Chapter 5
Spatial Dependence and Autocorrelation

5.1 Introduction

Spatial patterns are omnipresent in both environmental and ecological data (Wagner and Fortin 2005). In Chap. 4, we focused on point patterns to characterize the degree and spatial scale of aggregation or regular dispersion based on the *x*–*y* coordinates of point events (e.g., tree locations, nest locations). In Chap. 5, we are interested in interpreting spatial patterns of quantitative measures taken from sampling the environment.

The spatial analysis of such continuous variables falls under the realm of *spatial statistics*, traditionally aimed at quantifying spatial pattern and its statistical significance, and *geostatistics*, traditionally aimed at quantifying spatial variance and using this information to spatially interpolate data (Oliver and Webster 1991; Cressie 1993; Haining 2003; Dale and Fortin 2014). Spatial statistics can identify the spatial scale of patterns (i.e., the characteristic scale(s); Chap. 2). These approaches largely focus on quantifying and interpreting *spatial dependence*, or the similarity of a variable as a function of spatial location and/or geographic distance.

Tobler's first law of geography lies at the foundation of these issues, where "everything is related to everything else, but near things are more related than distant things" (Tobler 1970). To operationalize this fundamental idea, spatial statistics and geostatistics estimate spatial variance or covariance/correlation as a function of geographic distance. Tobler's law implies that at close distances, covariance or correlation of measurements should be high, yet as distances increase, the covariance or correlation should decline.

Spatial dependence can suggest key processes driving ecological patterns as a function of scale, such as spatial interactions among organisms as well as organism responses to environmental gradients that contain spatial dependence (Wagner and Fortin 2005) (Fig. 5.1). Yet spatial dependence can also be a nuisance for statistical inferences because data containing spatial dependence do not fulfill the common

R. Fletcher, M.-J. Fortin, *Spatial Ecology and Conservation Modeling*,
https://doi.org/10.1007/978-3-030-01989-1_5

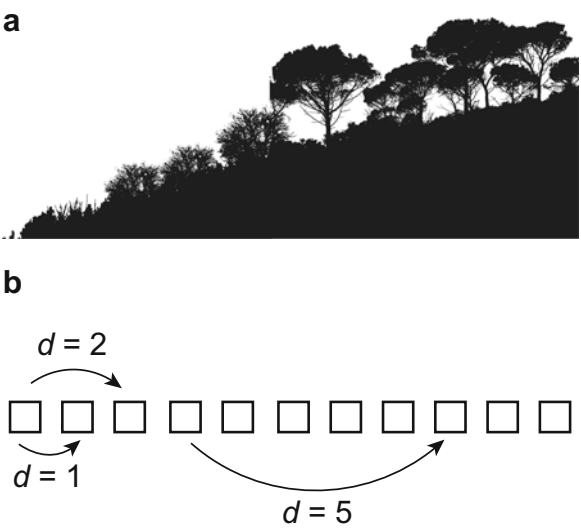

Fig. 5.1 The problem of spatial dependence. (**a**) When considering the environment across space, environmental measures such as elevation or canopy cover tend to be more similar at close locations and similarity declines with distance. (**b**) When sampling these gradients, we can interpret this spatial dependence based on lag distances between measurement locations. Shown are three lag distances ($d = 1, 3, 5$)

statistical assumption of independence assumed in many traditional statistics (Legendre 1993).

Here, our goals are to provide the key concepts needed to: (1) estimate the degree of spatial autocorrelation in data and potential spatial scale of the pattern; (2) understand how the estimated spatial variance can be used to interpolate and simulate spatial patterns using kriging and related approaches; and (3) identify the characteristic spatial scale(s) in the data using multiscale analysis (e.g., wavelet and spectral decomposition). We illustrate these concepts by extending our example from Chap. 4 on *Opuntia* cactus.

5.2 Key Concepts and Approaches

5.2.1 The Causes of Spatial Dependence

The terms *spatial dependence* and *spatial autocorrelation* are often used interchangeably, yet each term has a subtle different meaning based on why similarity of measurements in space occur (Table 5.1). To understand these differences, it is useful to make the distinction of whether spatial pattern is driven by *endogeneous* or *exogeneous* mechanisms (Bolker 2003). Endogeneous mechanisms are those that directly occur from the organism or processes being considered, which result in spatial pattern. Some common examples include localized dispersal generating spatial clustering of organisms, or social or grouping behavior (e.g., schooling fish or herds of ungulates). Exogeneous mechanisms, in contrast, are those that occur outside of the organism or process being measured, such as spatial aggregation of

Table 5.1 Common terms for spatial dependence in ecology

Term	Description
Anisotropy	When data have properties that are direction-dependent. Contrast to isotropy.
Correlogram	A plot of autocorrelation as a function of lag distance.
Endogenous process	A process that directly arises from the organism or response variable being considered, which results in patterns of spatial dependence.
Exogeneous process	A process that arises from outside of the organism or response variable being measured, such as spatial aggregation of resources or environmental gradients used by the organism of interest. Sometimes referred to as induced spatial dependence.
Isotropy	When data are uniform in all directions, which is frequently assumed in the analysis of spatial dependence.
Kriging	A method of interpolation for which the interpolated values are modeled via spatial covariance functions derived from variograms.
Scalogram	A plot of the wavelet variance as a function of a scaling factor related to distance.
Stationarity	When spatial pattern does not change over space or time (i.e., there is no trend in spatial dependence), which is frequently assumed in the analysis of spatial dependence.
Spatial autocorrelation	In a narrow sense, spatial dependence that arises from endogeneous processes.
Spatial dependence	Similarity in a response variable as a function of spatial location/distance, which can be driven by endogeneous or exogeneous processes.
Variogram	A plot of the spatial covariance as a function of lag distance. Different quantities are sometimes plotted, with the most common being semivariance.

resources or environmental gradients used by the organism of interest, which is sometimes referred to as "indirect" mechanisms and *induced spatial dependence* (Peres-Neto and Legendre 2010). In this context, spatial dependence is often considered a broad term for statistical spatial covariance that can be driven by both exogenous and endogenous processes. Note that this has also been referred to as *spatial legacy* (Peres-Neto and Legendre 2010). In contrast, spatial autocorrelation is sometimes considered as a certain type of spatial dependence that is driven by endogenous processes alone (Dale and Fortin 2014).

