7

Spatial relationship and multiscale analysis

Introduction

To determine the factors and processes that are responsible for the spatial pattern observed in ecological data, both correlative methods (this Chapter) and inferential methods (mostly regression; Chapter 8) can be used. The ecological data of interest can be univariate (a single variable, such as species abundance or plant height, and so on) or multivariate (several variables, such as species composition or environmental variables, and so on). With multivariate data, as with univariate data, the first step is always to evaluate the variables' spatial autocorrelation and to determine its significance. This can be achieved by using modified spatial autocorrelation coefficients or semi-variance functions for distance matrices based on the multivariate data, instead of equivalent measures for the original observations themselves (Wartenberg 1985; Wackernagel 2010; Dray *et al.* 2012; Legendre & Legendre 2012). Many current ecological studies are carried out over large study areas (risking non-stationarity) often combining several data types. Researchers need therefore to perform Exploratory Spatial Data Analysis (ESDA) to determine the key spatial scale(s) at which statistical analyses should be performed (Fortin *et al.* 2012a).

In this chapter, we present methods that quantify the spatial relationship between spatially autocorrelated variables (both univariate and multivariate) using correlation in the spatial context, and we then show how multiscale analysis can depict the spatial scales relevant to the spatial structure of the ecological data. In Chapter 8, which follows, we will summarize the key methods that are used to model spatial causality between spatially autocorrelated variables. The spatio-temporal analysis of such data will be covered in Chapter 11.

7.1 Correlation between spatially autocorrelated variables

To determine which environmental factors influence the variability of ecological data, it is common-place to compute the correlation among variables. Yet being spatially explicit, some of these variables may show various degrees of spatial autocorrelation. As we have mentioned before, in the presence of positive autocorrelation, the information provided by each of n observations or samples may be less than that from n independent observations, thus giving less than a full degree of freedom from each because of the dependency among the values (here due to spatial autocorrelation). It is therefore important to estimate the 'effective sample size' provided by n dependent observations, call it n', which is usually less than the nominal sample size, n, for independent data. The effective sample size for bivariate tests, such as correlation, can be estimated by adjusting for the amount of autocorrelation in the two variables (Haining 1990; Clifford *et al.* 1989; Dutilleul 1993a, b; Fortin & Payette 2002). In Chapter 8, we provide a more detailed account of how to compute the effective sample size, but here we present the basic principles.

The adjustment requires first computing the degree of spatial autocorrelation for each variable, call the two variables x and y, in k distance classes, producing a Moran's I spatial correlogram. Then we determine estimates of the covariance values for each of the k distance classes for each variable (call the matrices that contain them $\mathbf{S}_x$ and $\mathbf{S}_y$). Using a weighted connectivity matrix, $\mathbf{B}$, with $b_{ii} = 1/n - 1/n^2$ on the main diagonal and $b_{ij} = -1/n^2$ everywhere else, the effective sample size is:

$$n' = 1 + \frac{\mathrm{tr}(\mathbf{BS}_x)\mathrm{tr}(\mathbf{BS}_y)}{\mathrm{tr}(\mathbf{BS}_x\,\mathbf{BS}_y)}, \tag{7.1}$$

where tr is the trace of the matrix, the sum of the elements in the main diagonal. This is the Dutilleul (1993a, b) correction algorithm, which is more accurate than the original correction proposed by Clifford *et al.* (1989), especially when the nominal sample size is small.

When testing the significance of the linear correlation between two autocorrelated variables, the number of degrees of freedom is therefore n' – 2 instead of n – 2. As Fortin & Payette (2002) have shown, the estimated effective sample size is directly related to the estimated strength of spatial autocorrelation, and will vary depending on the number of distance classes used: fewer distance classes (larger distance class intervals) result in less spatial autocorrelation (i.e. less dependency is detected) such that the n' will be closer to n, while more distance classes (smaller distance class intervals) result in greater spatial autocorrelation (i.e. more dependency is detected) so that n' will be much smaller than n.

For the purpose of illustration, consider the correlation of species and environmental data sampled at 12 locations as depicted in Figure 7.1*a*. The spatial autocorrelation in the first distance class of 1.41 units is 0.311 ($p = 0.01$) for the species and –0.013 ($p = 0.30$) for the environmental variable. The linear correlation between the two variables is 0.601. When the significance of the correlation is assessed using the conventional n – 2 degrees of freedom (12 – 2 = 10), the probability is 0.038, leading us to believe that the correlation is statistically significant. However, the effective sample size, based on Dutilleul's correction, is 8.38, a value obviously smaller than n = 12. Using this effective sample size, the number of degrees of freedom, n' – 2, is 6.38 and the corresponding probability is 0.103, leading to the interpretation that the correlation is not significant. This simple example demonstrates clearly how the standard statistical interpretation would be inappropriate if data are spatially autocorrelated. Therefore researchers should always check first whether the data are spatially autocorrelated and then, if so, compute the appropriate effective sample size in order to assess the significance of the correlation.

7.2 Correlation of distance matrices

7.2.1 Mantel test

Whenever we are interested in summarizing the spatial pattern of species assemblages, or other multivariate data, with a single number, the spatial autocorrelation coefficients or semi-variance functions can be written as a cross-product between two matrices (Getis 1991; Dale *et al.* 2002). This general approach was first proposed by Mantel (1967) to quantify the spatial (x_i) and temporal (t_i) relationships of multivariate data from the same locations. The null hypothesis for a Mantel test is that the distances in some matrix $\mathbf{X}$ (based on species abundances at a set of sites, for example) are independent of the distances in another matrix $\mathbf{Y}$ (based on environmental variables, for example) sampled at the same locations. The Mantel statistic computes a linear relationship between two symmetric matrices:

$$Z_{\mathrm{M}} = \sum_{\substack{i=1 \\ i \neq j}}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} x_{ij}\, y_{ij}. \tag{7.2}$$

Z_{M} is the Mantel statistic, and x_{ij} and y_{ij} are similarity, dissimilarity, Euclidean distance, or connectivity matrix ($\mathbf{X}$, $\mathbf{Y}$) values of all the variables at each sampling location (Figure 7.2). The advantage of this matrix approach is that any dissimilarity or distance coefficient can be used (see Legendre & Legendre 2012 for a full list of possibilities), and this may be chosen to reflect better the community structure of the variables

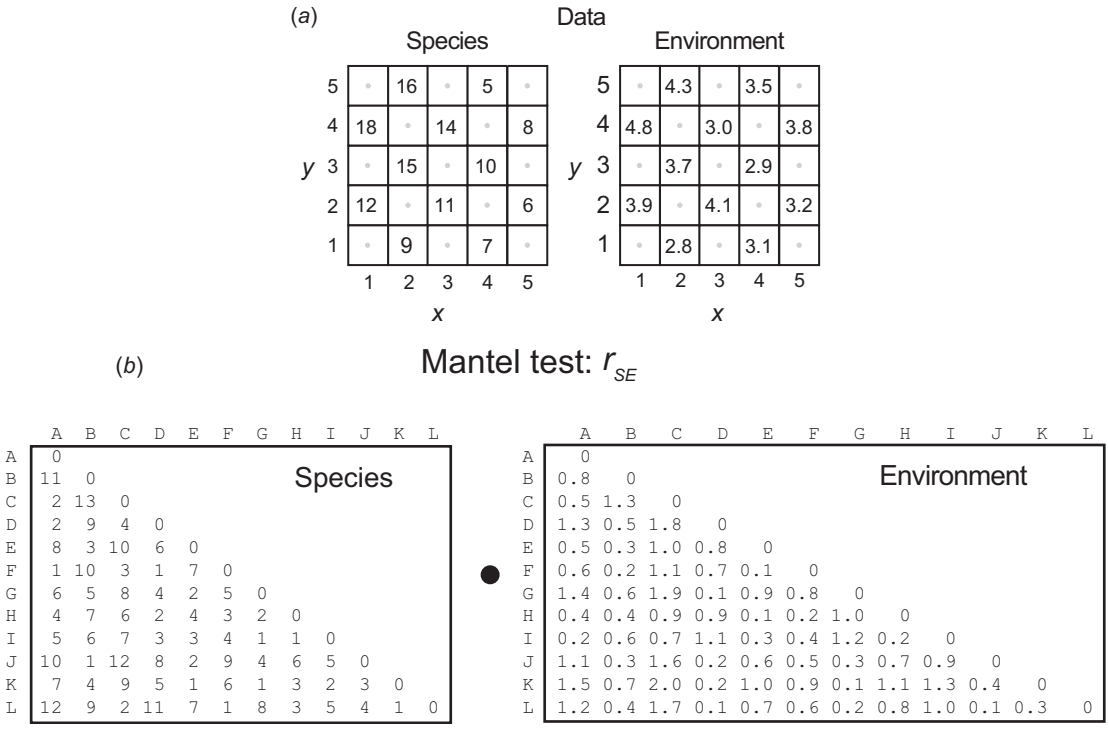


Figure 7.1 Correlation between spatially autocorrelated data. (*a*) This figure illustrates two variables (species and environment) sampled at the same 12 sampling locations. (*b*) The data converted into symmetrical Euclidean distance matrices (12 × 12) which are used to compute the standardized Mantel statistic, r_{SE}. Only the values of the lower triangulation part of the symmetrical distance matrices are shown. The '•' is the Hadamard product which is an element-by-element matrix multiplication.

of interest. Equation (7.2) can be rewritten in matrix notation:

$$Z_{\mathbf{XY}} = \mathbf{X} \bullet \mathbf{Y}. \tag{7.3}$$

$Z_{\mathbf{XY}}$ is the sum of all of the element-by-element products as indicated by '•' which is known as the Hadamard product between the distance matrices **X** and **Y** (Figures 7.1*b* and 7.3).

The measure $Z_{\mathbf{XY}}$ is unbounded, and to facilitate its interpretation, each matrix can be standardized before computing the Mantel correlation statistic to obtain a bounded r_{M} statistic, ranging from −1 to 1:

$$r_{\mathrm{M}} = \frac{1}{d-1} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left(\frac{x_{ij} - \overline{x}}{S_x} \right) \left(\frac{y_{ij} - \overline{y}}{S_y} \right), \tag{7.4}$$

where d is the number of elements in the lower (or upper) triangular part of each matrix, which is $n(n-1)/2$. From Eq. (7.4), it is evident that r_{M} in fact computes Pearson's linear correlation coefficient between the elements of two matrices and, as such, the Mantel test and the Pearson correlation values are closely related. Legendre & Fortin (2010) showed that the Mantel test is approximately the square root of the coefficient of determination R^2_{M} of the linear regression between the elements of the matrices. It is not a perfect equivalence because the sample size is n in the correlation, and $n(n-1)/2$ in the Mantel test (Figure 7.2). Also, the Mantel computation is performed using distance values rather than the raw data themselves, hence r_{M} values are not directly comparable to those of a Pearson's correlation coefficient (Dutilleul *et al.* 2000), but considerably weaker.

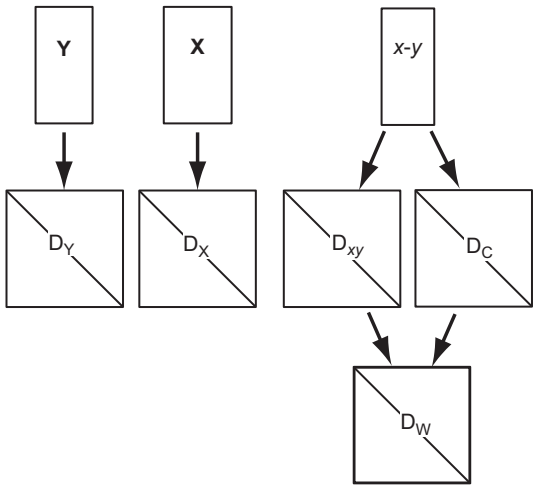


Figure 7.2 Flow chart illustrating how from the raw multivariate data, **X** (D_X), **Y** (D_Y) and *x*–*y* coordinates (Euclidean distance D_{xy}, binary connectivity matrix D_C, weighted connectivity matrix D_W such as a distance class interval matrix), square matrices can be obtained.

