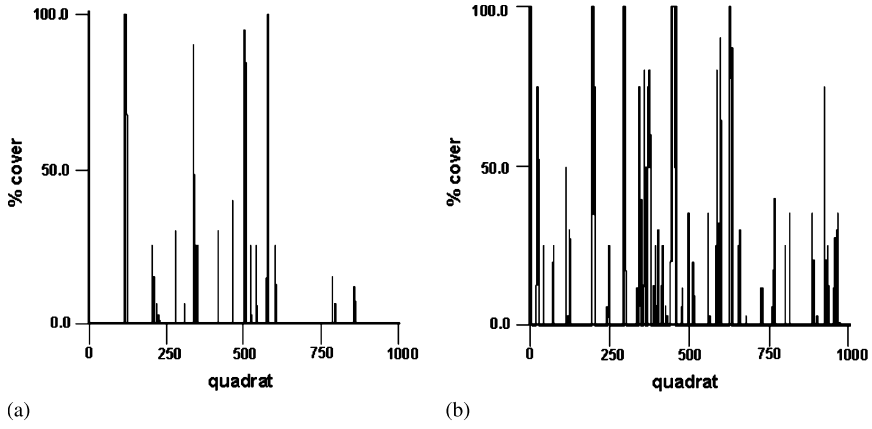

Fig. 9. Percent cover of (a) *Picea glauca* and (b) *Salix glauca*. One quadrat is equivalent to 0.1 m.

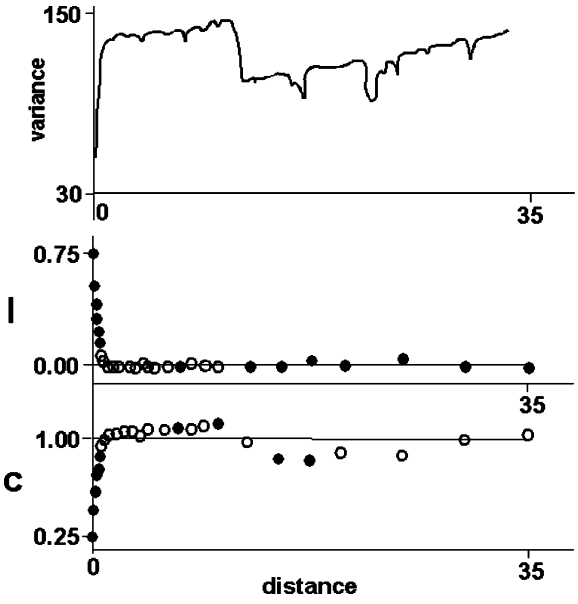

Fig. 10. Methods that measure variability plotted against varying extents in units of metres (ten quadrats) for *Picea glauca*. Above is local paired quadrat variance, PQV; underneath are two correlograms, Moran's I (middle graph) and Geary's c (lower graph). Filled symbols indicate significant individual lags ($p < 0.05$).

the analysis of relationships between the species, discussed below.

Relationships between the species

Using the raw cover data, correlations between species pairs were small (*Betula glandulosa* and *Picea glauca*, $r = 0.0188$; *Betula glandulosa* and *Salix glauca*, $r = -0.0577$; *Picea glauca* and *Salix glauca*, $r = 0.0016$). None was significant, before or after accounting for spatial structure using the Clifford et al. (1989) method. Cross-correlograms and cross-variograms also revealed little structure.

However, SADIE analysis of the cluster indices revealed strongly significant positive association between the spatial pattern in pairs of all three species (*Betula glandulosa* and *Picea glauca*, $X = 0.244$, $p < 0.001$; *Betula glandulosa* and *Salix glauca*, $X = 0.133$, $p = 0.001$; *Picea glauca* and *Salix glauca*, $X = 0.401$, $p < 0.001$). The method of Clifford et al. (1989) suggested an approximate halving of the effective degrees of freedom for each species pair; probability levels and confidence limits, from randomizations under the null hypothesis of no association, were adjusted accordingly. In a map of local association, χ_k, for *Picea glauca* and *Salix glauca* (Fig. 11), the dominance of plum shading over green demonstrates an overall positive association between the species. The positions of the shaded contours distinguishes areas in which relatively large local values occurred. For example, the plum shading between quadrats 333 and 336 reflects values of positive local association, arising from the coincidence of patches exceeding 40% cover for both species. Detrending the cluster indices by a quadratic function suggested that the association found was due mainly to larger-scale similarities between the species. All maps showed considerable clustering of local association towards the extreme right of the transect, where the coincidences of several gaps reflected long runs of quadrats where cover was exceedingly sparse or absent for all species.

The difference between the results for the simple correlation analyses and SADIE arise, as for the single species comparisons, because the former makes comparisons between variables that are abundance/density estimates, where an isolated large value contributes equally to the computed statistic as does a similar value in a patch. By contrast, the SADIE analysis is based upon the degree of spatial pattern, so isolated values are deliberately downweighted.

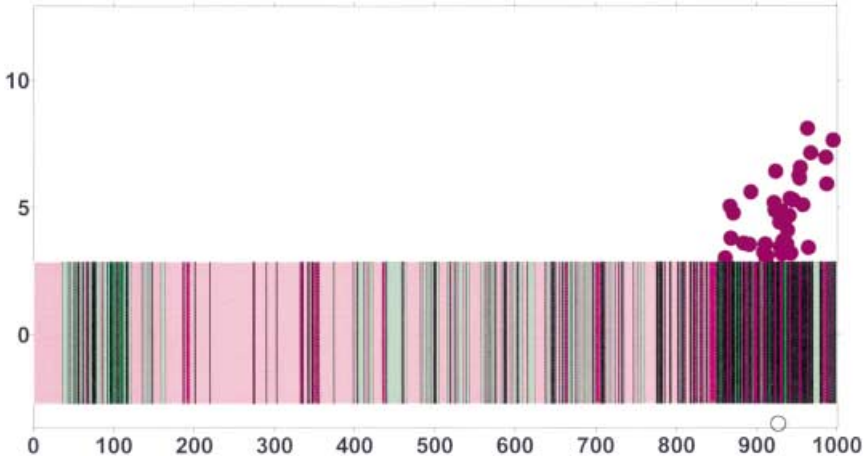

Fig. 11. SADIE local association (y-axis), χ_k, between *Picea glauca* and *Salix glauca* versus position of kth quadrat (x-axis). Symbols denote values of χ_k exceeding upper or lower critical values (25 expected); 35 filled plum circles indicate significant positive association, single open green symbol indicates negative dissociation. Variation of local association along transect is shown, for all values of χ_k, by shaded bands of colour in rectangle (plum positive; green negative; darker shades indicate greater extremes of association).