5.2.2 Why Spatial Dependence Matters

Given that spatial dependence is widespread in nature, why might we care? There are several answers to this question. First, there are practical reasons: when spatial dependence occurs, sampling locations within the range of dependence are no longer independent from each other. This issue is particularly troublesome, given that many common statistical tests assume that samples are independent. For instance, in a linear regression model, we often write:

$$y_i = \alpha + \beta x_i + \varepsilon_i, \tag{5.1}$$

where y is our response variable at location i, α is the intercept, β is the deterministic slope of the relationship of x with y, and ε is the error term. Where is independence assumed? In the error term of the model, we assume that errors are normally distributed with a mean of zero and a variance, σ^2, which is assumed to be iid—independent and identically distributed. This assumption means that each residual i (i.e., the difference between the observed and predicted value for i) is not dependent on other residuals and each comes from the same underlying distribution (see Chap. 6 for more details). Consequently, the problem of spatial dependence arises in our assumptions of the error in the model.

What is the problem if we violate this assumption? When spatial dependence occurs in our data and we ignore it, it often leads to type I error, where we infer significant patterns in the data that may in fact not exist. This is in contrast to a type II error, where we fail to conclude a significant pattern occurs when in fact exists. The issue of type I error arises because we are implicitly assuming that we have a larger sample size (and thus greater degrees of freedom) than we actually do, sometimes referred to a pseudo-replication (Hurlbert 1984). This assumption leads to artificially small estimates of uncertainty (or artificially high precision), such as standard errors (SEs) or confidence intervals (CIs) for parameter estimates, such as the SE of β in Eq. (5.1). Spatial dependence is thus thought to primarily bias our interpretation of the precision, not point estimates (e.g., we might adequately estimate β but not the SE or CI of β). Consequently, accounting for spatial dependence in statistical models, such as linear regression, may be necessary in some cases (see Chap. 6 for examples on how to do so). Alternatively, by identifying the scale(s) at which spatial dependence occurs, we may better design investigations to minimize problems of spatial dependence (Oliver and Webster 1991), such as spacing sampling locations at distances (i.e., lag distances) greater than the expected range of spatial dependence in the data.

The second reason why we might care about spatial dependence is that describing spatial dependence in our data may provide insights toward understanding key biological processes that generate the spatial patterns we are observing. For instance, when spatial dependence arises, is this pattern revealing the scale of social behavior, environmental variation in key resources, or dispersal (Brown et al. 1995; Koenig 1998; Fletcher and Sieving 2010; Cohen et al. 2016)? While quantifying spatial dependence alone may not provide rigorous answers to such questions, it may generate hypotheses or further predictions to help isolate the causes of spatial dependence.

Finally, spatial dependence can alter conclusions regarding conservation threats for many species and conservation strategies (Carroll and Pearson 2000; Landeiro and Magnusson 2011; Yoo and Ready 2016). For example, Koenig and Liebhold (2016) illustrated that there has been increasing spatial synchrony (one form of spatial dependence; see Chap. 10) in wintering birds across North America with warming temperatures over a 50 year time period. They emphasized that such

synchrony may have detrimental effects of population persistence through a reduction in demographic rescue (i.e., when dispersal reduces the probability of extinction of local populations).

5.2.3 *Quantifying Spatial Dependence*

There is a variety of ways to quantify spatial dependence. Here, we focus on the use of correlograms and semivariograms, which are complementary approaches frequently used in ecology and spatial statistics.

5.2.3.1 Correlograms

To understand how spatial statistics estimate spatial autocorrelation, it is useful to recall formulas for correlations, variances, and covariances. The spatial statistics we present emerge clearly from these classical statistics.

Recall the formula for a simple Pearson linear correlation, r, for two variables, z_1 and z_2:

$$r(z_1, z_2) = \frac{\sum_{i=1}^{n} (z_{1i} - \bar{z}_1)(z_{2i} - \bar{z}_2)}{\sqrt{\sum_{i=1}^{n} (z_{1i} - \bar{z}_1)^2 \sum_{i=1}^{n} (z_{2i} - \bar{z}_2)^2}} = \frac{\mathrm{Cov}(z_1, z_2)}{\sqrt{\mathrm{Var}(z_1)\mathrm{Var}(z_2)}}, \tag{5.2}$$

where $r(z_1, z_2)$ ranges from -1 to 1. The key is to extend this idea over space.