Therefore it is important to test the significance of the Mantel statistic, because it can seem quite small in value but still significant. The common outcome of having significant Mantel statistics that have small values can also arise due to the large sample sizes for the matrices.

Mantel statistics can be tested either by an approximate *t*-test when the sample size is large or by a randomization test when the sample size is small. The randomization is in a restricted form, so that it retains the relationships between pairwise distances measured between sampling locations; this is accomplished by randomly shuffling the rows and columns of one of the matrices (Fortin & Gurevitch 2001; Legendre & Legendre 2012). When the null hypothesis is true, the observed Mantel's statistic is expected to have a value near the middle of the reference distribution. With a significant relationship between the two matrices, whether positive or negative, the observed Mantel statistic is expected to be more extreme, either higher or lower, compared to the reference distribution. The precision of the probability value is directly related to the number of randomization iterations used. We recommend generating at least 10 000 randomizations (Manly 2006; Fortin *et al.* 2012*b*) where the observed statistic is included in the reference distribution (Hope 1968).

Taking the data illustrated in Figure 7.1 and computing the Mantel statistic based on Euclidean distances of the species and environmental variables, the Mantel statistic Z_M is 534 and the r_M is 0.31 ($p = 0.023$). Given that we are dealing with univariate data here, we can see how the r_M (0.31 for $n =$ 66 matrix elements) is related to the Pearson's correlation coefficient (0.601, $n = 12$ samples) and r^2 (0.36). For comparison, let us now compute Mantel statistics for multivariate data, in particular those depicted in Figure 7.4, based on the Euclidean distance matrices and a distance class matrix (Figure 7.5). The linear correlation between the species and the environmental variables is 0.37 but $p = 0.201$, which is not significant for the correlation between the two distance matrices. Using a real data set, with a tree species matrix, **X** (Euclidean distances between tree abundance data of 14 species; see Fortin 1992 for details), and a topography matrix, **Y** (Euclidean distances between relative elevations in metres), the r_M between them is both positive (0.26) and significant (Table 7.1). The Mantel test can also be computed between the same tree species matrix, **X**, and the geographical distance matrix, **Z** (Euclidean distances between sampling locations). In that case, r_M is also positive (0.23) and significant and corresponds to an averaged isotropic intensity of spatial structure for the tree species community for the entire study area (Table 7.1). Note that there is a strong positive relationship between the relative topography and the sample locations: 0.839. We will come back to this issue below.

There is a whole list of 'pros and cons' for the use of the Mantel test. The reason that the Mantel test is so popular is the fact that until recently it seemed to be the only test that allowed the computation of the correlation between two sets of multivariate data, based on a single synthetic distance measure (see Legendre & Fortin 2010 for suggestions of other

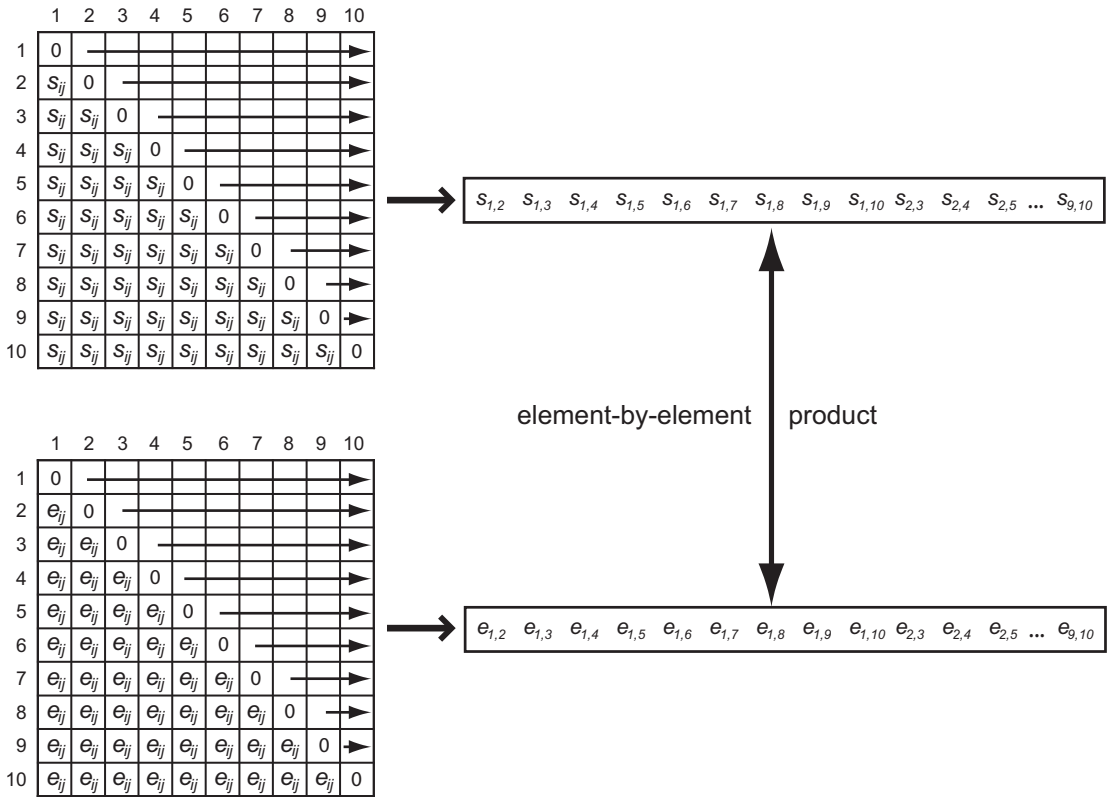
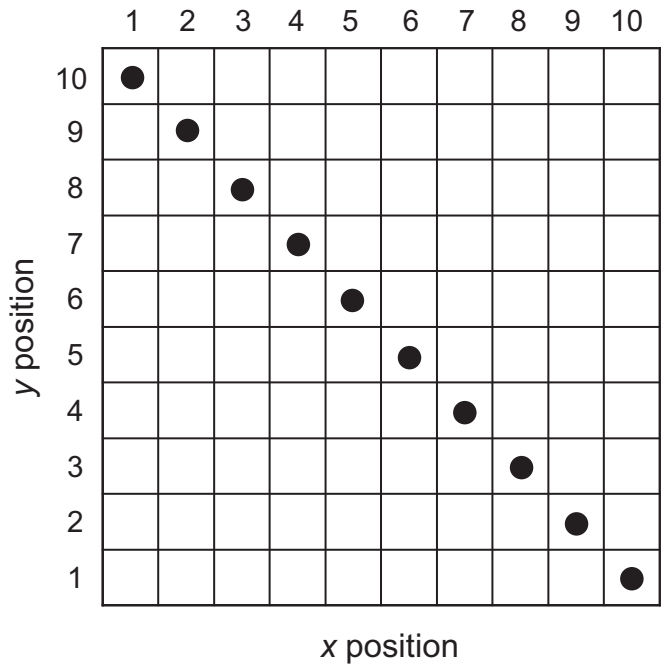


Figure 7.3 Illustration how the element-by-element matrix multiplication (Hadamard product) is performed to compute the Mantel statistic. The square distance matrices on the left can be unfolded into vectors from which the Hadamard product is computed between the elements of the $D_{Species}$ (s_{ij}) and the $D_{Environment}$ (e_{ij}) matrices.

methods). One main weakness of this approach is that we cannot determine how any one variable from either multivariate data set is related to any one variable in the other multivariate data set. Also, the Mantel test computes the linear correlation between the distances, even if the relationship is not actually linear (we can check this by plotting a scattergram of the two distance matrices), in which case the Mantel test is not a good evaluation of the relationship. To address this nonlinear relationship, Diezt (1983) suggested converting the distance measures to ranks before computing the cross-product. Thus, any monotonic relationship between the two matrices, even if it is not linear, can be estimated using the Mantel test; this is now equivalent to performing a Spearman correlation analysis rather than the Pearson approach.

Regardless of the type of relationship computed, linear or merely monotonic, the Mantel test is based on distance values rather than on the raw data, and the magnitude of the effect is often weaker than with the raw data, as illustrated above with the univariate data example (Figure 7.1). But what is really computed? It is the relationship between the distance measures, and pairs of sampling locations can have the same degree of dissimilarity either because they both have either high or low values. Consequently, the magnitude of r_M should not be used as the equivalent of the Pearson or Spearman correlation of the raw data. Furthermore, given that the significance



	Species			Environment			Coordinates	
	Spp.1	Spp.2	Spp.3	Soil	Moisture		*x*	*y*
1	10	11	0	1.1	0.15	1	1	10
2	50	8	0	1.4	0.20	2	2	9
3	60	6	1	2.2	0.25	3	3	8
4	70	5	1	2.5	0.30	4	4	7
5	25	9	0	2.6	0.15	5	5	6
6	75	7	1	3.3	0.35	6	6	5
7	80	2	1	3.7	0.40	7	7	4
8	80	3	1	5.5	0.45	8	8	3
9	75	2	2	5.6	0.50	9	9	2
10	90	1	2	5.9	0.55	10	10	1

Figure 7.4 Multivariate data of species abundance (three species) and environmental data (two variables) sampled at ten locations (*x*–*y* coordinates).

of the observed r_M is tested against the restricted randomization of the rows and columns of distance matrices (Mantel 1967), it can be significant even if the value is very small (quite close to zero) because the sample size for distances is much higher than the original number of sampled locations, *n*.

In addition, a misconception has developed over the years that the Mantel test can correct for the presence

Species matrix

	1	2	3	4	5	6	7	8	9	10
1	0	40	50	60	15	65	71	70	66	81
2	40	0	10	20	25	25	31	30	26	41
3	50	10	0	10	35	15	20	20	16	30
4	60	20	10	0	45	5	10	10	6	20
5	15	25	35	45	0	50	55	55	51	66
6	65	25	15	5	50	0	7	6	5	16
7	71	31	20	10	55	7	0	1	5	10
8	70	30	20	10	55	6	1	0	5	10
9	66	26	16	6	51	5	5	5	0	15
10	81	41	30	20	66	16	10	10	15	0

Environmental matrix

	1	2	3	4	5	6	7	8	9	10
1	0	0.3	1.1	1.4	1.5	2.2	2.6	4.4	4.5	4.8
2	0.3	0	0.8	1.1	1.2	1.9	2.3	4.1	4.2	4.5
3	1.1	0.8	0	0.3	0.4	1.1	1.5	3.3	3.4	3.7
4	1.4	1.1	0.3	0	0.2	0.8	1.2	3.0	3.1	3.4
5	1.5	1.2	0.4	0.2	0	0.7	1.1	2.9	3.0	3.3
6	2.2	1.9	1.1	0.8	0.7	0	0.4	2.2	1.8	2.2
7	2.6	2.3	1.5	1.2	1.1	0.4	0	1.8	1.9	2.2
8	4.4	4.1	3.3	3.0	2.9	2.2	1.8	0	0.1	0.4
9	4.5	4.2	3.4	3.1	3.0	1.8	1.9	0.1	0	0.3
10	4.8	4.5	3.7	3.4	3.3	2.2	2.2	0.4	0.3	0

