Spatial analysis in ecology

Marie-Josée Fortin, Mark R.T. Dale & Jay ver Hoef

Volume 4, pp 2051–2058

in

Encyclopedia of Environmetrics
(ISBN 0471 899976)

Edited by

Abdel H. El-Shaarawi and Walter W. Piegorsch

Spatial analysis in ecology

The first step in understanding ecological processes is to identify patterns. Ecological data are usually characterized by spatial structures due to spatial autocorrelation. *Spatial autocorrelation* refers to the pattern in which observations from nearby locations are more likely to have similar magnitude than by chance alone. The magnitude, intensity, as well as extent of spatial autocorrelation can be quantified using spatial statistics [3–5, 7]. Most ecological data exhibit some degree of spatial autocorrelation, depending on the scale at which the data were recorded and then analyzed. Ecological phenomena are also characterized by the multiple ecological processes that act upon them; these processes often operate at more than one spatial scale. Ecological data are a composite of several spatial scales: trends at macroscales; patches, gradients and patterns at meso- and local scales; and random patterns at local and microscales. The different processes and patterns at different scales are not necessarily linear or additive, and this contributes to the degree of spatial dependence in the data.

Spatial autocorrelation can have four sources, of which two are of direct scientific interest. There is a distinction between *spurious*, *interpolative*, *true* and *induced* autocorrelation [7]. Even when observations are statistically independent, spurious autocorrelation may be observed due to underlying processes affecting the spatial arrangement of the data. Interpolative autocorrelation arises when spatial response surfaces are smoothed, interpolated, or extrapolated. True spatial autocorrelation arises from causal interaction among nearby sample locations. Finally, spatial autocorrelation may be induced by a dependent variable through a causal relationship with another spatially autocorrelated variable. Spurious and interpolative autocorrelation are a nuisance. True and induced spatial autocorrelation arises from space–time processes (e.g. dispersal, migration; *see* **Point processes, spatial–temporal**) that are of direct scientific interest. This causality motivated researchers to use spatial autocorrelation as a clue or 'signature' left by the past action of space–time processes. This inherent spatial autocorrelation of ecological data has fundamental environmetric implications: it violates assumptions of independence required by many parametric inferential tests. To ensure independent samples, the use of a random sampling design is often recommended (*see* **Spatially constrained sampling**); this, however, will not remove true or induced spatial autocorrelation.

Properties of Spatial Statistics

Spatial statistics are based on assumptions. The most important one is the stationarity assumption, implying that the data should be normally distributed with the same mean and variance over the entire study area [3, 6, 7]. The spatial pattern should also be *isotropic*, i.e. the pattern shows the same intensity in all directions. A spatial pattern that varies according to direction is *anisotropic*.

Furthermore, both the size (i.e. area) and shape (e.g. rectangular rather than square) of the study plot affect the ability of most spatial statistics to estimate accurately the intensity, range and type of spatial pattern as well as its significance. This problem is known as the **edge effect**, where near the edge (boundary) of the study plot there are fewer neighboring sampling units than at the middle of the study plot. Hence, spatial patterns estimated at small and large spatial lags are based on fewer sampling units than those for intermediate spatial lags, affecting the accuracy of the estimated spatial pattern. Several corrections for edge effects have been proposed and are specific to the particularity of spatial statistical methods [3]; a simple rule of thumb is to compute a spatial pattern for half to two-thirds of the smallest edge length of the study plot [5].

Spatial Statistics: An Overview

Since the 1950s, several spatial methods of analysis have been developed and modified to improve our ability to detect and characterize spatial patterns. These stem from several fields of study (plant ecology, animal ecology, geography, mining engineering), having more or less different goals (explorative vs. inference), mathematical approaches (variance–covariance vs. count-based methods) and underlying assumptions (stationarity or pseudo-stationarity) [1–10].