Moran's I test statistic extends the standard Pearson correlation over space (increasing lag distances) to estimate the degree of spatial autocorrelation for a quantitative variable, z, as:

$$I = \frac{n}{\mathbf{W}} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \left(z_i - \bar{z}\right)\left(z_j - \bar{z}\right)}{\sum_{i=1}^{n} \left(z_i - \bar{z}\right)^2}, \tag{5.3}$$

where $\mathbf{W}$ is a weight matrix that describes the dependency between locations i and j. Typically, this is a neighborhood indicator matrix, where $w_{ij} = 1$ if i and j are adjacent and 0 otherwise. Note this matrix is often row standardized, such that $\sum_j w_{ij} = 1$. This statistic can also be calculated for different distance categories, or bins, to interpret spatial dependence as a function of distance as:

$$I(d) = \frac{n}{\mathbf{W}(\mathbf{d})} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(d)\left(z_i - \bar{z}\right)\left(z_j - \bar{z}\right)}{\sum_{i=1}^{n} \left(z_i - \bar{z}\right)^2}. \tag{5.4}$$

Notice how similar Moran's I and Pearson correlation coefficients are: in essence Moran's $I(d)$ is a Pearson's coefficient computed for one variable against itself according to increasing distances among sampling locations (d) (Fig. 5.1). The plot of $I(d)$ as a function of distance class is called a spatial *correlogram*: its shape helps to interpret how the spatial pattern varies with distance and to estimate the spatial scale of the pattern. When z is normally distributed and that there are enough pairs of sampling locations per distance class (usually more than 20 pairs), the $I(d)$ will vary between +1 (where positive values indicate positive spatial autocorrelation) and -1 (where negative values indicate negative spatial autocorrelation), while values close to 0 indicates the absence of spatial pattern. Thus, Moran's I behaves as a Pearson correlation coefficient, and it is frequently used by ecologists because of its intuitive interpretation. Yet, when $I(d)$ is computed with less than 20 pairs, its value can be greater than 1 or smaller than -1. To avoid this known "boundary" or "edge effect" (Chap. 4), correlograms are often computed only up ½ or ⅔ of the maximum distance between the sampling locations to ensure adequate sample size for each distance bin (Dale and Fortin 2014). Note that Eq. (5.3) provides a common, global test for spatial dependence, while Eq. (5.4) is typically only used for the generation of correlograms. We focus on the use of correlograms because they provide much richer and intuitive information regarding spatial dependence.

Moran's I is an isotropic (i.e., pooled in all directions) averaged value of spatial autocorrelation per distance class for the entire extent of a study area. To detect the potential for *anisotropy* (i.e., spatial autocorrelation that varies in different directions) in the spatial pattern, the estimation of spatial autocorrelation can be computed using both distance and angle classes (i.e., different directions).

As Moran's I is a dimensionless number, it can be compared across different variables. One limitation of Moran's I is that it is sensitive to outliers (e.g., one or a few points can generate significant, erroneous autocorrelation). This is why some researchers transform the data (e.g., log-transformation of the response variable) to reduce the impacts of outliers. Because of this sensitivity, a similar statistic, Geary's c, has been developed. Geary's c values range from 0 (positive spatial autocorrelation) to 2 (negative spatial autocorrelation) and 1 indicates the absence of spatial autocorrelation. Yet Geary's c is also somewhat sensitive to outliers. As Geary's c is in essence the standardized equivalent of the semivariance presented below, we will not focus on Geary's here (but see Dale and Fortin 2014 for details).

Significance for each Moran's I coefficient can be based on Monte Carlo randomizations or through normal approximations. If significance is assessed using normal approximations then the assumption of *stationarity* needs to be valid. Stationarity is a term that describes a situation where the process that generated the spatial pattern does not vary in across a study area (e.g., mean and variance are similar throughout the region of interest) (Haining 2003). As the same data are used to compute $I(d)$ at increasing distances, the $I(d)$ values are not independent. This is

the same statistical issue encountered previously regarding spatial point pattern analysis (Chap. 4), which requires the use of multiple comparison corrections. Therefore, a Bonferroni adjustment correction (or something similar) should be applied that accounts for the number of distance classes computed, k, to adjust the significance level (Brunsdon and Comber 2015). For instance, using a Bonferroni correction, for a $I(d)$ to be statistically significant, its probability needs to be smaller or equal to $0.05/k$ (e.g., for $k = 15$, the adjusted probability to be significant based on a Bonferroni correction is $0.05/25 = 0.003$).

5.2.3.2 Variograms

Geostatistics comes at the same goal of estimating spatial dependency through a slightly different means (Cressie 1993). Instead of starting with a correlation coefficient (i.e., standardized covariance) such as the Moran's I, geostatistics stem from the sample variance and covariance instead:

$$\mathrm{Var}(z) = \frac{1}{n-1}\sum_{i=1}^{n}\left(z_i - \bar{z}\right)^2, \tag{5.5}$$

$$\mathrm{Cov}(z_1, z_2) = \frac{1}{n-1}\sum_{i=1}^{n}\left(z_{1,i} - \bar{z}_1\right)\left(z_{2,i} - \bar{z}_2\right). \tag{5.6}$$

The semivariance, γ, is calculated as:

$$\gamma(d) = \frac{1}{2n(d)}\sum_{i}^{n(d)}\left[\left(z(x_i) - z(x_i + d)\right)\right]^2, \tag{5.7}$$

where z is the value of the variable at location x_i, and $n(d)$ is the number of pairs of sampling locations at distance class d. Note the similarities with the variance equation.

The term "semi" comes from the fact that we divide by 2 (it helps to stabilize the statistical properties of the metric). Again, plotting γ as a function of d produces a semivariogram, often simply referred to as a *variogram*. Note that semivariance is on the same units as the data (e.g., km). Unlike Moran's I, but like variance, $\gamma(d) \geq 0$ and there is no upper bound. For interpreting the shape of the semivariogram, small values (closest to 0) indicate strong spatial covariance (i.e., strong spatial pattern), whereas larger values indicate less spatial covariance (i.e., weak or no spatial pattern). It is a rule of thumb to only interpret $^2/_3$ of the total distance (extent) considered, similar to that for Moran's I (Cressie 1993; Dale and Fortin 2014); for larger distances, the $n(d)$ is typically too small for reliable inference.