Euclidean distance matrix

	1	2	3	4	5	6	7	8	9	10
1	0	1.4	2.8	4.2	5.7	7.1	8.5	9.9	11.3	12.7
2	1.4	0	1.4	2.8	4.2	5.7	7.1	8.5	9.9	11.3
3	2.8	1.4	0	1.4	2.8	4.2	5.7	7.1	8.5	9.9
4	4.2	2.8	1.4	0	1.4	2.8	4.2	5.7	7.1	8.5
5	5.7	4.2	2.8	1.4	0	1.4	2.8	4.2	5.7	7.1
6	7.1	5.7	4.2	2.8	1.4	0	1.4	2.8	4.2	5.7
7	8.5	7.1	5.7	4.2	2.8	1.4	0	1.4	2.8	4.2
8	9.9	8.5	7.1	5.7	4.2	2.8	1.4	0	1.4	2.8
9	11.3	9.9	8.5	7.1	5.7	4.2	2.8	1.4	0	1.4
10	12.7	11.3	9.9	8.5	7.1	5.7	4.2	2.8	1.4	0

Distance class matrix

	1	2	3	4	5	6	7	8	9	10
1	0	1	2	3	4	5	6	7	8	9
2	1	0	1	2	3	4	5	6	7	8
3	2	1	0	1	2	3	4	5	6	7
4	3	2	1	0	1	2	3	4	5	6
5	4	3	2	1	0	1	2	3	4	5
6	5	4	3	2	1	0	1	2	3	4
7	6	5	4	3	2	1	0	1	2	3
8	7	6	5	4	3	2	1	0	1	2
9	8	7	6	5	4	3	2	1	0	1
10	9	8	7	6	5	4	3	2	1	0

Figure 7.5 Corresponding Euclidean distance matrices of the species data, environmental variables, and *x*–*y* coordinates of Figure 7.4 as well as the distance class matrix where the spatial lag interval is 1.41 units.

of spatial autocorrelation in the data. This is not true *per se*; what the Mantel test can do is it can account for the relative distance between the sampled locations. In fact, the Mantel test assumes that there is no dependence among the distances, and that will not be the case in the presence of a spatial structure in the distance values. We will have to return to this issue in the next section, which discusses partial Mantel tests.

Table 7.1 Standardized Mantel (r_M) and partial Mantel ($r_{\mathbf{XY.Z}}$) values between the tree community data and the relative topography, controlling for location as the x–y coordinates of the sampling units ($n = 84$) using Euclidean distances. The probability is based on a reference distribution having 9999 randomizations plus the observed statistics

X	**Y**	**Z**	r_M	p for r_M
Tree	Topography		0.255	0.0001
Tree		Location	0.232	0.0001
	Topography	Location	0.839	0.0001
Tree	Topography	Location	0.113	0.0391

As a final comment in this section, we should describe the extension of the Mantel test approach to the calculation of a *Mantel correlogram*, which depicts the value of the Mantel statistic plotted against lag distance classes (i.e. the Mantel correlogram is a multivariable version of a spatial correlogram such as a Moran's I correlogram). The Euclidean distance matrix can be converted into a connectivity matrix, such as a distance class matrix (just as we do with Moran's I, where estimates of spatial autocorrelation are made at several distance classes), to compute a Mantel (multivariate) correlogram (Figure 7.6; Oden & Sokal 1986; Legendre & Fortin 1989; Borcard & Legendre 2012). In a recent simulation study, Borcard & Legendre (2012) showed that the Mantel correlogram's significance level is valid as calculated; that is, that the Type I error is correct. Hence the Mantel correlogram can be used to determine the spatial structure at increasing spatial lags. Using the same tree assemblage data mentioned above, a trend is identified (Figure 7.6) in our example by the Mantel correlogram with a spatial range at about 60 m and where the r_M at the first distance class is 0.09 ($p < 0.001$).

7.2.2 Partial Mantel tests and multiple-matrix regression

Correlation, as it is computed for a Mantel test between the distances of two matrices, can be attributable to unmeasured variables. To determine these potentially spurious correlations, a partial correlation approach can be used. Such a partial correlation procedure can be applied to distance matrices by extending the Mantel two-matrix test to a partial Mantel three-matrix test (Smouse *et al.* 1986). The partial Mantel test quantifies the relationship between two matrices, **X** and **Y**, while controlling for the effects of a third one, **Z**. The partial Mantel statistic, $r_{\mathbf{XY.Z}}$, is computed by detrending (or controlling or removing) the linear effects of the variables in matrix **Z** on those in matrix **X** and in matrix **Y**, using a linear regression (Figure 7.7) for **X** on **Z**, and for **Y** on **Z**. Then, with the residuals from both regressions, $\text{Res}_{\mathbf{X}|\mathbf{Z}}$ and $\text{Res}_{\mathbf{Y}|\mathbf{Z}}$, as its basis, a Mantel test is performed as in Eq. (7.3). Another way to compute the partial Mantel is to compute partial correlation using all three matrices at once (Legendre & Legendre 1998):

$$r_{XY.Z} = \frac{r_{XY} - r_{XZ}\, r_{YZ}}{\sqrt{1 - r_{XZ}^2}\,\sqrt{1 - r_{YZ}^2}}. \tag{7.5}$$

By comparing r_M values, computed with two and three matrices (partial Mantel test, $r_{\mathbf{XY.Z}}$), we can test alternative causal relationships among the three matrices as we do in a path analysis (Legendre & Troussellier 1988; Legendre & Fortin 1989; Leduc *et al.* 1992; Cushman *et al.* 2006; Legendre & Legendre 2012). Furthermore, the third matrix, **Z**, can be a design matrix (with dummy variables corresponding to treatments and control locations, see Fortin & Gurevitch 2001) or a set of covariables, allowing us to test specific hypotheses by coding as in a contrast matrix in ANOVA (Sokal *et al.* 1993) or by using geographic location as a surrogate for non-sampled variables (Fortin & Payette 2002).

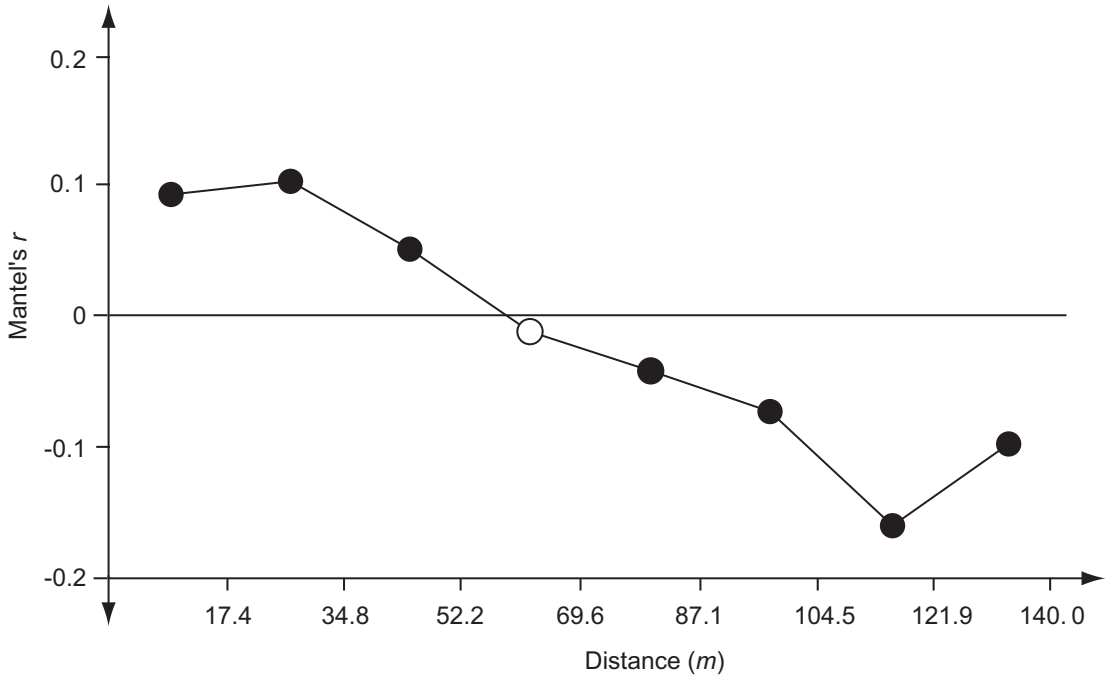


Figure 7.6 Mantel correlogram based on the abundance of 14 tree species where the standardized Mantel statistic, r_M, is plotted against the distance in metres. The overall spatial structure of the tree community is a trend with a spatial range of around 60 m.

Using the same univariate data already discussed (Figure 7.1), the partial Mantel test of the relationship between the species (S) and the environmental (E) variables controlling for the Euclidean distances between the sampling locations (L), $r_{SE.L}$, is 0.28 ($p = 0.045$). In comparison, the Mantel statistic between the species and environment, $r_M = r_{SE}$, was 0.31 ($p = 0.023$). Hence controlling for spatial distances between the sampling locations reduced the correlation between the dissimilarities of the two matrices, suggesting that one or two distance value equivalents are influencing Euclidean distances. Then, when computing the partial Mantel test with the multivariate tree data, as used to compute the Mantel test (Table 7.1), we can test the hypothesis that the influence of topography is closely related to the sampling locations. In doing so, the relationship between tree abundance and topography decreases to 0.113 and is barely significant at $\alpha = 0.05$. Consequently, we conclude that the topography influences tree abundances, but less so or less directly than we thought based on the Mantel test, because much of its influence can be accounted for by location.

When we control for the effects of a third matrix, **Z**, such as the Euclidean distance matrix for the sample locations, we are not controlling for the degree of spatial autocorrelation of the variables but only for the relative distance among the sampling locations. Furthermore, when the variables are strongly spatially autocorrelated, the restricted randomization (by rows and columns of the matrices) no longer provides unbiased outcomes, so that the significance of the partial Mantel test is not adequately evaluated (Oden & Sokal 1992; Guillot & Rousset 2013). This problem has been acknowledged recently, and different ways in which to restrict the randomization procedure have

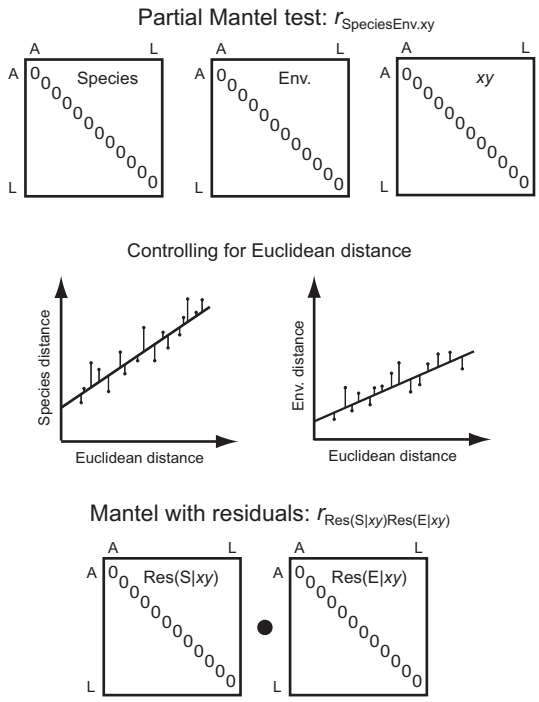


Figure 7.7 Partial Mantel test. This figure illustrates how the correlation between two variables (species and environment) sampled at the same 12 sampling locations can be computed by first factoring out the effect of a third matrix (here the *x*–*y* coordinates of the sampling locations), using linear regression and then computing a standardized Mantel statistic, $r_{\mathrm{Res}(S|xy)\ \mathrm{Res}(E|xy)}$, with the residuals of these two linear regressions. See text for more details.

been proposed (Dutilleul *et al.* 2000; Legendre 2000; Legendre & Fortin 2010). In essence, the problem stems from the lack of independence in the data so that the complete randomization procedure will produce inflated levels of Type I error (Cliff & Ord 1981). This is not unique to the Mantel test and the partial Mantel test, but affects most inferential tests, and we will expand on this topic in Chapter 8. Yet Oden & Sokal (1992) showed that when the spatial autocorrelation between the distance values of the matrix is weak, partial Mantel tests, as proposed by Smouse *et al.* (1986), can still be used because they are unlikely to reject the null hypothesis incorrectly.