Selection of appropriate methods for use with spatial data can be established either by the research objective (Table 1) or by the measurement types and sampling designs of the data (Table 2). Spatial statistics can also be classified according to the type

Table 1 Spatial statistics classified by objective

Objective	Spatial statistics
Exploration	Nearest neighbor, k-nearest neighbor Ripley's K (uni- and bivariate) Join-count Moran's I, Geary's c, semivariance γ Mantel test (multivariate)
Inference	Ripley's K (uni- and bivariate) Join-count Moran's I, Geary's c, semivariance γ Mantel test (multivariate)
Mapping (interpolation)	Trend surface analysis, kriging, splines, Voronoi polygons

of spatial structure that the methods are measuring or estimating: (a) global spatial structure (e.g. variance/mean ratio and aggregated indices); (b) spatial periodicity (e.g. spectral analysis, **wavelet** analysis, **fractal dimension**); (c) spatial intensity and range as a function of spatial lags (e.g. **Ripley's *K* function**, blocked quadrat variance methods, join-count, Moran's I, Geary's c, Mantel test, semivariance γ, SADIE); and (d) spatial interpolation (e.g. trend surface analysis, **kriging**, splines, Voronoi polygons). Below, we present only a few of the spatial statistics that are currently used in ecology. Several excellent spatial statistics textbooks and review articles are available for further detail about these methods [1–10].

Nearest Neighbor Distance

Given the above definition of spatial autocorrelation, it is expected that the $x-y$ coordinates of 'points' (e.g. individual plant stems) having a spatial structure are more likely to be spatially close than expected by chance alone. Following this simple idea, the **nearest neighbor method** measures the mean nearest distance for all points $\overline{d_i}$, where $i = 1$ for the first neighbor [3, 5, 10]. This analysis assumes that all the points (e.g. individual trees) in the study plot are surveyed and mapped. Then, the observed mean nearest distance is compared to the expected mean nearest distance. Under complete spatial randomness (CSR), counts in local areas follow a Poisson distribution and $\overline{d_i}$ follows a **Weibull distribution**, so that

$$\mathrm{E}(\overline{d_i}) = \gamma_i \left(\frac{A}{n}\right)^{1/2} \tag{1}$$

where γ_i is a constant which varies as a function of the ith neighbor analyzed, A is the total area sampled, and n is the number of points mapped.

Table 2 Spatial statistics classified by data measurement type and sampling design

	Data types	
Sampling design	Categorical/qualitative	Numerical/quantitative
Exhaustive census ($x-y$ coordinates)	Nearest neighbors k-Nearest neighbors Ripley's K (uni- and bivariate) Join-count	Aggregation indices (e.g. variance/mean, etc.)
Regular spacing (1D and 2D)	Block variance quadrat Spectral analysis Wavelet analysis Fractal dimension	Moran's I, Geary's c, Getis (global and local) Semivariance γ SADIE Mantel test (multivariate) Trend surface analysis, kriging, splines
Irregular spacing (1D and 2D)	Fractal dimension	Moran's I, Geary's c, Getis (global and local) Semivariance γ SADIE Mantel test (multivariate) Trend surface analysis, kriging, splines, Voronoi polygons

The size (i.e. area) of the study plot, A, will affect the expected nearest neighbor value. This technique has been extended to higher neighbors, k-nearest neighbors, and is also known as the *refined nearest neighbor method*.

Ripley's K Function

Following the spirit of the refined nearest neighbor analysis, **Ripley's *K* function** quantifies the spatial pattern intensity of points for various sizes of a circular search window [3, 10]. As with the nearest neighbor method, points correspond to the locations of discrete events (e.g. individuals). All the events of a given study plot need to be mapped (i.e. $x-y$ coordinate of each event). Ripley's K function computes the overall mean number of points lying within a circular search window of radius t:

$$\widehat{K}(t) = \frac{\lambda^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n} I_t(e_i, e_j)}{n} \quad \text{for } i \neq j \text{ and } t > 0 \tag{2}$$

where the point intensity, λ, is estimated as the density n/A, I_t is an indicator function that takes value 1 when e_j is within distance t of event e_i (and 0 otherwise), and n is the total number of events (Figure 1a). By using a circular window, Ripley's K function is an isotropic cumulative count of all points at distances from 0 to t. The expected number of events under a CSR process is πt^2.