Semivariance computed from observed data are called "empirical," "experimental," or "observed" variograms. Empirical variograms simply plot the semivariances

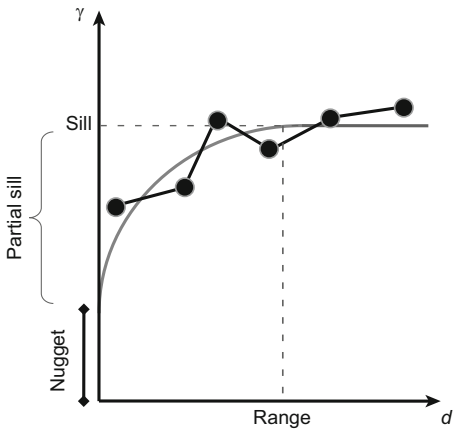

Fig. 5.2 Empirical and theoretical variograms, including the parameters of the variogram model. The empirical variogram shown as black points/line, while the theoretical variogram shown as a curved, asymptotic gray line

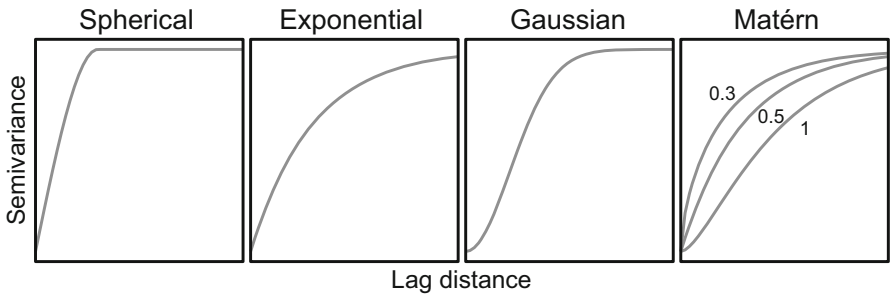

Fig. 5.3 Some common theoretical variogram models. For the Matérn variogram, different levels of kappa, a parameter in this model, are shown (note that when kappa = 0.5, the Matérn variogram becomes equivalent to the exponential variogram)

as a function of distance class, *d* (the spatial lag distance). Theoretical (or model-based) variograms can be fitted to an empirical variogram to spatially interpolate data at unsampled locations and to formally estimate the spatial scale of the pattern. In the presence of spatial pattern, three parameters relevant to interpreting the semivariance can be estimated from the fitted theoretical variogram: the range, sill, and nugget (Fig. 5.2). The *nugget* is the *y*-intercept at the origin that is greater than zero. It explains the variability in the data that occurs at very short distances, which could occur from measurement errors, sampling bias, or other random factors. The *range* indicates the distance up which the spatial dependence occurs, such that beyond the range, the data are no longer spatially autocorrelated. The *sill* is the value of semivariance beyond the estimated range, i.e., the variability that cannot be attributed to spatial autocorrelation. Note that some theoretical models assume that there is no sill (e.g., an exponential model; Fig. 5.3) while others assume that there is no nugget (i.e., the intercept = 0) (see Dale and Fortin 2014). If our interest is in spatial interpolation (which historically was the goal of variogram analysis and

geostatistics), we need to estimate parameters of model-based semivariograms and determine their relative fit to the data, using, for example, model-selection approaches (Burnham and Anderson 1998).

5.2.3.3 Kriging

To spatially interpolate across a region (e.g., make a predictive map of the response variable) based on the degree of spatial dependence, kriging is often used. *Kriging* is essentially a weighted moving average technique that uses estimates from a semivariogram (range, nugget, and sill) to perform spatial interpolation. More specifically, it is a set of linear regressions that determine the best combination of weights to interpolate across a region of interest by minimizing the variance from the spatial covariance in the data, where weights are derived from the estimates from the variogram (Dale and Fortin 2014; Oliver and Webster 2014).

The general form of the kriging model can be described as (Brunsdon and Comber 2015):

$$z = f(x_i) + v(x_i) + \varepsilon_i, \tag{5.8}$$

where $f(x_i)$ is a deterministic trend function (e.g., the response is non-stationary and may change with latitude or longitude), $v(x_i)$ describes the spatial dependence based on variogram parameters, and ε_i is the error. When there is no deterministic trend, *ordinary kriging* is used to interpolate based solely on the variogram parameters. In contrast, *universal kriging* assumes a large-scale, deterministic trend in the data, $f(x_i)$ (non-stationarity). This component is sometimes referred to as trend-surface analysis, which will be discussed in Chap. 6. Mathematical details about the different types of kriging algorithms can be found in Cressie (1993) and Haining (2003). Oliver and Webster (2014) provided a useful, practical tutorial on kriging.

In general, kriging is preferred for spatial interpolation in contrast to other simpler approaches. For example, a common, intuitive approach is inverse distance weighting (IDW) interpolation. IDW interpolates based on estimates that provide greater weight from nearby locations rather than distant ones. However, unlike kriging, this approach does not provide an objective means to determine the magnitude of distance-based weighting or the extent (maximum distance/limiting radius) for weighting. IDW also cannot provide SEs or other measures of uncertainty for predictions. Kriging, in contrast, has been shown to provide the best linear unbiased prediction for unsampled locations and can provide SEs for predictions. Reliable use of kriging requires proper estimation of the variogram model (Oliver and Webster 2014).

5.2.3.4 Some Extensions

For binary data, semivariance can be calculated with indicator functions by replacing $z(x_i)$ in Eq. (5.7) with an indicator function (Rossi et al. 1992). Monte Carlo randomizations are typically used to infer significance in such situations.