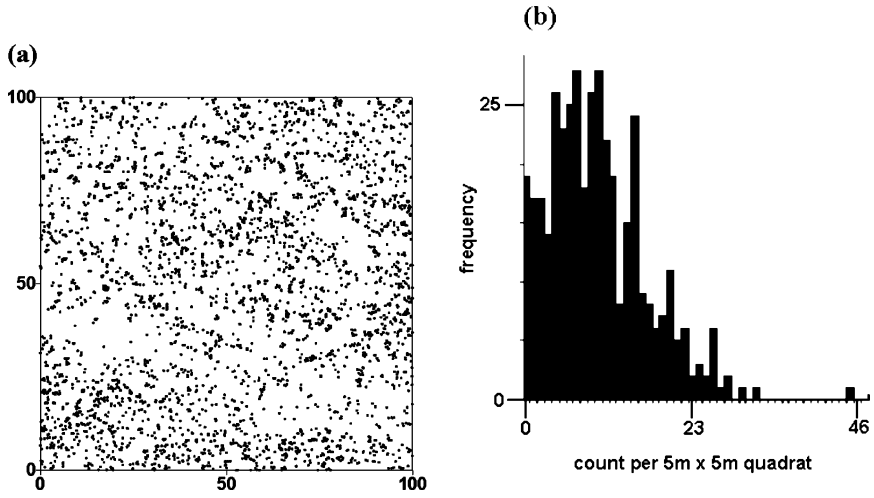

Fig. 12. (a) Locations of the 4357 *Ambrosia dumosa* individuals in the 100×100 m study arena. (b) Frequency distribution of the count of individuals per 5×5 m subarea quadrat.

Case study 2: mapped locations and volume of a desert shrub

The data (Miriti et al. 1998) are the locations and logarithmically-transformed estimated plant canopy volumes of all individual adults of the deciduous desert shrub *Ambrosia dumosa*, recorded on a single occasion in a 100×100 m area in the Colorado Desert (Wright and Howe 1987). These data form a subset of a larger and phased study of plant dynamics with different life stages, in which *Ambrosia dumosa* encompassed almost two-thirds of all recorded stems. The study site was selected deliberately to minimize heterogeneity attributable to environmental variation. Miriti et al. (1998), and see references within) were interested in the relative importance and effects of possible intra-specific interference and negative plant-to-plant interactions on spatial distribution, and to quantify the spatial scales over which major processes operated. Here, four versions of the data were analyzed: (**x**, **y**) point locations, derived (**x**, **y**, **z**) counts in 5×5 m contiguous subareas, (**z**) attributes (nearest neighbour distances and volumes), and derived (**x**, **y**, **z**) mean volume per plant in 5×5 m contiguous subareas.

Individuals of *A. dumosa* seemed to be distributed throughout the study arena with considerable, but small-scale aggregation (Fig. 12a). Two or three bands of relatively low or zero density, of width ca 15 m, appeared to run roughly north-west to south-east across the area. The frequency distribution of counts per 5×5 m quadrat shows a typically right-skewed distribution with a mean of 10.86 plants per quadrat and a similar mode (Fig. 12b); the variance/mean ratio was 4.3, which indicates significant numerical variance heterogeneity.

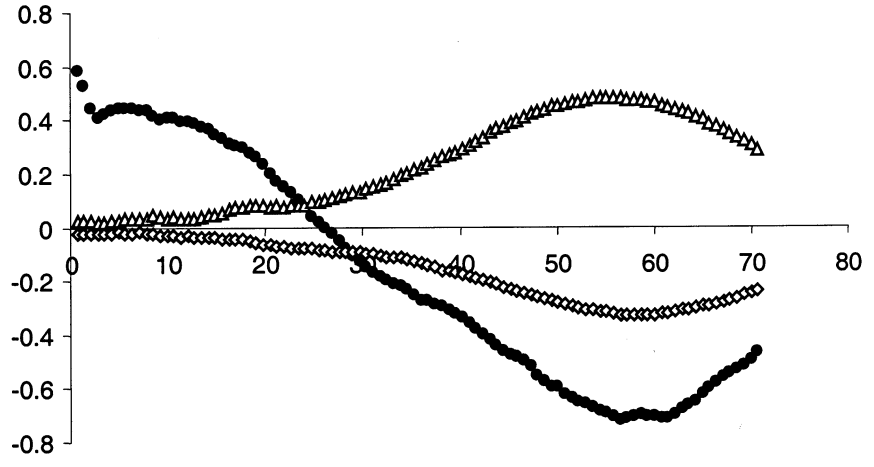

Fig. 13. Ripley's $\hat{L}(t)$ function for *Ambrosia dumosa* individuals, where t represents the radius in metres of the notional circle drawn around a randomly chosen plant. Also shown are upper, $R_u(t)$ and lower, $R_l(t)$ envelopes representing upper 97.5%-iles and lower 2.5%-iles under the null hypothesis of complete spatial randomness, derived from Monte Carlo randomizations. Here, for visual clarity, the three y-variables are each transformed by subtracting t, prior to plotting: $\hat{L}(t) - t$ are filled circles, $R_u(t) - t$ are open triangles, $R_l(t) - t$ are open diamonds. The x-axis represents t. Filled circles above the upper envelope indicate significant aggregation and those below the lower envelope indicate significant regularity.

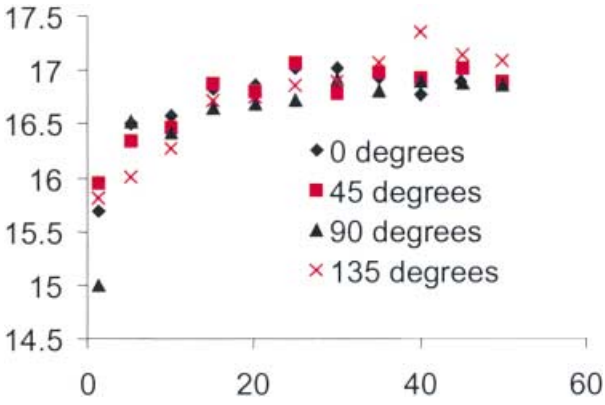

Fig. 14. Semivariance from directional variograms for *Ambrosia dumosa* counts, in 5 × 5 m quadrats, plotted against distance in metres, in directions: 0°, 45°, 90° and 135°.

Previous analyses by Miriti et al. (1998) showed that the total nearest neighbour distance was 386.0, which exceeded the expectation for a random distribution, using Donnelly's (1978) method, by over 15 times the standard deviation, indicating highly significant aggregation. They also used a hierarchical blocked quadrat variance method (Dale 1999), based on Morisita's (1959) index (the test for which is equivalent to Greig-Smith's (1952) test of the index of dispersion I), using as data the counts of individuals within contiguous 2^n m^2 square subareas, where n = 2, 3, and upwards. This gave an initial, rough indication of the relationship of how aggregation varied with subarea size. It showed moderate and non-significant heterogeneity, but which crucially decreased monotonically with subarea from the smallest, with side 2 m, and became virtually indistinguishable from random at the largest tested, with side 11.3 m. Hence, both these results confirmed the conclusions of Miriti et al. (1998) and the visual impression from Fig. 12a. There was considerable aggregation at the smallest spatial extents, consistent with some unspecified plant-to-plant interactions, but this seemed to decrease over medium extents of ca 100 m^2 with sides of 10 m, a distance beyond which an individual plant would expect to exert an influence over others.