To linearize and stabilize the variances [3], what can be called Ripley's $\widehat{L}(t)$ should be used instead of $\widehat{K}(t)$:

$$\widehat{L}(t) = \left(\frac{\widehat{K}(t)}{\pi}\right)^{1/2} \quad \text{or} \quad \widehat{L}(t) = t - \left(\frac{\widehat{K}(t)}{\pi}\right)^{1/2} \tag{3}$$

The expected value of $\widehat{L}(t)$ under a **Poisson process** is 0: positive values indicate *spatial clustering*, while negative values indicate *spatial segregation* (Figure 1a). Monte Carlo simulations of the Poisson point pattern process (i.e. CSR), or other more realistic point processes (*see* **Point processes, spatial**), are used to provide a confidence envelope of this function [2].

Ripley's K and L functions have been extended to analyze the spatial relationship between points from two event types where positive values indicate a positive interaction between the two variables, while negative values indicate spatial segregation or repulsion between them.

Blocked Quadrat Variance

Transects or grids of contiguous sampling units (e.g. quadrats) are used typically in plant ecology to identify changes in spatial structure along a gradient. The *blocked quadrat variance* methods [4] have been developed to identify spatial pattern in data of contiguous samples by computing the variance using various sizes of blocks of units as the search window. Several algorithms of this type have been proposed: for example, two-term local quadrat variance (TTLQV), paired-quadrat variance, and three-term local quadrat variance [4]. The TTLQV, for example, computes the variance from pairs of adjacent blocks of b units each as

$$V_2(b) = \frac{\sum_{i=1}^{n+1-2b}\left(\sum_{j=i}^{i+b-1} x_j - \sum_{j=i+b}^{i+2b-1} x_j\right)^2}{2b(n+1-2b)} \tag{4}$$

for a range of values of b. This variance is plotted as a function of b (Figure 1b) and peaks in the variance are interpreted as indicating scales of spatial pattern in the data [4].

Join-count Statistics

Spatial patterns for **binary data** (e.g. presence/absence) from adjacent sampling units (e.g. quadrats) or regions (e.g. counties) can be assessed using *join-count statistics* [10]. For the binary case, the null hypothesis states that neighboring regions are more likely to be of the same category, say 0 (white) or 1 (black), and therefore not described by a pattern of CSR. The observed join-count statistics (J_{BB} and J_{WW}) count the number of join encounters in adjacent regions having the same category; the corresponding J_{BW} statistic counts the number of adjacent regions not having the same category. Hence the J_{BB} and J_{WW} statistics assess the presence of positive spatial autocorrelation, while J_{BW} assesses the presence of negative spatial autocorrelation. Where the observation x_i is 0 or 1, the count of black–black (1–1) joins can be calculated as

$$J_{\mathrm{BB}} = \sum\sum w_{ij}Y_{ij}, \quad \text{where } Y_{ij} = x_i x_j \tag{5}$$

with w_{ij} a weight with value 1 if two samples, x_i and x_j, are adjacent ('joined') and 0 for all other cases. The count of black–white (1–0) joins is calculated as

$$J_{\mathrm{BW}} = \sum\sum w_{ij} Y^*_{ij}, \quad \text{where } Y^*_{ij} = 1 - \delta(x_i, x_j) \tag{6}$$

with δ the Kronecker delta function. The significance of the join-count statistics is achieved by computing a standard normal deviate (i.e. subtracting the mean and dividing by the **standard error**) using a two-tailed test to detect positive or negative spatial autocorrelation. Join-count statistics have been extended to multiple category data [10].

Spatial Autocorrelation Coefficients

Spatial intensity and scale of quantitative data from adjacent or noncontiguous sampling units (e.g.