Both Moran's $I(d)$ and the semivariance γ function can be extended to address spatial correlations between two variables, termed "cross-correlograms" and "cross-semivariograms," respectively (Goovaerts 1994; Wackernagel 2003). For instance, a cross-variogram between variables u and v can be defined as:

$$\gamma_{uv}(d) = \frac{1}{2n(d)} \sum_{i}^{n(d)} \left[\left(z_u(x_i) - z_u(x_i + d) \right) \right] \left[\left(z_v(x_i) - z_v(x_i + d) \right) \right], \tag{5.9}$$

Semivariance, like Moran's I, is a "global" statistic. These models have also been extended to estimate variation in local intensity of spatial dependencies, referred to as LISA, or Local Indicators of Spatial Association (Anselin 1995; Boots 2002). These local measures are sometimes used to identify hotspots of intensity across landscapes (Nelson and Boots 2008).

5.2.3.5 Statistical Nuisance

Finally, spatial dependence can often be more of a statistical nuisance issue for ecological and conservation problems. In these cases, we might want to adjust conventional analyses (e.g., linear regression) to deal with dependence. This can be done in a variety of ways (Keitt et al. 2002; Beale et al. 2010). For example, conventional generalized linear models (GLMs) have been extended to adjust for spatial dependence by directly modeling the covariance of the residuals (generalized least square regression, GLS). Another common practice is to assess whether or not spatial dependence remains after a conventional analysis. This is frequently done by calculating Moran's I on the residuals from a model (Dormann et al. 2007). If there is evidence for autocorrelation of the residuals, then the conventional analysis should be replaced with one that formally adjusts for spatial dependence. We will consider these approaches in detail in Chap. 6.

5.3 Examples in R

5.3.1 Packages in R

There are a few libraries to choose from for quantifying spatial dependence. We will focus on using `geoR` (Ribeiro and Diggle 2016), `spdep` (Bivand and Piras 2015), `gstat` (Pebesma 2004), `pgirmess` (Giraudoux 2018), and `ncf` (Bjørnstad and Falck 2001). `Spatial` (Venables and Ripley 2002) allows for limited geostatistical

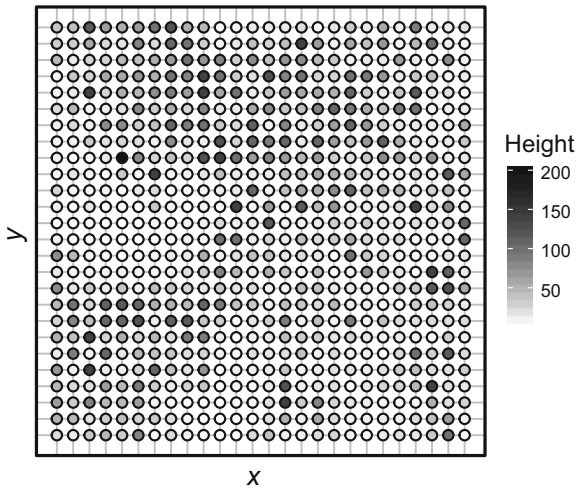

Fig. 5.4 A map of vegetation height measurements (cm) taken every 2 m on the plot

analysis (empirical correlograms and variograms), which comes with the VR Bundle when installing R. `spdep` has more options for correlograms and other spatial features (Bivand 2006). `geoR` provides a model-based variogram analysis based on maximum likelihood, while `gstat` has several geostatistics features, including the use of cross-variograms. We will also use the `ncf` package, which can fit spline (smoothed/non-parametric) correlograms and can provide a bootstrap approach for assessing statistical significance. We will implement kriging in with `geoR` and `gstat`.

5.3.2 The Data

As an example of interpreting spatial dependence, we return to the system considered in Chap. 4: old fields and prickly pear cactus (*Opuntia humifusa*) at the Ordway-Swisher Biological Station. In Chap. 4, we focused on data of *O. humifusa* locations in a 50 × 50 m plot, which were mapped using a high-resolution GPS (~30 cm error). Here, we focus on data from the surrounding matrix: samples of vegetation height taken systematically through the plot across a grid of sampling points spaced 2-m apart (Fig. 5.4) as part of a larger study on habitat loss and fragmentation (Fletcher et al. 2018). This information is relevant to movement of a pest insect considered in Chap. 4, *Chelinidea vittiger* (Schooley and Wiens 2004; Fletcher et al. 2014; Acevedo and Fletcher 2017) and we can use these measurements to first interpret spatial dependence of vegetation in the matrix and then create a map of vegetation height (via kriging) for understanding connectivity between *O. humifusa* patches (connectivity is covered in Chap. 9).

Our goals are to first interpret spatial dependence with the use of Moran's *I* and correlograms. We then use variograms to interpret the scale of spatial dependence and illustrate how model-based variograms can be used in kriging. Next, we

illustrate how kriging-related approaches can be used for generating spatial maps, similar in concept to what we illustrated in Chap. 3 with neutral landscape models. We end by introducing approaches to interpret multiscale spatial dependence.

5.3.3 Correlograms

We first import and visualize our data (`'cactus_matrix.csv'`). We will use a couple of different packages for calculating Moran's *I* and correlograms, contrasting what each can provide. We contrast these options because each uses different methods to infer statistical significance of potential spatial dependence, each varies in the complexity of coding required, and each can be helpful under different circumstances.

```
# load the matrix data into R:
> matrix <- read.csv(cactus_matrix.csv', header = T)
> head(matrix, 3)

##
 x y Height
1 0 0 35
2 0 2 65
3 0 4 75
```

With the data loaded, we can plot the data in several ways to interpret it. For example, we plot variation in vegetation height (`Height`) based on *x*–*y* coordinates, using a gray scale (with 12 breaks using the `cut` function) to fill points (using `pch=21`, which allows the fill of points to differ) to visualize variation in the matrix (Fig. 5.4).

```
> plot(matrix[,"y"] ~ matrix[,"x"],
  pch =21, bg = gray.colors(12)[cut(matrix[,3], breaks = 12)])
```