To quantify further the relationship between aggregation and spatial scale of distance, t, we plotted Ripley's $\hat{L}(t)$ function versus t for the (**x**, **y**) point location data (Fig. 13). This confirmed one indication from the blocked quadrat variance analysis, that there was significant and strong aggregation from the smallest scale studied, a distance of 0.7 m, which declined almost monotonically with t. However, this did not cross the upper envelope until ca 24 m. Unusually, $\hat{L}(t)$ continued to decline, indicating regularity by falling substantially below the lower envelope, and not until 60 m did

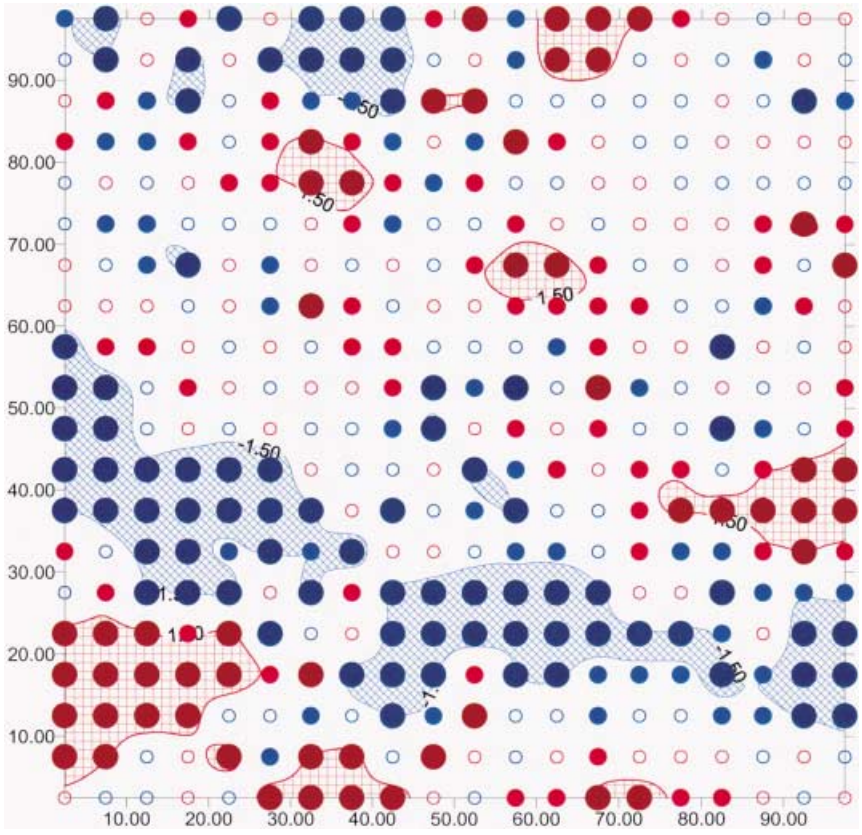

Fig. 15. Overlaid contour and classed post maps of SADIE clustering indices for counts of *Ambrosia dumosa*, in 5 × 5 m quadrats. Red shading and darker, larger filled red circles indicate strong patchiness with index values > 1.5; blue shading and darker, larger filled blue circles indicate strong gaps with index values < −1.5. Medium-sized filled circles: units with clustering that exceeds expectation (> 1 or < −1). Open circles: clustering below expectation (< 1 or > −1).

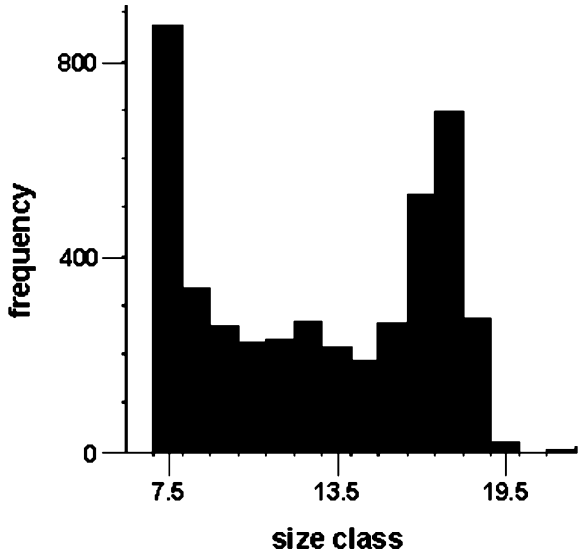

Fig. 16. Frequency distribution of *Ambrosia dumosa* logarithmically-transformed plant canopy volume per plant.

it increase again relative to the randomizations. This is not thought to reflect an important facet of the data. It was probably a manifestation of the change in intensity along the diagonal that runs from lower-left to upper-right of the area, which appears to cycle 2.5 times as it traverses the two large relatively empty bands referred to above. The wavelength of this cycle is ca $(40\sqrt{2})$ m = 57 m which coincides well with the value of t for which $\hat{L}(t)$ is a minimum. The abundance of individuals ensures that there is considerable autocorrelation in the plot of $\hat{L}(t)$ vs t. Hence, it would not only be unwise to infer that true regularity existed at the scale of 60 m, but also to infer that the distance at which the spatial process became random was 24 m, or indeed to attempt a precise estimate of this distance. In summary, the most important information conveyed by this analysis relates to the existence of the strongest pattern at the smallest distances.

The following analyses were done for the counts of individuals in 5×5 m contiguous subareas. Directional variograms were calculated (Fig. 14). The large nugget variance and shallow slope up to the sill supported the conclusion above, that the major structure in the data occurred at distances smaller than were resolvable by the subarea size. The variogram range also confirmed the indication, from $\hat{L}(t)$, that there was some larger-scale structure up to distances of ca 25 m. This distance was confirmed by the results from both Moran's I and Geary's c. Additionally, the variograms showed some mild anisotropy in the 135° direction for distances > 40 m, providing further support for the existence of the bands referred to above, although no directionality was detected by an angular correlation analysis.

The SADIE analysis confirmed the presence of large-scale patches of varying size up to ca 400 m^2 (mean patch cluster index = 1.29; p = 0.025), and gaps (mean gap cluster index = -1.35; p = 0.048). The most notable feature was a long gap ca 10 m wide stretching from left to right across almost the entire width of the hectare block (Fig. 15), and supporting further the existence of the bands of low density.