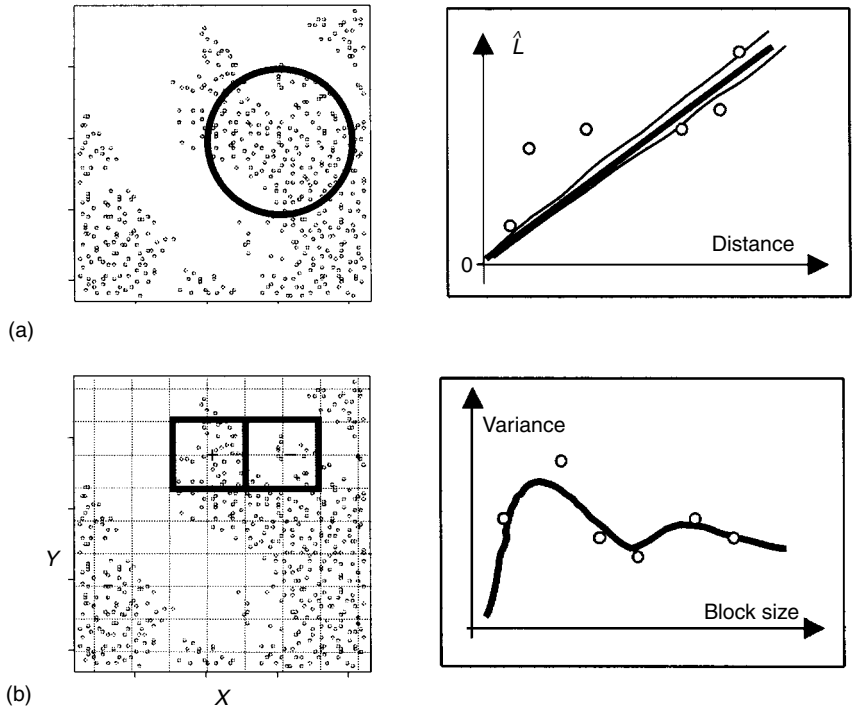

Figure 1 (a) The left graph is the map of the $x-y$ coordinates of point data (e.g. trees) where the circle is Ripley's K search window of radius t used to count the observed number of points. The right graph is the plot of observed $\hat{L}(t)$ (plotted as small open circles) against distance (t). The expected $L(t)$ is depicted by the heavy diagonal line. The two lines above and below the diagonal line indicate the confidence envelope. At small distances the spatial pattern is significantly aggregated, while at large distances the pattern is significantly segregated. (b) The left graph illustrates how point data (e.g. trees) can be sampled using contiguous sampling units. The rectangle is the TTLQV search window of block size $b = 2$ (see (4)). The right graph is the plot of the TTLQV variance against the block size. The first peak indicates the scale of the first spatial aggregation of the data while the second weaker peak at larger distance indicates the periodicity of the spatial structure. (c) The left graph illustrates how point data (e.g. trees) can be sampled using contiguous sampling units. The circle illustrates the spatial distance class used in Moran's I to compute spatial autocorrelation at increasing spatial lags, d. Here, the observed Moran's I at the spatial lag of $d = 2$ is computed using only the search region between the dotted circle and the next solid circle (i.e. $d > 1$ and $d = 2$). The right graph is the spatial correlogram where for small distances the spatial autocorrelation is positive and significant. The patch size is found just after the second distance class, i.e. when the values of spatial autocorrelation switch from positive values at short distances to negative values at intermediate and large distances. (d) The left graph is as in (c). The right graph is the experimental variogram where the range (a) is, as for Moran's I, between the second and third distance classes, hence where the sill starts. The nugget effect is almost zero