In correlogram (and variogram) analyses, we should truncate the range of lag distances at which we consider spatial dependence to approximately ½ to ⅔ the total distance observed. We can determine this distance by creating a pairwise distance matrix from the sampling locations. Because of the small spatial scale at which this plot occurs, we do not need to worry about projections for this calculation.

```
#calculate a distance matrix
> coords <- cbind(matrix$x, matrix$y)
> colnames(coords) <- c("x", "y")
> distmat <- as.matrix(dist(coords))

#maximum distance to consider in correlogram/variogram
> maxdist <- 2/3 * max(distmat)
```

To interpret spatial dependence with Moran's *I*, we start with the simplest approach and package and work up to less simple, but more flexible approaches. The first is the `pgirmess` package, which is a wrapper package for the `spdep` package. The `spdep` package has several useful spatial analysis functions, but it is less user-friendly than some other common spatial packages. The `pgirmess` package is more user-friendly in this way (but less flexible). We will use `pgirmess` and then contrast it to the `ncf` and `spdep` packages for generating different types of correlograms. In this package, we use the `correlog` function, specifying the coordinates for each sample and the measurement (i.e., height). We also specify that we want to use `method = "Moran"` (this package can also calculate Geary's *c*), the number of distance classes to consider, and we ask for the test to be two-sided (i.e., testing for both the potential of positive and negative spatial dependence).

```
> library(pgirmess)

#correlog from pgirmess
> correlog.pgirmess <- correlog(coords, matrix$Height, method =
  "Moran", nbclass = 14, alternative = "two.sided")

#summary
> head(round(correlog.pgirmess, 2))

##
 dist.class coef p.value n
[1,] 4.45 0.19 0.00 21692
[2,] 9.36 0.08 0.00 37708
[3,] 14.27 −0.01 0.22 51132
[4,] 19.18 −0.04 0.00 55500
[5,] 24.09 −0.02 0.00 61012
[6,] 28.99 −0.01 0.12 58540
```

In the above code, we find that the `correlog` function creates a matrix that contains each distance class considered (with `dist.class` reflecting the center of each bin), the Moran coefficient for that distance, the *p*-value, and the sample size (number of pairs of locations used) for that distance. This package uses normal approximations to test for the significance of spatial autocorrelation (i.e., it assumes the response variable is normally distributed and uses asymptotic theory to derive *p*-values). This approximation can be fast and relatively easy to implement, but it makes some key assumptions (e.g., normality in the residuals of the response data). We can then plot the correlogram

```
#correlogram plot
> plot(correlog.pgirmess)
> abline(h = 0)
```

The plot provides a visualization of Moran's *I* as a function of distance, with distances of significant spatial dependence shown in red (Fig. 5.5a is a generalized plot that contrasts this approach with those described below). This analysis suggests that positive spatial dependence is significant out to approximately 10 m, with some

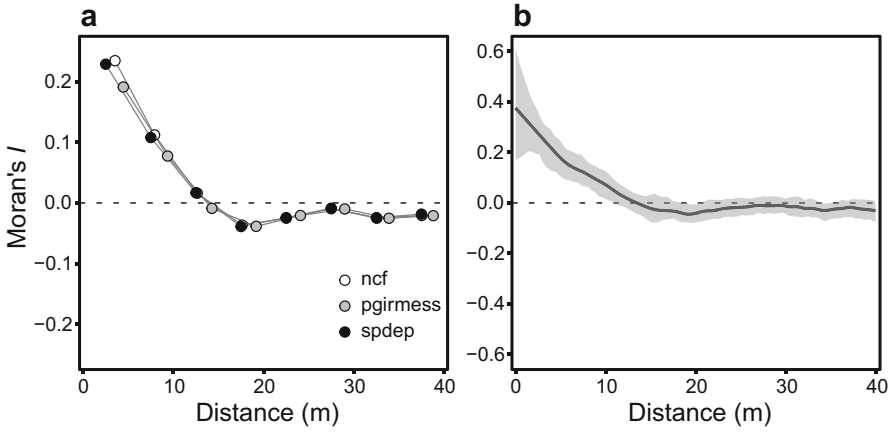

Fig. 5.5 (**a**) Contrasting distance-binned correlograms and (**b**) a spline correlogram. For (**a**), three approaches taken from different R packages are shown. For (**b**), the bootstrap confidence envelopes are shown. `Pgrimess` uses normal approximations to infer significance, `ncf` uses Monto Carlo permutations, while `spdep` is the most flexible package and can use either normal approximations or permutations to infer significance

evidence of negative spatial dependence at moderate distances. Note that this approach defaults to calculating Moran's *I* for distances up to the maximum distance in the observed data; however, we should ignore distances beyond ½ to ⅔ of the maximum distance.

An alternative approach is to use the `ncf` package. This package can provide non-parametric tests of significance for correlograms. It can also provide spline correlograms (Bjørnstad and Falck 2001). In spline correlograms, Moran's *I* is estimated with cubic splines that provide a smooth relationship across a variable of interest (more on this in Chap. 6) such that binning of distances is not required. This aspect is one benefit of using this package. This package has two approaches for interpreting potential significance of the correlogram. The first is a bootstrap approach (Efron 1979) to generate pointwise confidence intervals for the correlogram, such that evidence for spatial dependence is inferred when the confidence intervals do not overlap zero. Bootstrapping is a resampling technique used for inferring uncertainty in sample estimates and/or statistical significance in data. Bootstrapping involves resampling the data with replacement many times, where for each sample the variable of interest is calculated (in this case, Moran's *I*). The distribution of values of the estimate can then be used to approximate confidence intervals. A second approach `ncf` uses is the use of Monte Carlo permutations to generate a null envelope for spatial dependence, analogous to what we used in Chap. 4 for point patterns.