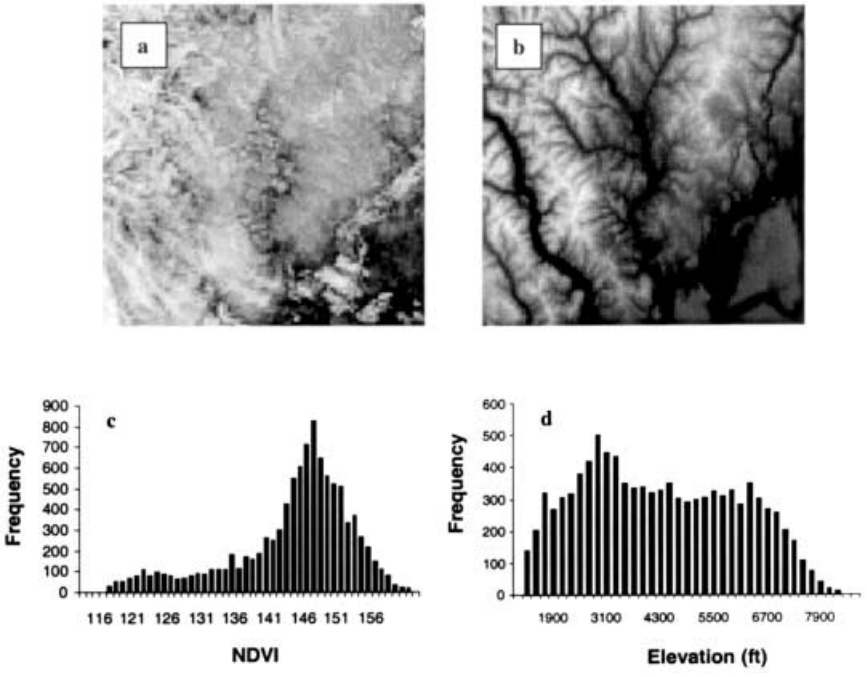

Fig. 17. (a) The 100×100 grid of NDVI derived from AVHRR data from 7 to 20 August, 1992, collected over the Cascade Mountains. Darker shades represent smaller values. (b) The 100×100 grid of elevation data from the same location. (c) Frequency distribution of NDVI and (d) elevation in feet.

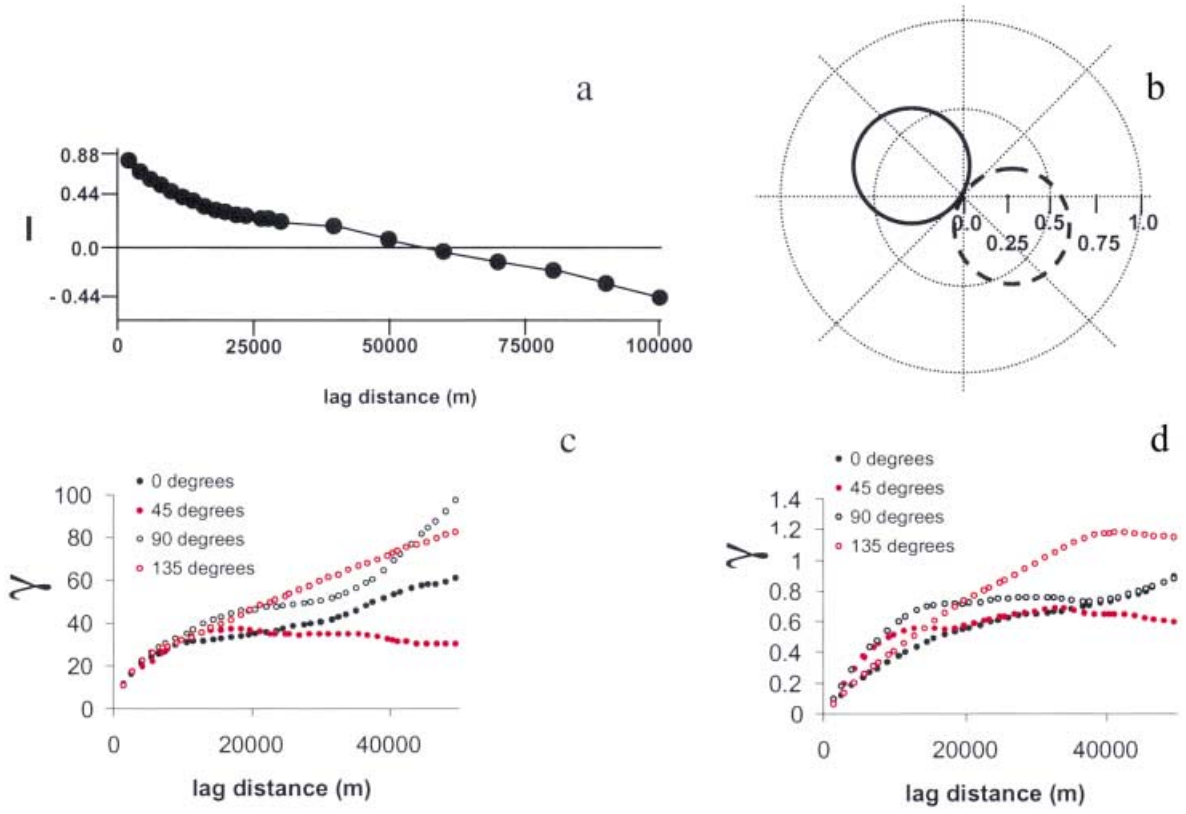

Fig. 18. (a) Moran's I correlogram for NDVI; all lags are significant ($p < 0.05$). (b) Angular correlation for NDVI. (c) Directional variogram for NDVI and (d) elevation.

There was a very strong linear relationship between logarithmically-transformed plant canopy volume and counts, and when mapped these two derived (**z**) variables appeared very similar. However, a frequency distribution of transformed volume per plant (Fig. 16) showed a clearly bimodal distribution with relatively many plants in the smallest size class and a second more symmetric mode for much larger plants. Such bimodality is characteristic of many long-lived plants in which seeds readily germinate, but mortality of smaller individuals is high until a certain size threshold is reached, after which survival probabilities increase until senescence. Also, mapping revealed considerable spatial pattern in the values of volume per plant within the 5×5 m contiguous subareas. Broadly, volumes per plant were greater in the left-hand half of the study arena ($x < 50$, mean transformed volume per plant = 12.93 with SEM 0.121; $x > 50$, mean transformed volume per plant = 12.42 with SEM 0.099). Clustering was significant (SADIE indices both $p = 0.0002$), with two regions of relatively low volume per plant (top centre and lower-right side, each 200–300 m^2) and one 200 m^2 patch of greater than average volume, at the lower-left of the area. The reason for spatial structure in these growth differences between plants is unknown.

Case study 3: remotely-sensed AVHRR data in the Cascade Mountains

Since 1982, the Advanced Very High-Resolution Radiometer (AVHRR) satellite imager (Eidenshink 1992) has collected image data from daily coverage of the Earth, using a nominal 1-km sampling rate. We downloaded archived and processed area-referenced AVHRR data, collected from 7 to 20 August 1992, from the EROS Data Center web site ⟨http://edcwww.cr.usgs.gov/landdaac/1KM/comp10d.html⟩. We extracted a 100×100 matrix of ca 1 km square cells located on the east side of the Northern Cascade Mountains, Washington (Fig. 17a). The normalized difference vegetation index (NDVI) (James and Kalluri 1993, Townshend et al. 1994), the difference of near-infrared (channel 2) and visible (channel 1) reflectance divided by their sum, were also provided. Positive values of NDVI indicate green vegetation; negative values indicate non-vegetated surface features such as water, barren rock, ice, snow or clouds. To minimize data storage requirements, NDVI values were rescaled as integers in the interval [0,200]. A US Geological Survey digital elevation model from the same region (e.g. Thelin and Pike 1991) was registered with the

NDVI data and resampled to the same 1×1 km grid cells (Fig. 17b). Frequency distributions for NDVI (Fig. 17c) and elevation (measured in feet, Fig. 17d) exhibited only slight skewness and no transformation was needed before analysis. Specific questions posed for these data were: what are the spatial patterns of NDVI and elevation and how similar are they? Are these two variables correlated?