The partial Mantel test provides a way to address causality among the matrices using a path analysis framework (Legendre & Troussellier 1988; Legendre & Legendre 2012). The partial Mantel, however, accounts for only one matrix and there are several cases where more matrix sets could explain the variability of species assemblage. Hence the partial Mantel has been extended to a regression approach, with the 'multiple regression on distance matrices' method (MRDM; Hubert & Golledge 1981; Smouse *et al.* 1986; Manly 2006; Legendre *et al.* 1994; Lichstein 2007; Zapala & Schork 2006; among others). Here there is a clear 'response' distance matrix ($\mathbf{D}_{\mathrm{Y}}$) and a series of predictor distance matrices ($\mathbf{D}_{\mathrm{X1}}$, $\mathbf{D}_{\mathrm{X2}}$, . . ., $\mathbf{D}_{\mathrm{X}n}$) that are unfolded into their corresponding vector form on which a multiple regression is then performed (Figure 7.8). The significance of the regression is tested by a randomization procedure. Legendre *et al.* (1994) suggested using a stepwise selection procedure to determine which predictor vectors are the most parsimonious to be retained in the regression. Nowadays, ecologists and evolutionary biologists mostly choose the Akaike information criterion (AIC) as a model selection tool (Akaike 1974; Burnham & Anderson 2002; Johnson & Omland 2004). AIC is a log-likelihood model that assumes a least-squares estimation with normally distributed errors from the residual sums of squares (RSS):

$$\mathrm{AIC} = n + n \times \log(2\pi) + n \times \log(\mathrm{RSS}/n) + 2(p+2), \tag{7.6}$$

where n is the sample size, p is the number of predictors in the model with the addition of two to account for the intercept and the error term. There are some issues, however, when using AIC to select predictor vectors: (1) it should not be used with linear mixed models; and (2) regression with vector dependence between the distances can occur, implying that the sample size in a MRDM is inflated to $n(n-1)/2$, and a correction should be used to account for the strength of dependence present (Van Strien *et al.* 2012).

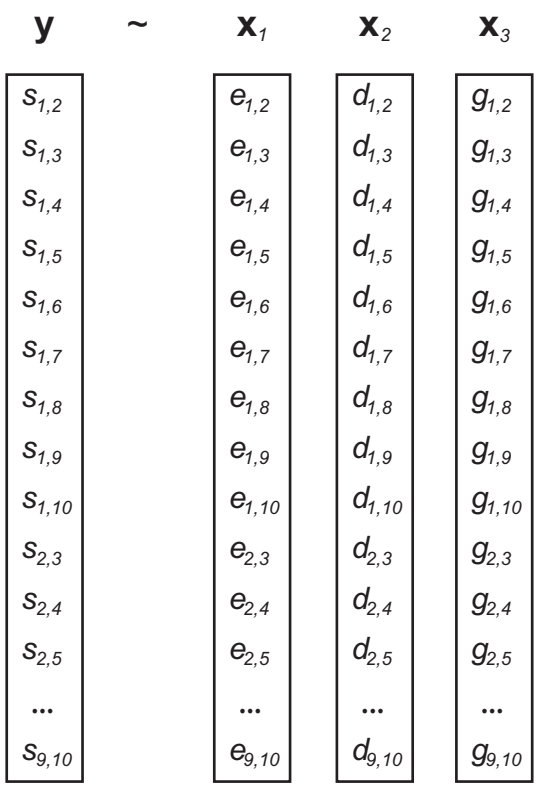


Figure 7.8 Multiple regression on distance matrices. After having unfolded each distance matrix as illustrated in Figure 7.3, the multiple regression of the response vector **y** of the distances of the species assemblage data against the predictor vectors ($\mathbf{x}_1$: distances of environmental data, $\mathbf{x}_2$: Euclidean distances, $\mathbf{x}_3$: distances of geomorphologic data) can be computed.

7.3 Canonical (constrained) ordination

One of the fundamental concerns with Mantel tests, which we have already mentioned, is that they should be used only to test hypotheses at the level of the distance matrix. Hence the Mantel test evaluates the correlation of the summary behaviour of multivariate data that has been collapsed into a single distance or dissimilarity value. The resulting correlation between the distances of the matrices is therefore a global outcome for all of the variables and it is not possible to identify which variable(s) contribute most to the strength of the correlation. To identify which variables contribute most to the relationship among sets of multivariate data, canonical ordination techniques such as canonical correspondence analysis, CCA, or redundancy analysis, RDA (Legendre & Legendre 2012) can be used, as we will illustrate in this section. Here the term 'canonical' refers to the relationship between two sets of variables. Redundancy analysis (RDA; Rao 1964; Legendre & Legendre 2012) is an ordination technique which corresponds to a multiple regression for multivariate data, which maximizes the variance explained by the linear relationship between the canonical axis of response matrix **Y** and the explanatory variables in matrix **X**. Therefore the RDA is also known as a direct ordination technique or a canonical ordination analysis because the relationship between two matrices is computed at the matrix level using the eigenvalue–eigenvector properties of the multivariate matrices. RDA is computed by the following equation, which uses eigenvalue notation from matrix algebra:

$$\left(\mathrm{S_{YX}}\mathrm{S_{XX}^{-1}}\mathrm{S'_{YX}} - \lambda_k \mathbf{I}\right)\mathbf{u}_k = 0, \tag{7.7}$$

where $\mathrm{S_{YX}}$ is the matrix of covariance between **Y** and **X** while $\mathrm{S'_{YX}}$ is its transpose, $\mathrm{S_{XX}^{-1}}$ is the inverse of the variance–covariance matrix of the predictor variables in **X**, and k is an index of the dimensions analysed. The other notation includes: **I** is the identity matrix, with 1s on the main diagonal and 0s elsewhere, λ_k is the kth eigenvalue, and $\mathbf{u}_k$ is the kth eigenvector. Any RDA ordination of the **X** axes is linearly constrained by the variables in **Y**. The best linear fit between the species and environmental ordination axes is obtained by an iterative procedure. RDA is, in essence, an extension of a principal component analysis (PCA) where the correlation between the two matrices is linear. Consequently RDA is a Euclidean distance representation of the objects in the multidimensional space.

Often the relationship between species and environmental variables is induced by other underlying factors

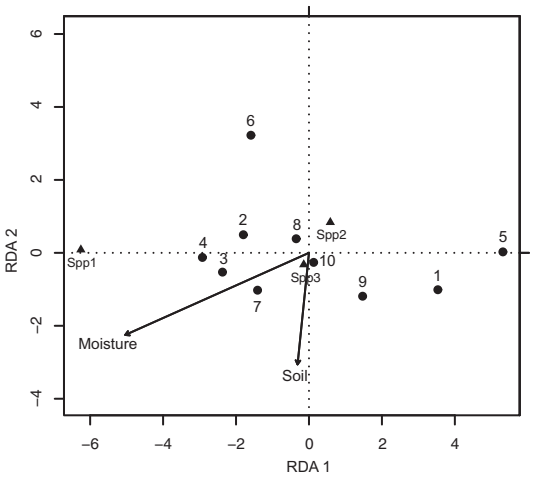


Figure 7.9 Partial RDA based on the 10 sampled data used in Figure 7.4: the relationship between species assemblage and the environmental variables are constrained by the *x*–*y* coordinates of the sampling locations. The first axis explained 60.44% of the variance and the second 0.026%. The plot depicts well that only Spp1 is related to the environmental variable 'moisture'. Dots and numbers indicate the positions of the sampling locations in the RDA plot.

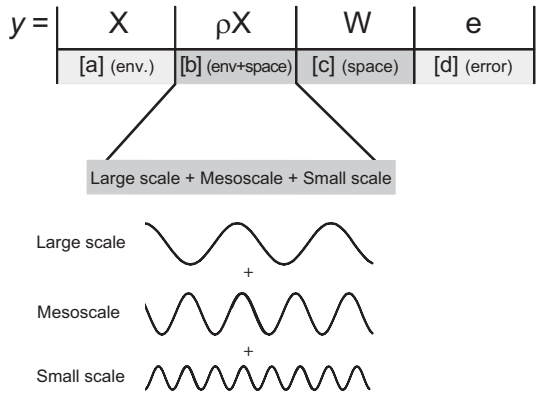


Figure 7.10 Variance partitioning analysis where the variability of the response variable *y* is explained by: (1) solely the environmental variables, X, and is known as the fraction [a]; (2) the spatially autocorrelated portion, ρ, of the predictor, environmental, variables [b]; and (3) [c]] 'space' which may be due to a mismatch between the spatial sampling design and the spatial scales that act on the response and predictor variables, or nonlinear relationships. Then the residual errors form the fraction [d]. Here we explicitly depict that the spatial autocorrelated structure can be due to the addition of spatial patterns (here for illustration purposes a series of stationary sine–cosine patterns according to the spatial scales along a transect) occurring at three spatial scales (local/small scale, landscape/mesoscale, and regional/large scale).

such as climate, topography, geomorphology, or historical events. One way to account for or to control for the effects of these other variables is to use partial canonical ordination techniques such as partial CCA or partial RDA (Borcard *et al.* 1992; Legendre & Legendre 2012). The advantage of such ordination and partial ordination techniques over the Mantel and partial Mantel tests is that the relationship is computed on the raw data values rather than on distance measures derived from them. Also, these techniques permit us to disentangle the relative contributions of each variable to the overall relationship between the two matrices **Y** and **X**.

The partial RDA of the multivariate data in Figure 7.4, relating the species assemblage and environmental variables while constrained by location in the form of *x*–*y* coordinates, is shown in Figure 7.9. The plot in Figure 7.9 clearly indicates that only the variable 'Spp1' is related to the environmental variable of moisture. Based on a randomization procedure, however, this partial RDA does seem to be significant ($F_{2,6} = 4.59$, $p = 0.059$, from 9999 randomizations). By next performing a partition of the variance into four components ([a] to [d] in Figure 7.10) and adjusting for the number of parameters (see Peres-Neto *et al.* 2006), we find the following: (1) overall 75.17% of the variance is explained by the environmental data and location in *x*–*y* coordinates, components [a+b+c]; (2) 22.72% is explained solely by the environmental variables; (3) 55.09% is explained by the spatially structured environmental variables; (4) −2.19% due to the other space components [c]; and (5) 24.83% is not explained, the residual component [d]. The main reason for the negative contribution in the space

fraction [c] is that there is a high collinearity between the response and predictor variables, all showing spatially autocorrelated values based on Moran's I computed at the distance of 1.41 units: 'Spp1' = 0.187; 'Spp2' = 0.499; 'Spp3' = 0.428; 'soil' = 0.784; and 'moisture' = 0.631. In some circumstances, it is recommended to detrend the data and analyse the detrended data (Legendre & Legendre 2012). Unlike the Mantel test, which cannot distinguish the contributions of individual variables to the overall result, the RDA could determine that the relation between the species assemblage is mostly due to the 'Spp1' response to the values of 'moisture'.