To use the `ncf` package, we need to either detach the `pgrimess` package or call the relevant function in `ncf` differently. This is because one of the functions we will use in the `ncf` package, `correlog`, has the same name as the one used above for `pgrimess`. If we do not want to detach `pgirmess` we can call the function from `ncf` as `ncf::correlog`.

```
> library(ncf)

#correlogram with Monte Carlo test
> correlog.ncf <- ncf::correlog(x = matrix$x, y = matrix$y, z =
  matrix$Height, increment = 5, resamp = 99)
> plot(correlog.ncf)
> abline(h = 0)
```

With this approach, we find similar evidence for spatial dependence based on the Monte Carlo permutations (Fig. 5.5a). However, in this case we may interpret that positive spatial dependence occurs at slightly greater distances than observed when using normal approximations with `pgrimess`. Note that in this function, the entire distance range is also considered although we should ignore distances beyond ½ to ⅔ of the maximum distance for inferences.

We can contrast these results with the use of spline correlograms with the `spline.correlog` function. We request a bootstrapping approach to infer significance in this situation.

```
#spline correlogram with 95% pointwise bootstrap CIs
> spline.corr <- spline.correlog(x = matrix$x, y = matrix$y, z =
  matrix$Height, xmax = maxdist, resamp = 100, type = "boot")

#plot with point-wise 95% CIs from bootstrap
> plot (spline.corr)
```

This correlogram suggests that the slight negative spatial dependence identified with the above approaches at moderate distances is too weak to infer a statistical pattern (i.e., the bootstrap confidence intervals overlap zero) (Fig. 5.5b).

Finally, we illustrate using the `spdep` package for correlograms. This package provides greater flexibility in the development of correlograms than other packages. For example, we can generate indicator correlograms for binary (0, 1) response data (see above). With this package, we will first calculate a general Moran's *I* that is sometimes used as an overall test of spatial dependence in data (Bivand et al. 2013). We then show how to create a customized correlogram using a similar approach.

To interpret spatial dependence with the `spdep` package, we must manually create the spatial weights matrix, **W**, shown in Eqs. (5.3) and (5.4). Note that `spdep` actually stores **W** in a list format rather than matrix format, because in many cases the former can be more compact and take up less storage computationally. Calculating **W** can be accomplished using the `knearneigh`, the `dnearneigh`, or the `cell2nb` functions. Here, we use the `dnearneigh` function, which creates a list, where each element is a vector for the neighbor IDs for each sample. Neighbors are identified based on distances specified with `dnearneigh`. The `knearneigh` finds the *k* closest neighbors, which could vary in distance with some sampling designs, while the `cell2nb` identifies data on regular grids, like that used here, but it is less generalizable so we do not focus on this function. Below we specify `d1 = 0` (minimum distance) and `d2 = 3` (maximum distance), which with these data results in an eight-neighbor function (i.e., queen's rule) (Fig. 5.6).

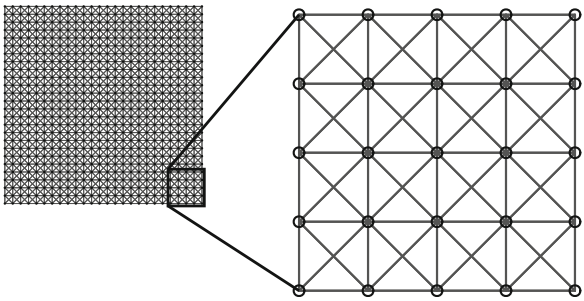

Fig. 5.6 Neighborhood matrix used for calculating Moran's *I* from $d = 0$–3. This matrix identifies neighbors based on an eight-neighbor rule (Queen's rule; see Chap. 3)

```
> library(spdep)

#make a neighborhood list:
> neigh <- dnearneigh(x = coords, d1 = 0, d2 = 3, longlat = F)

#plot the neighborhood
> plot(neigh, coordinates(coords))
```

To then calculate Moran's *I*, we convert the `neigh` object to a spatial weights list. As part of this procedure, we specify `style='W'`, which means that we will create a row-standardized **W**:

```
> wts <- nb2listw(neighbours = neigh, style = 'W', zero.policy = T)
```

With these spatial weights, we can now calculate Moran's *I*. `spdep` allows for inferring significance through normal approximations using `moran.test` (similar to `pgirmess`) or through Monte Carlo permutations using `moran.mc` (similar to `ncf`):

```
> mor.mc <- moran.mc(x = matrix$Height, listw = wts,
  nsim = 999, zero.policy = T)
> mor.norm <- moran.test(x = matrix$Height, listw = wts,
  randomisation = F, zero.policy = T)
> mor.mc

##
 Monte-Carlo simulation of Moran I

data: matrix$Height
weights: wts
number of simulations + 1: 1000

statistic = 0.27366, observed rank = 1000, p-value = 0.001
alternative hypothesis: greater

> mor.norm
```

```
##
 Moran I test under normality

data: matrix$Height
weights: wts

Moran I statistic standard deviate = 13.819, p-value < 2.2e-16
alternative hypothesis: greater
sample estimates:
Moran I statistic Expectation Variance
 0.2736595356 −0.0014814815 0.0003964261
```

In this case, both approaches yield identical estimates of Moran's *I* (0.274) and both provide a global test suggesting that spatial dependence is statistically significant.