Declining values in Moran's I correlogram (Fig. 18a) and increasing values in the directional variogram (Fig. 18c) indicated strong, positive autocorrelation in vegetation across the map, while their monotonic nature across the full range of distances indicated the presence of a large-scale trend or gradient of vegetation. This trend was present in every direction except for a 45° angle. The angular correlation diagram (Fig. 18b) confirmed this directionality, achieving maximal values between 90° and 135°. This revealed anisotropy may be seen clearly in the raw data (Fig. 17a); highest NDVI values are to the north-west and lowest values to the south-east. This pattern probably reflects the distribution of forested areas along the eastern slope of the North Cascade Mountains. Forests are more abundant at higher elevations and high desert vegetation is dominant at the lower elevations to the south-west. The correlogram, angular correlation diagram and anisotropy directions for elevation were so similar to those for NDVI that they are not shown. The directional variogram for elevation (Fig. 18d) confirms the anisotropy shown by the trend and direction of the strongest gradient evident in the map (Fig. 17b).

Rossi et al. (1992) pointed out that a large-scale trend, such as the anisotropy found here for NDVI and elevation, could obscure patterns of spatial dependence at smaller lag distances. Here, it is unclear from the directional variograms (Fig. 18c, d) whether anisotropy exists in autocorrelation at short lag distances or is limited to the large-scale trend.

In order to investigate further the scale of autocorrelation and anisotropy, we removed the large-scale trend in both data sets by fitting a model of the form: $z_{x,y} = a + b_1x + b_2y$. Estimates of the parameters for NDVI were: $a = -172.4$, $b_1 = -0.000173$, $b_2 = 0.000107$; and for elevation: $a = -55119$, $b_1 = -0.0303$, $b_2 = 0.0258$. The directional variograms for the detrended NDVI residuals from this model (Fig. 19a) differed from those discussed above; they reached a maximum (sill), and indicated less anisotropy, although the range appeared slightly longer in the 135° direction than for other angles. The directional variograms of the detrended elevation data (Fig. 19b) showed precisely the same differences, but additionally, for the 90° angle the elevation variogram exhibited a slight "hole effect", first increasing then decreasing with increasing lag distances. This pattern is characteristic of some sort of oscillatory undulation or repetitive fluctuation (Isaaks and Srivastava 1989) here probably caused by the regular alternation of valleys and ridges. The same pattern also appears, although less distinctly, in the 90° variogram (Fig. 19a) of the NDVI data.

Overall, there was a striking similarity between the NDVI and elevation variograms. Both had nugget effects near zero for the raw data, indicating the existence of a very continuous, smooth surface. The ranges of both detrended sets of data were between 10 000 and 20 000 m, presumably reflecting the average distance between mountain peaks. However, the similarity of the variograms cannot be used alone to infer their association, since many different spatial arrangements may share the same variogram (Liebhold and Sharov 1998). We calculated a Pearson correlation coefficient of 0.4916 between raw NDVI and elevation; a "naïve" test

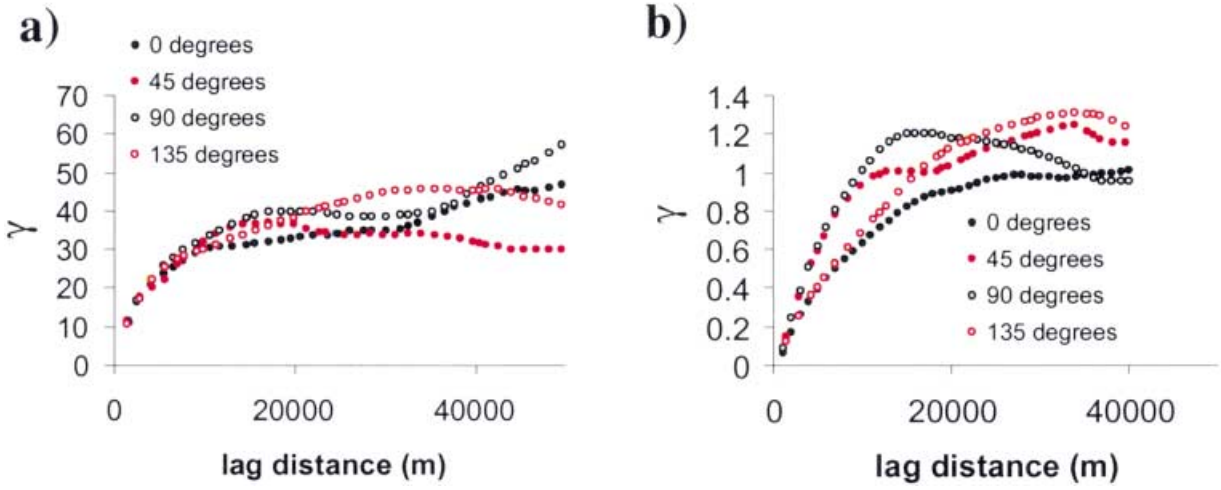

Fig. 19. Directional variograms of detrended data. (a) NDVI, (b) elevation.

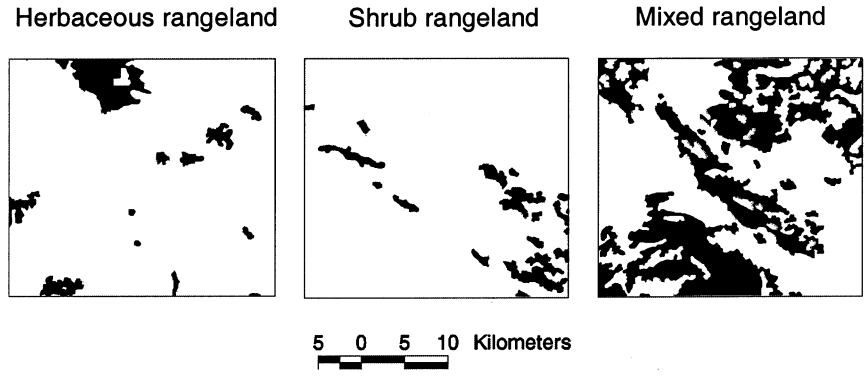

Fig. 20. Vector representation of land use and land cover (LULC) data.

for association without adjusting for spatial autocorrelation would indicate significance ($p < 0.0001$). However, the existence of autocorrelation in both variables violates the assumption of independence among samples and the test of correlation requires adjustment (Clifford et al. 1989, Bocard et al. 1992, Dutilleul 1993). This is because the presence of spatial autocorrelation indicates that the amount of information in a set of data is less, and may be considerably less, than that in a sample of the same size where there is none. This loss of information is therefore allowed for by an adjustment that leaves the magnitude of the correlation unchanged, but alters the degrees of freedom of the test. We used the variograms in Fig. 18 to adjust the degrees of freedom from 9998 to 14.52 using the method of Clifford et al. (1989); the revised, valid test was not significant ($p = 0.0679$). Hence, the observed large magnitude of the correlation was almost completely due to common and large-scale spatial structure; after adjustment for this effect, correlation was only marginally significant.