We used RDA in this example, because the species response to the environmental data is approximately linear due to the relatively narrow range of the environmental gradient. Where the environmental gradient is wide and the species response is unimodal, CCA should be used instead (ter Braak 1986). As a technical detail, CCA, unlike the RDA, can preserve the χ^2 distance among the locations in the ordination plot. Palmer (1993) provided a useful explanation of the characteristics of CCA and when it should best be used.

7.4 Multiscale analysis

The main aim of ecological studies is to understand which processes and environmental factors act on ecological response variables to generate patterns whether the patterns are spatial, temporal or phylogenetic. In Chapters 1 and 2 of this book we have discussed at length the fact that several factors and processes can produce more than one spatial pattern at more than one scale (e.g. large trends and local patchiness) embedded in data. At each spatial scale (local, regional and broad), there are different environmental factors and processes that are responsible for shaping the spatial patterns we observe (Table 7.2, Figure 7.10; Currie 2007; Fortin *et al.* 2012a). According to these various spatial scales, both the spatial resolution (data model) and the ecological data type vary, including vector (point, line, polygon) and raster (pixel, cell, quadrat) types. It is therefore imperative to use the appropriate analytical method to determine the key or most prominent spatial scales in ecological data according to both the data model and data type.

The effects of space on ecological data are not limited to the actual x–y coordinates of the samples with the resulting spatial autocorrelation and spatial dependency patterns, but also include a 'spatial legacy' (the current spatial pattern is partly due to the spatial pattern in previous times) and 'spatial contingency' (the current spatial pattern at location i is influenced by the spatial pattern at locations surrounding location i) (Chapter 1; Peres-Neto & Legendre 2010). All these various components of space are often used as surrogate predictor variables in multiple regressions and constrained ordinations in the absence of available factor variables. Through the years several attempts have been made to add space through predictor

Table 7.2 The relationships among spatial scales, the data type and key ecological processes, with examples

Local scale	Regional scale	Broad scale
Data model: vector	Data model: vector, raster	Data model: raster
Data type: qualitative and quantitative	Data type: qualitative and quantitative	Data type: presence : absence
Foraging	Resource use	Abiotic and climate factors
Biotic factors/Intraspecific	Biotic factors/Interspecific	Speciation/Extinction
Population dynamics	Metapopulation	Species range shift
Species interactions	Metacommunity	Species assemblage shift
Daily movement	Natal dispersal	Migration

variables by (1) simply using the *x*–*y* coordinates of the samples (e.g. in a multiple regression); (2) combining the *x*–*y* coordinates using a polynomial regression such as the trend surface analysis (Borcard *et al*. 1992; see Chapter 6) which models only the global larger spatial scales; (3) defining a neighbourhood matrix based on local adjacency or connectivity among the sampling locations (Pelletier *et al*. 1999); (4) defining a Euclidean distance matrix based on *x*–*y* coordinates of the samples, as it is done with the Mantel and partial Mantel tests already mentioned in this chapter; (5) defining a least-cost or quantitative resistance matrix based on landcover type values between the samples, as it is done with the Mantel and partial Mantel tests (Cushman *et al*. 2006; Spear *et al*. 2010); and (6) applying eigenfunction-based methods using the *x*–*y* coordinates of the sample locations (Griffith 1996; Diniz-Filho & Bini 2005; Dray *et al*. 2006; Dray *et al*. 2012). The advantage of the eigenfunction methods over the other approaches is their ability to provide predictors for more than one spatial scale, therefore reflecting a broader spectrum of potential factors that influence the spatial patterns observed in ecological data.

When the processes can be assumed to be stationary, a multiscale analysis can be performed using a multiscale ordination (MSO) based on multivariate variograms (Wagner 2003). A MSO analysis of multivariate species data produces variance–covariance matrices that match the various spatial scales as determined by an empirical variogram matrix (Chapter 6), allowing us to determine the spatial scale of individual species and of species composition. A generalized MSO method can also be used with canonical ordination axes (Wagner 2004; Couteron & Ollier 2005) and can be used to analyse beta-diversity.

In this section we present two types of decomposition analysis that can be used when the assumption of stationary is not met: (1) the generalized Moran's eigenvector maps (hereafter MEMs; Dray 2011; Dray *et al*. 2012) which provides spectral decomposition based on an eigenfunction-based method for sampled data (i.e. qualitative or quantitative vector data); and (2) multiresolution analysis (MRA) based on hierarchical wavelet transform for continuous population data (i.e. quantitative raster data).

7.4.1 Generalized Moran's eigenvector maps

As we mentioned in Chapter 3, there are several ways to capture the essence of the spatial structure of sampled data into a spatial weighted matrix, $\mathbf{D_W}$ (Figure 7.2): (1) a connectivity matrix (1 = connected; 0 = not connected), (2) a Euclidean distance matrix, or (3) a weighted matrix based on either cost or resistance values or on probabilities. With any such $\mathbf{D_W}$ matrix, we can use any of several spectral decomposition methods based on eigenfunctions to produce sets of spatial eigenvectors that are orthogonal and thus uncorrelated with each other.

In 2002, Borcard & Legendre proposed using a matrix of truncated Euclidean distances ($\mathbf{D}_{\text{trunc}}$) as the $\mathbf{W}$ matrix to perform a spectral decomposition based on a principal coordinate analysis (PCoA); they called this approach the 'Principal Coordinates of Neighbour Matrices' (PCNM). A truncated distance matrix, $\mathbf{D}_{\text{trunc}}$, allows the separation of the effects at small distances from those operating at large distances. Hence the determination of the truncation distance threshold (a Euclidean distance) is a crucial step in the production of these orthogonal spatial eigenvectors. In the absence of external knowledge, the threshold value can be based on a Minimum Spanning Tree (see Chapter 3) constructed on the *x*–*y* coordinates of the sampling locations by selecting the longest link of that tree as the threshold distance. For computational purposes, four times this threshold value is designated as the numerical equivalent of 'far away' and replaces all the distances larger than this value; it is also inserted on the main diagonal of the $\mathbf{D}_{\text{trunc}}$ matrix, so that each sampling location is not considered to be connected to itself. The PCoA of the $\mathbf{D}_{\text{trunc}}$ matrix, which is the next step, produces n – 1 non-zero eigenvalues and n − 1 non-zero eigenvectors. (Perhaps surprisingly, some of the eigenvalues are negative, but that is due to effects of truncation, see Dray *et al*. 2006.) Each of the n – 1 spatial eigenvectors is related to a specific spatial scale. When the sampling locations are regularly spaced, the corresponding spatial eigenvectors give sine waves. The spatial eigenvectors are produced with decreasing periods so that both large and local spatial scales can

be determined from the set of Euclidean distances between the sampling locations. The orthogonal spatial eigenvectors thus produced can then be used as surrogates for the effects of 'space' forming a matrix of explanatory variables in a RDA, partial RDA, or multiple regression.

Since the Borcard & Legendre paper of 2002, there have been several clarifications of the PCNM approach leading to mathematical refinements of the spectral decomposition based on eigenfunctions (Griffith 1996; Dray *et al.* 2006; Griffith & Peres-Neto 2006; Peres-Neto & Legendre 2010; Borcard *et al.* 2011; Dray 2011; Dray *et al.* 2012; Legendre & Legendre 2012). It turns out that PCNM is, in fact, a specific case of a broader family of spectral decomposition methods based on eigenfunctions and the PCNM method is now more usually referred to as a distance-based MEM (dbMEM) of the generalized Moran's eigenvector maps (Dray 2011; Dray *et al.* 2012; Legendre & Legendre 2012). The generalized MEM differs from the dbMEM in its use of a $\mathbf{D_W}$ matrix that is the Hadamard product of two matrices, a connectivity matrix and a similarity matrix, the product of which is then centred by removing the column- and row-means. The eigenanalysis of this centred $\mathbf{D_W}$ symmetric matrix produces both eigenvalues and eigenvectors where the eigenvalues are Moran's I coefficient values. This relation between the eigenvalues and Moran's I can be better understood by rewriting the Moran's I using a matrix notation:

$$I(x) = \frac{n}{\mathbf{1}^T W \mathbf{1}} \frac{x^T \Omega x}{x^T \mathbf{H} x}, \tag{7.8}$$

where x is a vector of the quantitative data, n is the sample size. $\boldsymbol{W}$ is the spatial weighted matrix, (here $\mathbf{D_W}$), $\mathbf{H}$ is a centring operator ($\mathbf{I} - \mathbf{1}^T\mathbf{1}/n$; with $\mathbf{I}$ the identity matrix, and $\mathbf{1}$ a vector of ones), and $\Omega = \mathbf{H}W\mathbf{H} = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \mathbf{u}_i^T$. The $\mathbf{u}_i$ are normalized eigenvectors that maximize each successive value produced for Moran's I while maintaining orthogonality. Given Eq. (7.8), the $I_{\max}$ is obtained with the first eigenvalue (λ_1; i.e. the largest eigenvalue and largest value of Moran's I) and $I_{\min}$ with the last eigenvalue (λ_n; i.e. the smallest eigenvalue and smallest value of Moran's I). The sign of Moran's I coefficient of spatial autocorrelation depends on the sign of the eigenvector and its associated eigenvalue.

Given that the MEM is based on an undirected connectivity matrix it is assumed that the spatial patterns are isotropic. There could be circumstances however where directionality between the sampling locations affects the degree of similarity of the data. Such anisotropic cases can occur when the data are sampled along streams. The MEM can therefore be modified to account for directionality using a directional connectivity matrix (Blanchet *et al.* 2008a, 2011; Mahecha & Schmidtlein 2008). Such directional MEMs are called the asymmetric eigenvector maps (AEMs).

One drawback of these eigenfunction decomposition methods, however, is that they produce $n - 1$ orthogonal spatial eigenvectors, which may be too many! When the sample size is small this is not so much an issue but with a large sample size one needs to select a subset of these eigenvectors. Jombart *et al.* (2009) proposed a multiscale pattern analysis based on a PCA of a matrix of the coefficients of determination, the multiple regressions between the ecological data, and the spatial eigenvectors, while Blanchet *et al.* (2008b) proposed a forward selection procedure (Legendre & Legendre 2012).

To illustrate the MEM multiscale decomposition analysis, we simulated data on a lattice of 128 × 128 units, following the approach proposed by Legendre *et al.* (2002) (Figure 7.11). This provides data with several spatial scales of pattern: (1) a large north–south trend; (2) a large patchy pattern generated according to a spherical variogram with a spatial range of 100 units; (3) small-scale patchiness generated according to a spherical variogram with a spatial range of 20 units; and (4) random noise with a mean of 0 and variance of 1 (Figure 7.12). Four subsamples were taken from this dyadic grid (i.e. a power of 2; here 2^7) using random and systematic sampling designs with 100 and 400 samples. Based on these subsamples and their respective Minimum Spanning Trees (Figure 7.13), the threshold distance values to

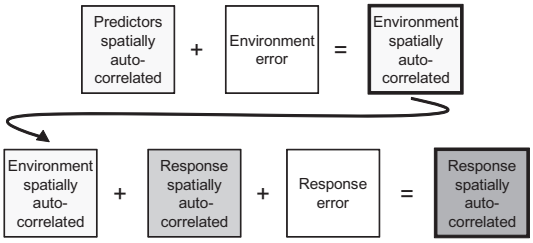


Figure 7.11 Flow chart on how predictor (e.g. environmental variables) and ecological response variables can be simulated using a conditional simulation method based on a variogram as implemented by Legendre *et al.* (2002) to include spatial autocorrelation structure in the data. Environmental error is added to spatially autocorrelated predictors to produce the spatially autocorrelated environmental data. The spatially autocorrelated predictors can be the sum of several predictors having different spatial scales. Then spatially autocorrelated environmental data are the starting point of the spatially autocorrelated ecological response data where a new spatially autocorrelated variable is simulated. The final spatially autocorrelated ecological response is the sum of the spatially autocorrelated environmental and response as well as the response error.