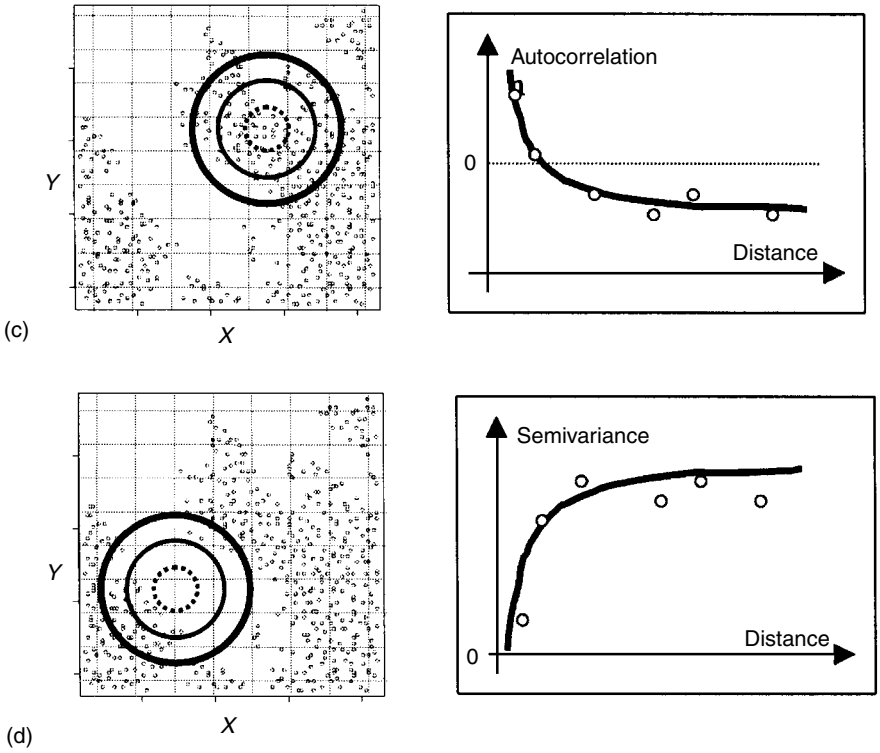

Figure 1 (*Continued*)

quadrats) can be estimated using spatial autocorrelation coefficients such as Moran's I and Geary's c [3, 5, 7]. Moran's I computes the degree of correlation between the values of a variable as a function of spatial lags. This coefficient is structurally comparable to a Pearson's product–moment correlation coefficient and computes the deviation between the values of the variable and its mean (see below). Moran's I varies from -1 (negative autocorrelation) to 1 (positive autocorrelation), with an expected value close to zero in the absence of spatial autocorrelation [specifically, $E(I) = -(n-1)^{-1}$]. Geary's c, on the other hand, measures the difference among values of a variable at nearby locations. It behaves somewhat like a distance measure and varies from 0 for perfect positive autocorrelation, to about 2 for a strong negative autocorrelation. In the absence of significant spatial autocorrelation, the expected value $E(c)$ is 1.

Moran's I coefficient is computed for increasing distance class of spatial lag d:

$$I(d) = \frac{\dfrac{\sum\sum w_{ij}(d)(x_i - \bar{x})(x_j - \bar{x})}{W(d)}}{\dfrac{\sum (x_i - \bar{x})^2}{n}} \tag{7}$$

while Geary's c coefficient is

$$c(d) = \frac{\dfrac{\sum\sum w_{ij}(d)(x_i - x_j)^2}{2W(d)}}{\dfrac{\sum (x_i - \bar{x})^2}{(n-1)}} \tag{8}$$

where $w_{ij}(d)$ are elements of a weight matrix for which a value of 1 indicates that a pair of two samples, x_i and x_j, are in the same distance class d and a value of 0 indicates all other cases. $W(d)$ is the sum of $w_{ij}(d)$. Outlier values affect the estimation of spatial autocorrelation by these coefficients. In the case of Moran's I, outlier values will bias estimation of the mean, and therefore either

under- or overestimate the calculation of spatial autocorrelation based on the deviation to that biased mean. In the case of Geary's c, biases occur when outlying values of squared differences with other values have more weight on the estimated values of spatial autocorrelation.

The plot of Moran's I or Geary's c against the distance classes d is called a *spatial correlogram* (Figure 1c). The distance at which the value of spatial autocorrelation crosses the expected value, E(I) or E(c), indicates the range of the patch size or simply the spatial range of the pattern. To assess statistical significance, a test must take into account the fact that individual coefficients of a spatial correlogram are not independent of one another. Therefore, significance is assessed by using a Bonferroni method that approximates the adjusted significance probability for multiple testing, in this case multiple distance classes [7]. Under the Bonferroni adjustment, the probability level, α', used to test the entire spatial correlogram is determined by dividing the probability level, say $\alpha = 0.05$, by the number of distance classes, say $k = 6$ (e.g. $\alpha' = \alpha/k = 0.05/6 = 0.0083$). A correlogram is deemed significant if the significance level of at least one individual coefficient is lower than the α' level.