Case study 4: land cover in Monterey County, California

Land use land cover (LULC) data are area-referenced data developed by the US Geological Survey (USGS) for every portion of the conterminous USA at either a 1:100 000 or 1:250 000 scale. They are available from the USGS web site in either vector or raster format (Anon. 1986). Data are derived by manual interpretation of aerial photographs, incorporating information from earlier land-use maps and field surveys. Land-use cover is classified according to the scheme developed by Anderson et al. (1976).

The data used in this case study consisted of the south-eastern corner of the 1:100 000 Monterey quadrangle located in the central coast region of California. The study arena spanned 30.4 km from east to west and 27.2 km from south to north. Only land areas classified into three land cover classes were included: herbaceous rangeland (code 31); shrub and brush rangeland (code 32); and mixed rangeland (code 33). All analyses were performed on data projected to the Universal Transverse Mercator projection (zone = 10) (Fig. 20). The analysis described here is limited to those descriptive statistics (McGarigal and Marks 1995) that measure patch size, edge density and shape from polygonal data (Table 2). Mixed rangeland was the dominant land cover type and also had the greatest mean patch size. Both fragmentary indices, patch size CV and edge density, were also largest for mixed rangeland. Also, both the mean shape index and the area-weighted shape index were greatest for the mixed rangeland cover type, indicating a greater complexity of polygon shape.

Omnidirectional indicator variograms (Isaaks and Srivastava 1989) were calculated for each set of binary data in raster format (200 m pixels) representing the presence-absence of the three land covers. Relative variograms were calculated for normalized data values from each cover type, allowing direct comparison of their spatial structure despite their different means and variances. Variogram range was largest for herbaceous rangeland and smallest for the shrub and brush rangeland (Fig. 21a). The range reflects both patch size and within-patch (gap) size (Woodcock et al. 1988). This explains in part why shrub rangeland had the smallest mean patch size (Table 2) and also the shortest range. Woodcock et al. (1988) also reported how increases in the object size variance resulted in a more rounded shape of the variogram, confirmed here by that for shrub rangeland (Fig. 21a) which also had the lowest patch-size CV (Table 2). However, note that the herbaceous rangeland had the longest range, even though its patch size statistics were both smaller than that for mixed rangeland, probably because of the larger total area of the latter. This again illustrated how the range of an indicator variogram reflects the size of gaps as well as patches.

Table 2. Summary statistics for the land use and land cover data, based on descriptors of patches, edges and shape for polygonal data describing herbaceous, shrub and mixed rangelands.

Land type	% of total area	Mean patch size (ha)	Number of patches	Patch size standard deviation (ha)	Patch size coefficient of variation	Total edge length (m)	Edge density (m ha^{-1})	Mean shape index	Area-weighted mean shape index
Herbaceous rangeland	7.3	469.5	13	875.9	186.5	140 000	1.7	1.59	1.9
Shrub rangeland	5.2	243.6	18	330.6	135.7	177 000	2.1	1.63	2.2
Mixed rangeland	36.8	1404.1	22	3090.8	220.1	676 200	8.1	1.97	4.6

Directional variograms for shrub rangeland (Fig. 21b) demonstrated anisotropy; the range was considerably longer in the north and northwest directions. This difference appeared to reflect both an elongation of patches in a consistent north-westerly direction, as well as the arrangement of patches along a band running in that same direction (Fig. 20). This anisotropy probably reflects ridge topography that may be associated with vegetation. The hole effect (depression at ca 2300 m) in the north-east directional variogram (Fig. 21b) reflected the alternating presence and absence encountered when traversing the study arena in a direction perpendicular to the angle of patch elongation (Fig. 20).

Discussion

We have attempted to offer, for a limited number of sets of data, some flavour of the kinds of outputs, inferences and interpretations possible from spatial analysis. Readers must realize that we have not attempted an exhaustive account. Many methods are capable of extension to provide further features, to meet specific needs. For example, almost all of the local quadrat variance methods detect the average size of patches and gaps in a phased pattern along a transect, but new local variance (Galiano 1982) detects the size of the smallest phase of the pattern, whether it be patches or gaps. This, when combined with other local quadrat variance methods, allows separate estimation of patch and gap size.

We give only four recommendations to ecologists with spatial data in search of methods with which to analyze them. First, make extensive use of simple visualization techniques such as graphs and mapping (Tufte 1997, Carr 1999) as a first step to understanding spatial characteristics in the data. Second, select statistical methods that are available and appropriate for the data type. Third, select a method that can answer pertinent questions and provide relevant spatial information. Reference to Table 1 may help the selection process. Given the suitability of several methods, it would be invidious and naïve to attempt to go beyond this to recommend which single specific technique to use. Readers must form their own conclusions, informed by the theoretical properties of the methods and their performance in the above case studies. Most methods are distinct and no single one can identify all of the spatial characteristics in data. Therefore, our fourth recommendation is to employ several different techniques. We do not believe that any methods we have examined intrinsically provide redundant information, fail to detect real patterns, or identify patterns that are not real.

In certain circumstances, it may be sensible to widen the possible range of techniques that may be brought to bear on a problem, by converting between data types.

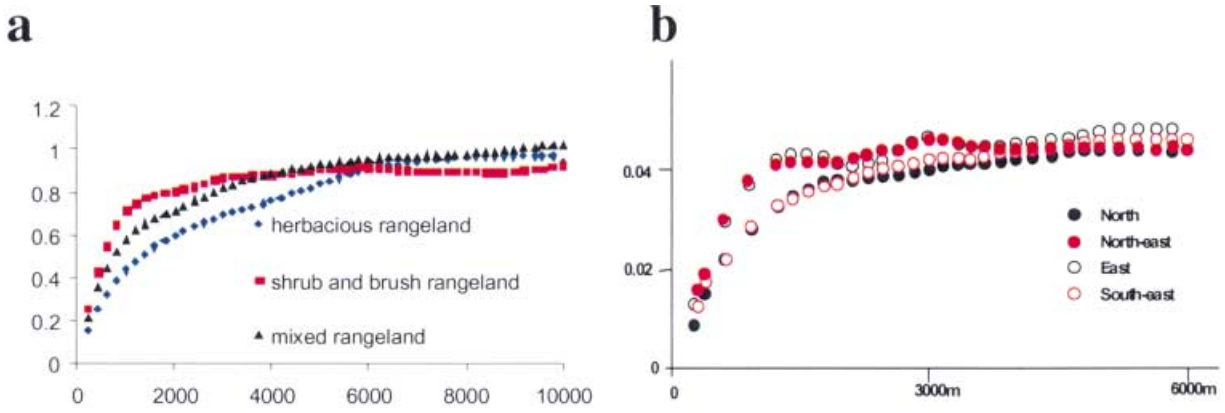

Fig. 21. (a) Omnidirectional relative variograms for LULC data, with semivariance plotted against lagged distance; (b) directional variograms for the shrub and brush rangeland data in (a).