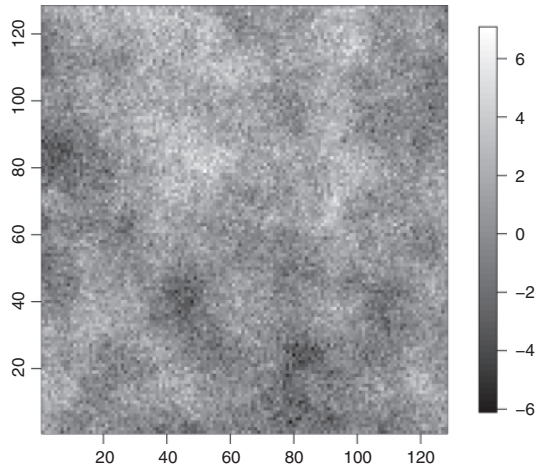


Figure 7.12 Simulated spatially autocorrelated response data (grid of 128 × 128 points) using the flowchart illustrated in Figure 7.11. The data are the sum of three spatially autocorrelated variables based on spherical variogram: (1) one gradient north–south; (2) spatial patchiness having a spatial range of 100 units; (3) spatial patchiness having a spatial range of 20 units; and (4) random noise.

truncate the Euclidean distance matrix into small versus 'far' distances are: 14.0 for the 100 random points; 10.0 for the 100 systematic points; 14.14 for the 400 random points; 6.0 for the 400 systematic points. Using these thresholds, the Moran's I eigenvectors were computed. For comparison, Figures 7.14 and 7.15 show the same eigenvector maps at the 1st, 10th, 25th and 30th rank for the 100 subsamples and 1st, 10th, 50th and 100th for the 400 subsamples. As mentioned above, the sine–cosine wave pattern is more obvious for the systematic subsamples than for the random ones. Yet when the 'random' sample size increases from 100 to 400, the shape pattern of the eigenvectors looks more like a sine–cosine wave, due to the higher density of points. When performing a forward selection procedure (Legendre & Legendre 2012) to determine the key spatial scales, the number of significant eigenvectors drops drastically (the number of spatial scales) for both sample sizes, where the largest spatial scales (i.e., the lowest ranked eigenvectors) are those retained to explain the pattern of the simulated data. The variance partitioning of the effects of the x–y coordinates (i.e. fraction [b]) and the significant eigenvectors (i.e. fraction [c]) show that the MEMs explain between 32% and 40% of the simulated spatial pattern while x–y coordinates explain only 7% to 15% (Table 7.3). Overall, the results are consistent, given the good spatial coverage of the samples over the entire simulated extent; but if the sample size is small relative to the extent of the study, the results can be quite different, resulting in less of the total variance being explained by the MEM, as Gilbert & Bennett (2010) have demonstrated.

These results illustrate the ability of MEMs to identify the key spatial scales and to help explain the variability of the simulated data. Although MEMs were used here as surrogate predictor variables, they can help to determine the spatial scales at which environmental or other processes may affect ecological variables.

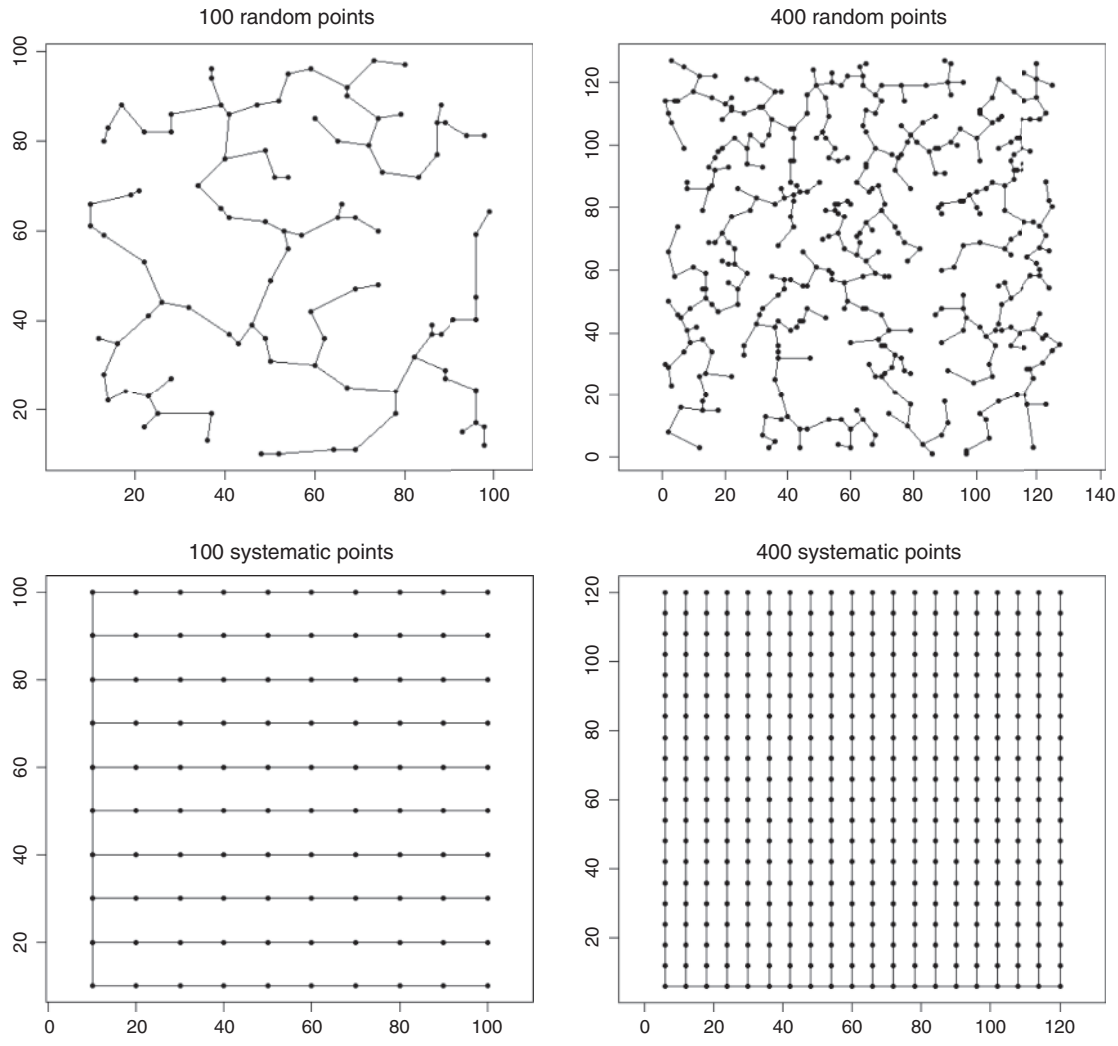


Figure 7.13 Four subsamples of the 128 × 128 points (Figure 7.12): random or systematic samples having 100 or 400 points. The lines linking the points are the Minimum Spanning Trees for each sample. The threshold distance values for the Moran's eigenvector maps are based on the Minimum Spanning Trees: 14 for the 100 random points; 10 for the 100 systematic points; 14.14 for the 400 random points; six for the 400 systematic points.

7.4.2 Multiresolution spectral decomposition analysis based on wavelets

Unlike the eigenfunction-based decomposition methods, multiresolution decomposition analysis is based on the spectral signal of quantitative data observed at a set of contiguous locations. Such contiguous data should be in the form of a dyadic grid (i.e. the dimension of the lattice being a power of 2), forming either a transect of one dimension or a two-dimensional lattice. A spectral decomposition of the signal of the observed data can be achieved using

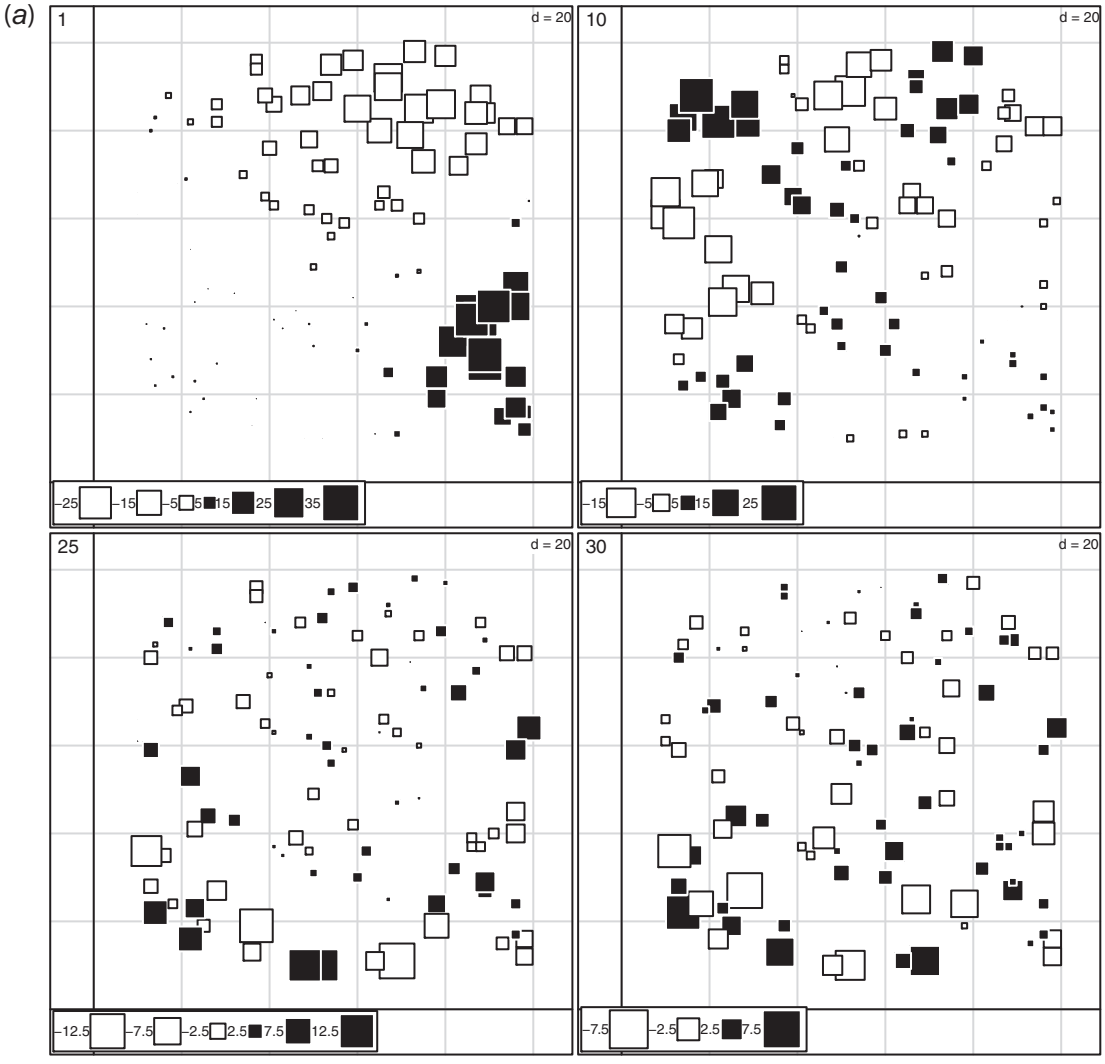


Figure 7.14 Moran's eigenvector maps of the 100 points: (*a*) random points and (*b*) systematic points. Here the 1st, 10th, 25th and 30th eigenvector maps are shown. The positive values are in black and the negative ones in white. The size of the square is proportional to their magnitude.