Spatial autocorrelation coefficients provide an averaged isotropic estimation of spatial autocorrelation intensity at each distance class. When working with plant spatial structures, however, most species will show some degree of spatial anisotropy due to differential response to environmental conditions. Spatial anisotropy can be detected by calculating and comparing the spatial autocorrelation for pairs of locations grouped not only by distance class but also by direction. Membership to the distance and direction classes is given by the weight matrix. Anisotropy is identified when patch sizes vary according to direction.

The spatial autocorrelation coefficients described above are computed for the entire study plot, producing global statistics. There are cases, however, where local estimation of the intensity of spatial autocorrelation may reveal interesting insights on the local spatial processes. One can then use a *local index of spatial association* (LISA) [1]. For example, both Moran's coefficient and Geary's c can be calculated at each site i separately to give indices of local association or autocorrelation [1].

Experimental Variogram

Ecologists have shown an increasing interest in geostatistical methods to identify and model spatial pattern [3, 6, 9]. One begins by estimating parameters that characterize the spatial structure of the data in terms of spatial variance using an experimental **variogram**, and then uses these parameters to interpolate values at unsampled locations via kriging.

Spatial pattern intensity and range are estimated from the experimental variogram (Figure 1d), which estimates the *spatial semivariance function* $\gamma(d)$:

$$\hat{\gamma}(d) = \frac{\sum\sum w_{ij}(d)(x_i - x_j)^2}{2W(d)} \qquad (9)$$

Similar to Geary's c autocorrelation coefficient, the semivariance is a distance function; the difference between these two coefficients lies mainly in the fact that the semivariance lacks a denominator that standardizes the spatial autocorrelation estimation. The semivariance is therefore not bounded, making comparisons among variables difficult. Geary's c and the semivariance are alike in their response to outlier values since they are both based on the squared differences among the values of a given variable.

As for the spatial correlogram, the shape of the variogram obtained with sampled data – the experimental variogram – allows the description of the overall spatial pattern and the estimation of spatial autocorrelation parameters: (a) the *spatial range*, a, where the variable is spatially influenced by the same underlying process; (b) the *nugget effect*, C_0, which is the estimate of the error inherent in the measurements (sampling design and sampling unit size) and environmental variability; and (c) the *sill* $(C_0 + C_1)$ that quantifies the spatial pattern intensity, where C_1 is the spatial variance component [3, 6, 9].

Theoretical Variogram

According to the shape of the experimental variogram, different theoretical variogram models can be used to predict values at unmeasured locations, which is known as **kriging**. The most commonly used models include: the linear model, the exponential model, the spherical model, and the Gaussian model. Since most of the spatial variance signal is in the first part of the variogram (up to the spatial range), high-quality parameter fitting for short-distance lags is

very important. Generalized least squares (*see* **Least squares, general**), **maximum likelihood estimation**, and restricted maximum likelihood methods can be used to make the choice of theoretical variogram models more objective [3, 6]. Optimal estimation of these parameters is crucial for the subsequent kriging steps and the resulting interpolated map. Hence, a spatial range that is too short will result in a spiky map: spatial autocorrelation structure will be included only in interpolated locations near the sampling points and all the other locations will have the mean intensity value (the modeled sill value). Too large a spatial range will conversely result in a very smooth map.

Kriging

The kriging procedure, also known as BLUP (best linear unbiased prediction), returning the observed values at sampling locations, interpolates values using the intensity and shape of the experimental and modeled variogram of the data (using a neighborhood and/or distance search radius), and provides the standard errors of the interpolated values [3, 6, 9]. These prediction errors have often been used to optimize sampling design by identifying areas where sampling effort should be increased or decreased. These errors are, however, a function of the selected theoretical variogram model and not of the raw data. Caution should therefore be taken in interpreting the meaning of these estimation errors when optimizing sampling design. Kriging is one type of spatial interpolator. Other spatial interpolation techniques exist that also produce smoothed interpolated maps (see Tables 1 and 2).