Examples are the formation of counts by conversion of data from point locations to contiguous subareas, the taking of samples, and transformation from polygonal to contiguous subareas (i.e. vector to raster format). For the first two the degree of loss of information must be considered (Dungan et al. 2002). For the third, note that both geographic entities and fields can be represented either in vector (represented by origin, length and direction), raster (regular tessellation with square elements) or TIN (irregular tessellation with triangular elements) form. The choice of form depends largely on the application of the data (Burrough and McDonnell 1998). The vector form assures better accuracy and efficient data storage, while raster data can be processed faster and are adapted easily to deal with changes in modes of analysis. Conversion of polygonal data from vector to raster form was illustrated in case study 4. Indeed, the raster format is exploited increasingly as the primary data type in landscape ecology, because of the growing utilization of remotely-sensed data in that discipline (and see case study 3, above). The conversion facilitated the application of a wider range of statistical methods, such as indicator variograms, frequently used in disciplines other than landscape ecology. A further example of the generic use of statistical techniques is the texture and "lacunarity" measures employed in remote sensing and landscape ecology (Plotnick et al. 1993), that calculate variance across gliding windows. Dale et al. (2002) note their similarity to the local quadrat variance methods used in plant ecology.

Acknowledgements – This work was conducted as part of the "Integrating the Statistical Modeling of Spatial Data in Ecology" Working Group supported by the National Center for Ecological Analysis and Synthesis (NCEAS). NCEAS is a center funded by the NSF (Grant # DEB-94-21535) of the USA, the Univ. of California-Santa Barbara, the California Resources Agency, and the California Environmental Protection Agency. IACR-Rothamsted receives grant-aided support from the Biotechnology and Biological Sciences Research Council of the U.K.

References

Anderson, J. R. et al. 1976. A land use and land cover classification system for use with remote sensor data. – U.S. Geol. Surv., Prof. Pap. 964.

Anon 1986. Land use land cover digital data from 1:250 000 and 1:100 000-scale maps, data user guide 4. – U.S. Geol. Surv.

Anselin, L. 1995. Local indicators of spatial association – LISA. – Geogr. Anal. 27: 93–115.

Bailey, R. G. and Ropes, L. 1998. Ecoregions: the ecosystem geography of the oceans and continents. – Springer.

Bjørnstad, O. N., Ims, R. A. and Lambin, X. 1999. Spatial population dynamics: analyzing patterns and processes of population synchrony. – Trends Ecol. Evol. 14: 427–431.

Bliss, C. I. 1941. Statistical problems in estimating populations of Japanese beetle larvae. – J. Econ. Entomol. 34: 221–232.

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

Bradshaw, G. A. and Spies, T. A. 1992. Characterizing canopy gap structure in forests using wavelet analysis. – J. Ecol. 80: 205–215.

Burrough, P. A. and McDonnell, R. A. 1998. Principles of geographical information systems. – Oxford Univ. Press.

Carr, D. B. 1999. New templates for environmental graphics: from micromaps to global grids. – Proc. of the 9th Lukacs Symp., Bowling Green State Univ., Bowling Green, Ohio, USA, 23–25 April.

Chatfield, C. 1985. The initial examination of data. – J. R. Stat. Soc. A 148: 214–253.

Clark, P. J. and Evans, F. C. 1954. Distance to nearest neighbour as a measure of spatial relationships in populations. – Ecology 35: 23–30.

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

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

Clifford, P., Richardson, S. and Hémon, D. 1989. Assessing the significance of the correlation between two spatial processes. – Biometrics 45: 123–134.

Cressie, N. A. C. 1991. Statistics for spatial data. – Wiley.

Dale, M. R. T. 1999. Spatial pattern analysis in plant ecology. – Cambridge Univ. Press.

Dale, M. R. T. and Zbigniewicz, M. W. 1997. Spatial pattern in boreal shrub communities: effects of a peak in herbivore densities. – Can. J. Bot. 75: 1342–1348.

Dale, M. R. T. and Mah, M. 1998. The use of wavelets for spatial pattern analysis in ecology. – J. Veg. Sci. 9: 805–814.