Fourier transform analysis that assumes stationarity, which requires the underlying process has the same characteristics throughout the full spatial extent covered by the data. If this is true, then the resulting decomposition's spatial scales will have a sine–cosine wave pattern. Yet, especially with large data sets, the likelihood of a process being stationary over an entire study area is low. In such circumstances, a wavelet transform analysis can be used instead because it does not require stationarity. Indeed,

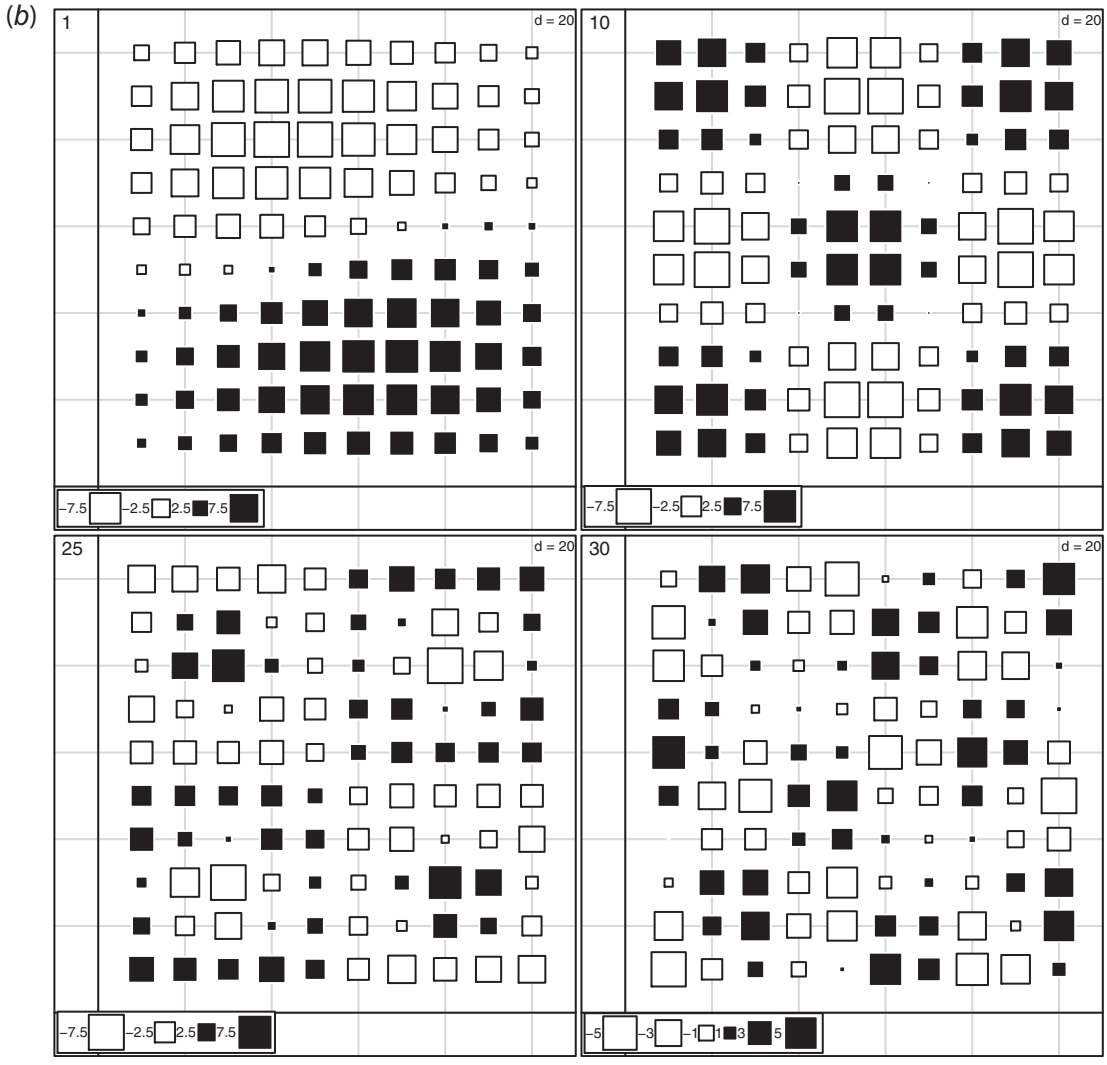


Figure 7.14 (*cont.*)

wavelet analysis can identify locally different types of spatial patterns, as we describe in Chapters 5 and 9. Here we stress the fact that, as for block variance methods (see Chapter 5), wavelet analysis can be performed using wavelet templates not only of various shapes but also of several spatial scales or resolutions (Daubechies 1992; Torrence & Compo 1998; Saunders *et al.* 2005). Hence analysis using wavelets of multiple resolutions also allows the identification of the key spatial scales of the data (Keitt & Urban 2005; Keitt & Fischer 2006; James & Fortin 2012).

There are different ways in which the multiple resolution analysis can be performed, providing spatial scales that are either orthogonal or not. To obtain

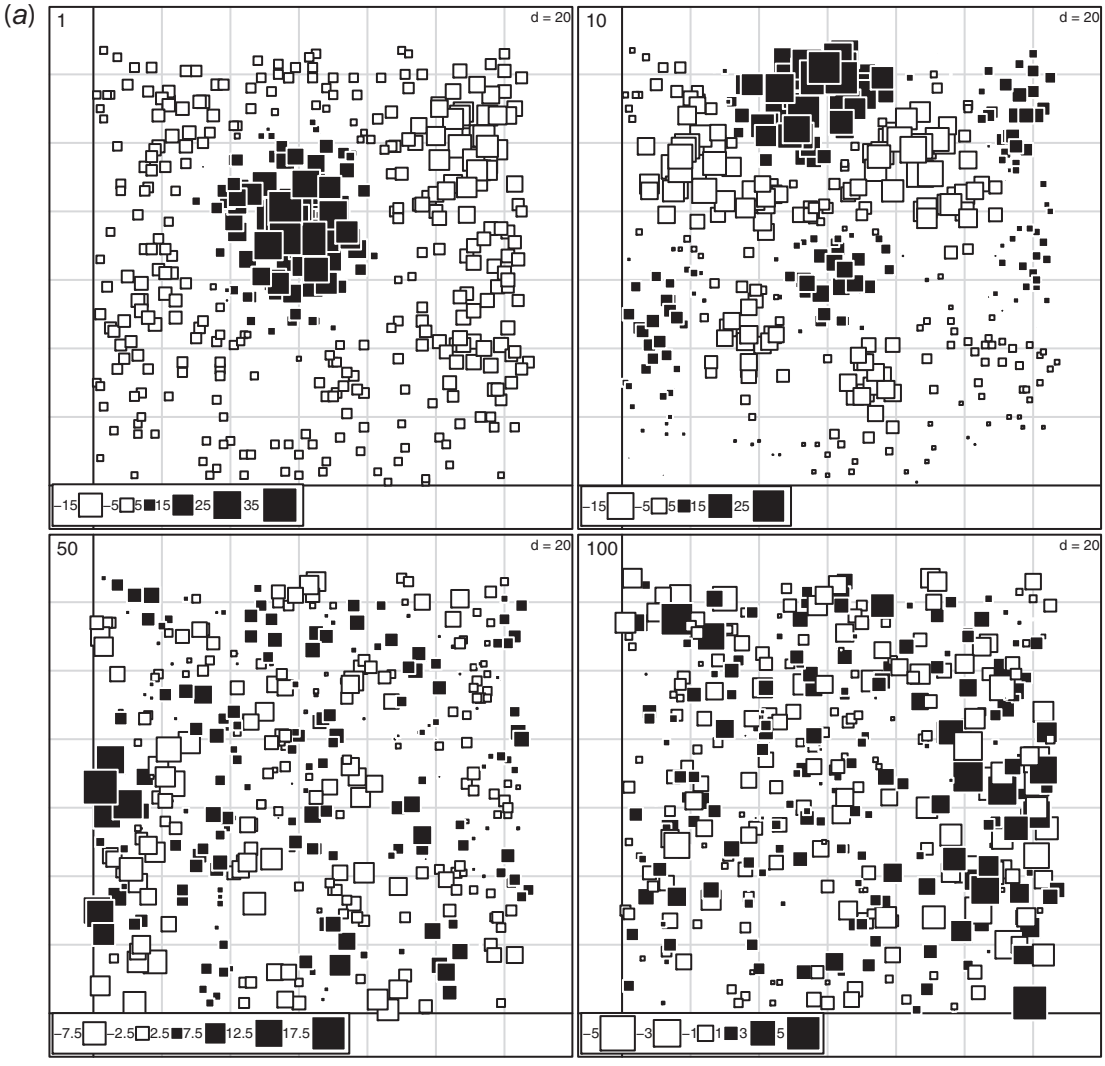


Figure 7.15 Moran's eigenvector maps of the 400 points: (*a*) random points and (*b*) systematic points. Here the 1st, 10th, 50th and 100th eigenvector maps are shown. The positive values are in black and the negative ones in white. The size of the square is proportional to their magnitude.

orthogonal spatial scales as powers of 2, a discrete decomposition wavelet transform (DWT) is applied (see Torrence & Compo 1998). The alternative, of computing the continuous decomposition wavelet transform (CWT), can produce nonorthogonal results that will include redundant spatial signals throughout the range of decomposed spatial scales. The maximum overlap discrete wavelet transform (MODWT) method is a compromise between the DWT and CWT and it will result in fewer spatial scales while keeping the original spatial resolution of the data (James & Fortin 2012). Here we will focus on the DWT method because

(*b*)