Mantel Test

The spatial methods presented above compute the spatial pattern for a single variable at a given time or compare the relationship between the spatial patterns of two variables. In some studies, it is interesting to estimate the spatial structure of a set of variables (e.g. in a community study). This can be achieved by using the *Mantel test* [5, 7], which is a linear estimate of the relationship between the two square distance matrices based of the degree of relationship of two sets of variables taken at the same sampling locations. The Mantel statistic, Z, sums the products between corresponding elements of the distance matrices:

$$Z = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} B_{ij} \quad \text{for } i \neq i \tag{10}$$

where A is the variable distance matrix and B comprises the actual Euclidean (spatial) distances among the n sampling units (e.g. quadrats). The Mantel statistic, Z, can be normalized into a product–moment correlation coefficient, r, that varies from -1 to 1. The normalized Mantel statistic corresponds to an averaged isotropic intensity of spatial autocorrelation for this set of variables over the entire study plot.

Significance is assessed either by using an **exact randomization technique** to construct a reference distribution, or by using an asymptotic t-approximation [7]. Given that the relationship between the two matrices is based on distance values rather than on the raw data, the degree of relationship, here spatial autocorrelation, is often weaker than what would have been expected using the raw data. One should therefore not be too concerned about gauging the intensity of the relationship, but rather whether or not significance is indicated.

Another way to compute the spatial autocorrelation with the Mantel test is to first classify the Euclidean distances according to spatial distance classes (d), as in Moran's I, and then to compute the normalized Mantel statistic, the variable distance matrix A and this new spatial distance class matrix. The result is a multivariate spatial correlogram (see [7] for more details).

More Spatial Statistics

The field of spatial statistics is a very active domain of research in ecology. Many new methods (see Tables 1 and 2) have been developed or modified in the last few decades, and to have summarized them here is beyond the scope of this entry. More mathematical details can be found in the literature (among others, [1]–[10]).

References

[1] Anselin, L. (1995). Local indicators of spatial autocorrelation – LISA, *Geographical Analysis* **27**, 93–115.

[2] Burrough, P.A. & McDonnell, R.A. (1998). *Principles of Geographical Information Systems*, Oxford University Press, Oxford.

[3] Cressie, N.A.C. (1993). *Statistics for Spatial Data*, Wiley, New York.

[4] Dale, M.R.T. (1999). *Spatial Pattern Analysis in Plant Ecology*, Cambridge University Press, Cambridge.

[5] Fortin, M.-J. (1999). Spatial statistics in landscape ecology, in *Landscape Ecological Analysis. Issues and Applications*, J.M. Klopatek & R.H. Gardner, eds, Springer-Verlag, New York, pp. 253–279.

[6] Goovaerts, P. (1997). *Geostatistics for Natural Resources Evaluation*, Applied Geostatistics Series, Oxford University Press, New York.

[7] Legendre, P. & Legendre, L. (1998). *Numerical Ecology*, 2nd English Edition, Elsevier, Amsterdam.

[8] Perry, J.N. (1995). Spatial analysis by distance indices, *Journal of Animal Ecology* **64**, 303–314.

[9] Rossi, R.E., Mulla, D., Journel, A.G. & Franz, E.H. (1992). Geostatistical tools for modeling and interpreting ecological spatial dependence, *Ecological Monographs* **62**, 277–314.

[10] Upton, G.J.G. & Fingleton, B. (1985). *Spatial Data Analysis by Example,* Vol. 1: *Point Pattern and Quantitative Data*, Wiley, New York.

(*See also* **Landscape ecology**; **Multivariate kriging**; **Poisson cluster process**; **Space–time covariance models**; **Variogram estimation**)

MARIE-JOSÉE FORTIN, MARK R.T. DALE & JAY VER HOEF