Dale, M. R. T. et al. 2002. Conceptual and mathematical relationships among methods for spatial analysis. – Ecography 25: 558–577.
David, F. N. and Moore, P. G. 1954. Notes on contagious distributions in plant populations. – Ann. Bot. Lond. 18: 47–53.
Diggle, P. J. 1983. Statistical analysis of spatial point patterns. – Academic Press.
Donnelly, K. P. 1978. Simulations to determine the variance and edge-effect of total nearest neighbor distance. – In: Hodder, I. (ed.), Simulation studies in archaeology. Cambridge Univ. Press, pp. 91–95.
Dungan, J. et al. 2002. A balanced view of scale in spatial statistical analysis. – Ecography 25: 626–640.
Dutilleul, P. 1993. Modifying the t-test for assessing the correlation between two spatial processes. – Biometrics 49: 305–314.
Eidenshink, J. C. 1992. The 1990 conterminous U.S. AVHRR data set. – Photo. Engin. Rem. Sens. 58: 809–813.
Falsetti, A. B. and Sokal, R. R. 1993. Genetic structure of human populations in the British Isles. – Ann. Human Biol. 20: 315–329.
Fisher, R., Thornton, H. and Mackenzie, W. 1922. The accuracy of the plating method of estimating the density of bacterial populations. – Ann. Appl. Biol. 9: 325–359.
Galiano, E. F. 1982. Détection et mesure de l'hétérogénéité spatial des espèces dans les pâturages. – Acta Oecol. Oecol. Plant. 3: 269–278.
Getis, A. 1991. Spatial interaction and spatial autocorrelation: a cross-product approach. – Environ. Plann. A 23: 1269–1277.
Getis, A. and Ord, J. K. 1995. The analysis of spatial association by use of distance statistics. – Geogr. Anal. 24: 189–206.
Greig-Smith, P. 1952. The use of random and contiguous quadrats in the study of the structure of plant communities. – Ann. Bot. 16: 293–316.
Gustafson, E. J. 1998. Quantifying landscape spatial pattern: what is the state of the art? – Ecosystems 1: 143–156.
Haase, P. 1995. Spatial pattern analysis in ecology based on Ripley's K-function: introduction and methods of edge correction. – J. Veg. Sci. 6: 575–582.
Hanski, I. A. and Gilpin, M. E. 1997. Metapopulation biology: ecology, genetics and evolution. – Academic Press.
Hassell, M. P., Comins, H. and May, R. M. 1991. Spatial structure and chaos in insect population dynamics. – Nature 353: 255–258.
Hurlbert, S. 1990. Spatial distribution of the montane unicorn. – Oikos 58: 257–271.
Isaaks, E. H. and Srivastava, R. M. 1989. An introduction to applied geostatistics. – Oxford Univ. Press.
James, M. E. and Kalluri, S. 1993. The pathfinder AVHRR land data set: an improved coarse resolution data set for terrestrial monitoring. – The Int. J. Rem. Sens. 15: 3347–3364.
Jumars, P. A., Thistle, D. and Jones, M. L. 1977. Detecting two-dimensional spatial structure in biological data. – Oecologia 28: 109–123.
Keitt, T. et al. 2002. Accounting for spatial pattern when modeling organism-environment interactions. – Ecography 25: 616–625.
Korie, S. et al. 1998. Analysing maps of dispersal around a single focus (with Discussion). – Ecol. Environ. Stat. 5: 317–350.
Korie, S. et al. 2000. Spatiotemporal associations in beetle and virus count data. – J. Agricult. Biol. Environ. Stat. 5: 214–239.
Legendre, P. and Fortin, M.-J. 1989. Spatial pattern and ecological analysis. – Vegetatio 80: 107–138.
Legendre, P. et al. 2002. The consequences of spatial structure for the design and analysis of ecological field surveys. – Ecography 25: 601–615.
Liebhold, A. M. and Sharov, A. A. 1998. Testing for correlation in the presence of spatial autocorrelation in insect count data. – In: Baumgartner, J., Brandmayr, P. and Manly, B. F. J. (eds), Population and community ecology for insect management and conservation. Balkema, pp. 111–117.
Liebhold, A. M., Rossi, R. E. and Kemp, W. P. 1993. Geostatistics and geographic information systems in applied insect ecology. – Annu. Rev. Entomol. 38: 303–327.
Lloyd, M. 1967. Mean crowding. – J. Anim. Ecol. 36: 1–30.
Matern, B. 1986. Spatial variation, number 36 in lectures notes in statistics. – Springer.
May, R. M. (ed.) 1976. Theoretical ecology. – Blackwell.
McGarigal, K. and Marks, B. J. 1995. FRAGSTATS: spatial pattern analysis program for quantifying landscape structure. – USDA For. Serv. Gen. Tech. Rep., PNW-GTR-351, USDA For. Serv. Pacific Northwest Res. Stn.
Miriti, M., Howe, H. F. and Wright, S. J. 1998. Spatial patterns of mortality in a Colorado Desert plant community. – Plant Ecol. 136: 41–51.
Moran, P. A. P. 1950. Notes on continuous stochastic phenomena. – Biometrika 37: 17–23.
Morgan, K. A. and Gates, J. E. 1982. Bird population patterns in forest edge and strip vegetation at Remington Farms, Maryland. – J. Wildl. Manage. 46: 933–944.
Morisita, M. 1959. Measuring the dispersion and the analysis of distribution patterns. – Mem. Fac. Sci., Kyushu Univ., Ser. E. Biol. 2: 215–235.
Mugglestone, M. A. and Renshaw, E. 1996. A practical guide to the spectral analysis of spatial point processes. – Comp. Stat. Data Anal. 21: 43–65.
Oden, N. L. 1984. Assessing the significance of a spatial correlogram. – Geogr. Anal. 16: 1–16.
Oden, N. L. and Sokal, R. R. 1986. Directional autocorrelation: an extension of spatial correlograms in two dimension. – Syst. Zool. 35: 608–617.
Ord, J. K. and Getis, A. 1995. Local spatial autocorrelation statistics: distributional issues and an application. – Geogr. Anal. 27: 286–306.
Perry, J. N. 1998a. Measures of spatial pattern for counts. – Ecology 79: 1008–1017.
Perry, J. N. 1998b. Measures of spatial pattern and spatial association for counts of insects. – In: Baumgartner, J., Brandmayr, P. and Manly, B. F. J. (eds), Population and community ecology for insect management and conservation. Balkema, pp. 21–33.
Perry, J. N. and Woiwod, I. P. 1992. Fitting Taylor's power law. – Oikos 65: 538–542.
Perry, J. N. et al. 1999. Red-blue plots for detecting clusters in count data. – Ecol. Lett. 2: 106–113.
Peuquet, D. 1984. A conceptual framework and comparison of spatial data models. – Cartographica 21: 66–113.
Peuquet, D., Smith, B. and Brogaard, B. 1999. The ontology of fields. – Rep. of a Specialist Meeting held under the Auspices of the Varenius Project, Bar Harbor, Maine, 11–13 June, 1998, ⟨http://www.ncgia.ucsb.edu/~vanzuyle/varenius/ontologyoffields_rpt.html⟩.
Plotnick, R. E., Gardner, R. H. and O'Neill, R. V. 1993. Lacunarity indices as measures of landscape texture. – Landscape Ecol. 8: 201–211.
Ripley, B. D. 1977. Modelling spatial patterns. – J. R. Stat. Soc. B 39: 172–212.
Ripley, B. D. 1988. Statistical inference for spatial processes. – Cambridge Univ. Press.
Rosenberg, M. S. 2000. The bearing correlogram: a new method for analyzing directional spatial autocorrelation. – Geogr. Anal. 32: 267–278.
Rossi, R. E. et al. 1992. Geostatistical tools for modeling and interpreting ecological spatial dependence. – Ecol. Monogr. 62: 277–314.
Silvertown, J. et al. 1992. Cellular automaton models of interspecific competition for space – the effect of pattern on process. – J. Ecol. 80: 527–534.

Simon, G. 1997. An angular version of spatial correlations, with exact significance tests. – Geogr. Anal. 29: 267–278.
Skellam, J. G. 1952. Studies in statistical ecology. I. Spatial pattern. – Biometrika 39: 346–362.
Sokal, R. R. and Oden, N. L. 1978. Spatial autocorrelation in biology. 1. Methodology. – Biol. J. Linn. Soc. 16: 199–228.
Sokal, R. R., Oden, N. L. and Thomson, B. A. 1998a. Local spatial autocorrelation in biological variables. – Biol. J. Linn. Soc. 65: 41–62.
Sokal, R. R., Oden, N. L. and Thomson, B. A. 1998b. Local spatial autocorrelation in a biological model. – Geogr. Anal. 30: 331–354.
Southwood, T. R. E. 1966. Ecological methods. – Methuen.
Stoyan, D. and Stoyan, H. 1994. Fractals, random shapes and point fields, methods of geometrical statistics. – Wiley.
Taylor, L. R., Woiwod, I. P. and Perry, J. N. 1978. The density-dependence of spatial behaviour and the rarity of randomness. – J. Anim. Ecol. 47: 383–406.
Thelin, G. P. and Pike, R. J. 1991. Landforms of the conterminous United States – a digital shaded-relief portrayal. – USGS.
Townshend, J. R. G. et al. 1994. The 1-km AVHRR global data set: needs of the International Geosphere Biosphere Program. – The Int. J. Rem. Sens. 15: 3319–3332.
Tufte, E. R. 1997. Visual explanations: images and quantities, evidence and narrative. – Graphics Press.
Turner, M. G. and Gardner, R. H. 1991. Quantitative methods in landscape ecology. – Springer.
Upton, G. J. G. and Fingleton, B. 1985. Spatial data analysis by example. Vol. 1: point pattern and quantitative data. – Wiley.
ver Hoef, J. M., Cressie, N. A. C. and Glenn-Lewin, D. C. 1993. Spatial models for spatial statistics: some unification. – J. Veg. Sci. 4: 441–452.
Watt, A. S. 1947. Pattern and process in the plant community. – J. Ecol. 35: 1–22.
Wiens, J. A. 1989. Spatial scaling in ecology. – Funct. Ecol. 3: 385–397.
Winder, L. et al. 2001. Modelling the dynamic spatio-temporal response of predators to transient prey patches in the field. – Ecol. Lett. 4: 568–576.
Woodcock, C. E., Strahler, A. H. and Jupp, D. L. B. 1988. The use of variograms in remote sensing: I. Scene models and simulated images. – Rem. Sen. Environ. 25: 323–348.
Wright, S. J. and Howe, H. F. 1987. Pattern and mortality in Colorado desert plants. – Oecologia 73: 543–552.