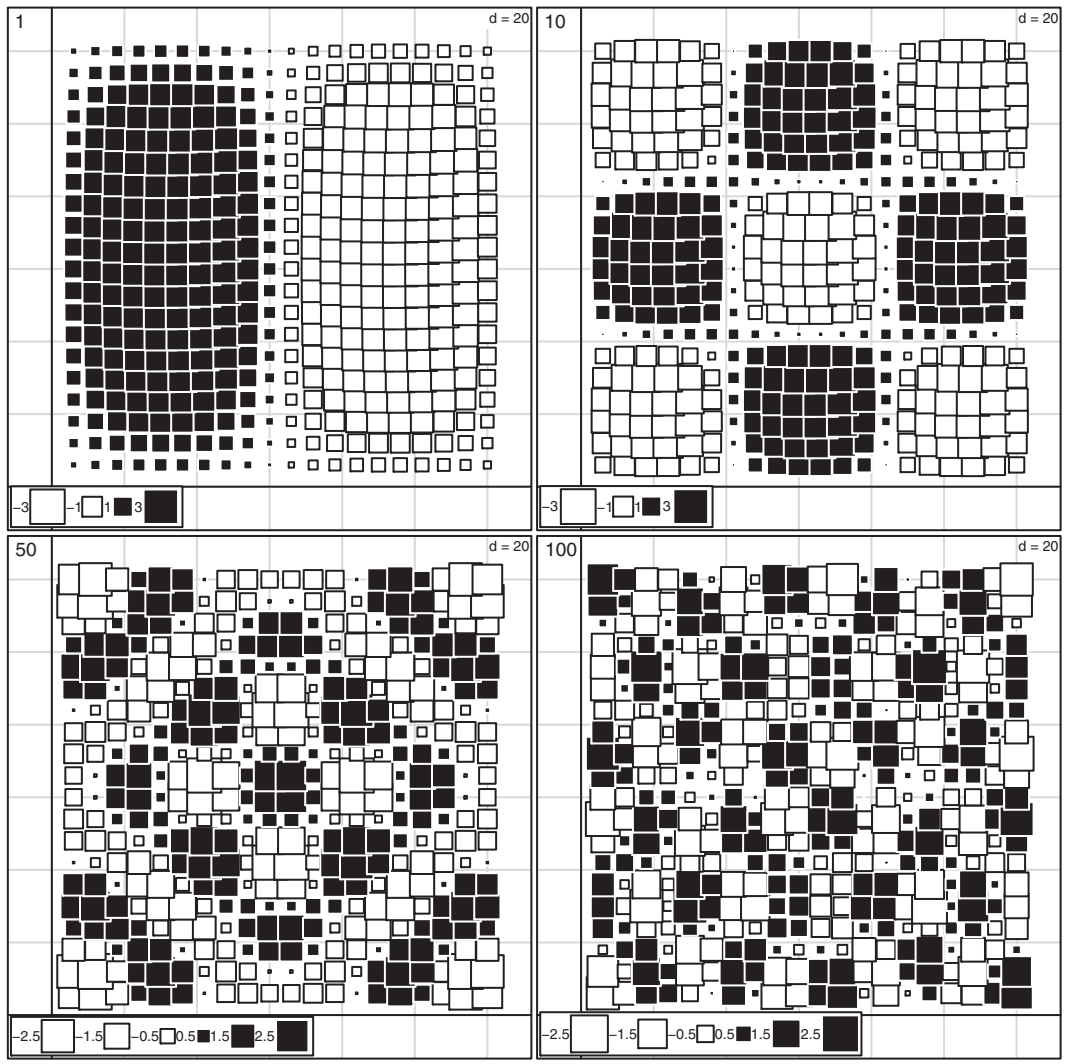


Figure 7.15 (*cont.*)

it is more directly comparable to the MEM, resulting in orthogonal scales.

The discrete wavelet transform, DWT, for the data series f, indexed by the location variable s, is based on a wavelet function, $\Psi(s)$. The transform is computed as follows:

$$\mathrm{DWT}_{\psi} f(m,n) \propto \sum_{s} f(s)\Psi^{*}_{m,n}(s), \tag{7.9}$$

where m is the dilation or scaling coefficient which determines the wavelet's width and thus the scale to which the wavelet is sensitive, and n is the translating or shifting coefficient which determines the wavelet's

Table 7.3 Summary of the results of the MEM and variance partitioning of the four subsamples

	100 random	100 systematic	400 random	400 systematic
Positive eigenvectors	33	44	120	189
Significant eigenvectors	5	6	25	40
Rank of the first four eigenvectors for comparison	1, 3, 4, 6	1, 3, 7, 12	1, 2, 4, 7	2, 4, 6, 8
Variance (%) explained by fraction [b]	13	7	7	15
Variance (%) explained by fraction [c]	34	33	32	40
Unexplained variance (%) explained (fraction [d])	53	60	61	45

position in the data grid. The wavelet function, now scaled and located, is $\Psi_{m,n}(s) = 2^{-m}{}_{\times}\Psi(2^m s - n)$ for our spatial analysis using the wavelet defined by the function $\Psi(s)$. (As a technical detail, the asterisk in Eq. (7.9) indicates that this is a complex conjugate, which is one of a pair of complex numbers with a real part and an imaginary part of opposite signs). On a dyadic grid, this DWT decomposes the signal into discrete orthogonal spatial scales by dilating the wavelet defined by $\Psi(s)$. In general, the results of wavelet analysis are presented by a two-dimensional plot of the wavelet transform scores, coded by colour, with the dimensions of the plot being scale (m) on one axis and location (n) on the other. Averaging the squares of the scores over all locations give the equivalent of a 'variance as a function of scale' plot, as with the quadrat variance methods described in Chapter 5. Such diagrams are often called 'scalograms' for obvious reasons.

The number of spatial scales that can be computed by this discrete wavelet scaling approach is given by the dyadic grid size of the observed data. In the case of the simulated data illustrated in Figure 7.12, the grid size is 2^7 and so the maximum number of scales we can examine is 7. Then the multiresolution decomposition analysis provides the proportion of the total wavelet variance of the data that corresponds with each decomposed spatial scale in the DWT approach. Figure 7.16 shows the spatial pattern analysis of the simulated data at each orthogonal spatial scale. Remember that the DWT is applied directly to the quantitative grid data and not to the Euclidean distance between the sampling locations as it was the case for the MEM. When a priori knowledge is available, we can combine the spatial pattern from two or more scales. In the absence of other knowledge about the key spatial scales, a wavelet variance scalogram, which plots the wavelet variance as a function of wavelet width (proportional to pattern scale) by averaging over all locations, can depict the relative contribution of the decomposed spatial scales. For the simulated data presented here as an illustration, the prominent spatial scales of pattern are 2^1, 2^4, and 2^6 (Figure 7.17), based on the positions of the peaks in the variance plot. Carl & Kühn (2010) and Vepakomma & Fortin (2010) used this multiresolution decomposition analysis method to remove the effects of some spatial scales that contain small-scale spatial autocorrelation on the data to examine only the larger spatial scales of the data in a subsequent regression analysis. This decomposition procedure can therefore be used to detrend data when the spatial patterns are not linear or not stationary.

7.5 Concluding remarks

More and more ecologists are designing their studies over larger areas with the aim of understanding which processes affect species' spatial arrangement. Given the availability of novel data types, novel spatial and statistical methods are needed to evaluate and understand them. In this chapter, we have shown how the

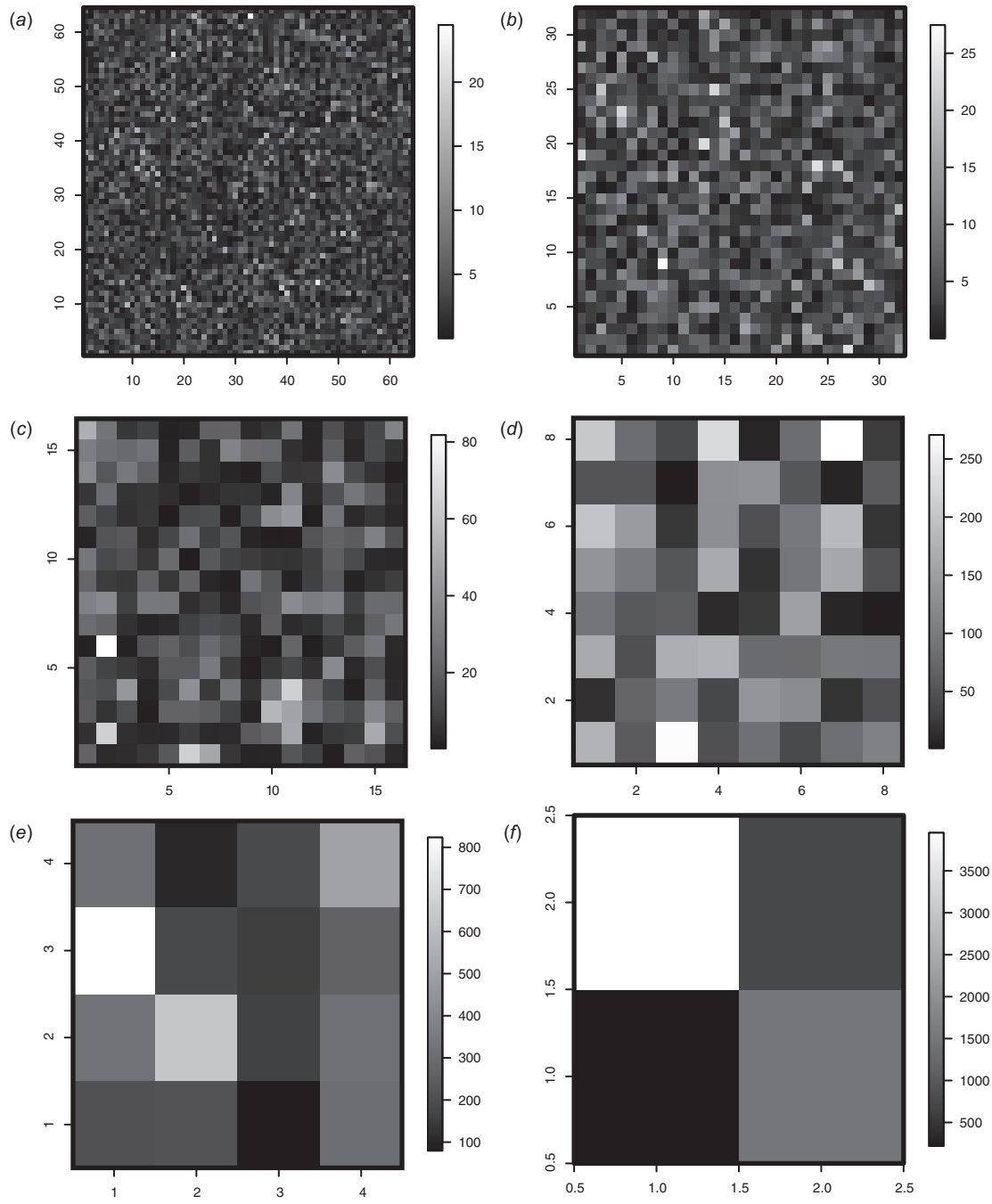


Figure 7.16 A multiresolution decomposition based on a discrete wavelet transform (DWT) of the dyadic simulated data of Figure 7.12 (128 × 128 units): (*a*) spatial scale 2^1; (*b*) spatial scale 2^2; (*c*) spatial scale 2^3; (*d*) spatial scale 2^4; (*e*) spatial scale 2^5; and (*f*) spatial scale 2^6.

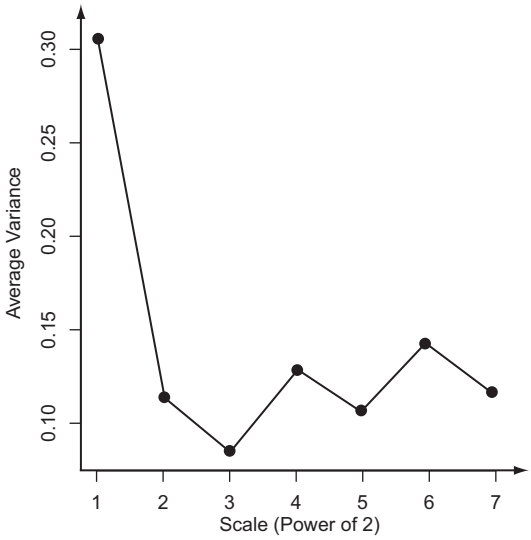


Figure 7.17 Wavelet scalogram of the simulated data illustrated in Figure 7.12. The key orthogonal spatial scales are 2^1, 2^4 and 2^6.

fields of spatial analysis and multivariate analysis have merged (Dray *et al.* 2012) to allow the multiscale analysis of ecological data. This new field of multivariate multiscale spatial analysis is very dynamic and advancing rapidly, and it is to be expected that more developments will occur in the years and decades to come. Yet it is important to remember that the use of spatial eigenvectors as spatial predictors in inferential statistics does not provide knowledge about any process-based predictors. Hence, when possible, environmental variables that are process-based should be used first, and the use of spatial eigenvector predictors should be the last resource, given that their explanatory power is limited to the study area extent and location, not allowing extrapolation to other areas. In the absence of alternatives, however, these approaches do provide ecologists with useful tools for the analysis of data in a spatial context and with a range of spatial scales.