Now we take the above approach calculate Moran's *I* for specific lag distance categories, generate a permutation value for each category, and then put together for a correlogram. We first create a data frame for storing the output and then provide a `for` loop that repeats the above process for each lag distance.

```
#first, create a df for storing data
> correlog.sp <- data.frame(dist = seq(5, 0.5 * max(distmat), by
  = 5), MoransI = NA, Null.LCL = NA, Null.UCL = NA, Pvalue = NA)

#Calculate Moran's I for lag distances
> for (i in 1:nrow(correlog.sp)){

 d.start <- correlog.sp[i, "dist"] − 5
 d.end <- correlog.sp[i, "dist"]
 neigh <- dnearneigh(x = coords, d1 = d.start, d2 = d.end,
 longlat = F)
 wts <- nb2listw(neighbours = neigh, style = 'W', zero.policy
 = T)
 mor.i <- moran.mc(x = matrix$Height, listw = wts, nsim = 99,
zero.policy = T)

 #summarize results from spdep
 correlog.sp[i, "dist"] <- (d.end + d.start)/2
 correlog.sp[i, "MoransI"] <- mor.i$statistic
 correlog.sp[i, "Null.LCL"] <- quantile(mor.i$res, p = 0.025)
 correlog.sp[i, "Null.UCL"] <- quantile(mor.i$res, p = 0.975)
 correlog.sp[i, "Pvalue"] <- mor.i$p.value
 }
> plot(y = correlog.sp$MoransI, x = correlog.sp$dist)
> abline(h = 0)
> lines(correlog.sp$dist, correlog.sp$Null.LCL, col = "red")
> lines(correlog.sp$dist, correlog.sp$Null.UCL, col = "red")
```

We have now seen several ways to calculate correlograms, each of which provides different benefits and limitations. In this case, the correlograms showed

generally similar patterns (Fig. 5.6). The use of normal approximations to interpret the significance of spatial autocorrelation (`pgirmess` and `spdep` packages) can be helpful with large data sets, where Monte Carlo tests can prove computationally expensive. However, Monte Carlo tests can be helpful when data are not normally distributed. The `ncf` package provides a means to not resort to binning of lag distances, which can be helpful and provides a straightforward bootstrapping procedure to infer significance. The `spdep` package allows great flexibility for calculating correlograms, but is less user-friendly.

5.3.4 Variograms

To illustrate empirical and model-based semivariograms, we use both the `geoR` and `gstat` packages. We primarily focus on the `geoR` package because it enables likelihood-based comparisons (e.g., AIC) between model-based variograms, which is useful for identifying the best variogram model for inferences and interpolation (Oliver and Webster 2014), and it provides an interesting Monte Carlo approach. The `gstat` package provides more options for different types of model-based variograms and can calculate cross-variograms so we briefly illustrate its use as well. We first create a `geoR` object that consists of the *x*–*y* coordinates and the value at each coordinate, which in this case is vegetation height. We refer to the measurements at sampling locations as *z*.

```
#load packages
> library(geoR)
> library(gstat)

#create a geoR object
> geo.veg <- as.geodata(matrix)
```

The `geoR` package provides a useful scheme for visualizing the raw data: `plot(geo.veg)` provides a four-panel plot. The first panel shows the sampling locations, where the measurements, *z* (vegetation height in this example), are shown as a color ramp, with low values being blue and high values red. The second and third panels show *z* values as a function of *x* and *y* coordinates. These panels can help for visually interpreting whether there is potential anisotropy in the data (directionality or trend in *z* as a function of *x*–*y* locations). The final panel provides a histogram (and density plot) of the *z* values.

We can calculate the empirical variogram for the data using the `variog` function in the `geoR` package. We will set the maximum distance considered based on our above code. Note that to calculate an empirical variogram, we bin lag distances. `geoR` will automatically do that for us, but we can also manually define the break points in the lag distance categories used for the semivariogram (Fig. 5.7):

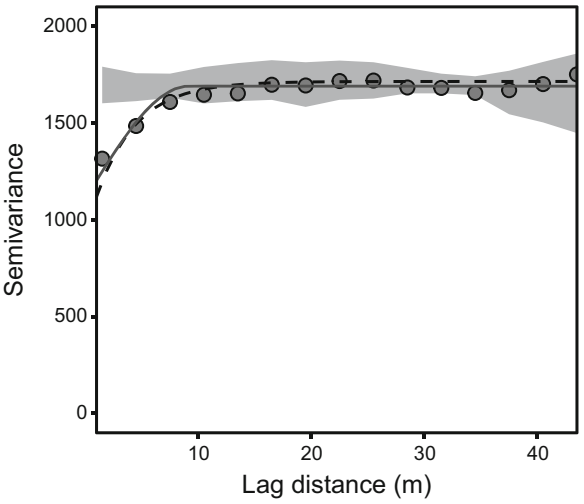

Fig. 5.7 Empirical and theoretical (exponential, black dashed, and spherical, gray solid, models) variograms for interpreting spatial dependence of vegetation height. Also shown are the 99% null pointwise envelopes (shaded region)

```
#Empirical semivariogram
> emp.geoR <- variog(geo.veg, max.dist = maxdist)
> plot(emp.geoR)

#standardize break points to a minimum 3-m lag distance
> emp.geoR <- variog(geo.veg, max.dist = maxdist,
  breaks = c(seq(0, maxdist, by = 3)))
> plot(emp.geoR)
```

In `gstat`, we can create empirical variograms by first creating an object that `gstat` can read (specifying the coordinates of the data) and then using the `variogram` function in `gstat`:

```
> gstat.veg <- matrix
> coordinates(gstat.veg) <- ~x + y
> emp.gstat <- variogram(Height ~ 1, cutoff = maxdist, width =
  3, gstat.veg)
> plot(emp.gstat)
```

Comparing the two packages illustrates that they provide essentially identical empirical variograms.

The above variograms assumed *isotropy*—no directionality in spatial dependence. We can subset our data based on direction to visually consider whether there might be evidence for anisotropy in spatial dependence using the `variog4` function in `geoR` or by adding the `alpha` argument to the `variogram` function in `gstat`. In both cases, data are subset such that four varigrams are calculated for the 0°, 45°, 90°, 135° directions (Fig. 5.8a), where 0° covers the range from −22.5° to 22.5°, 45° covers 22.5° to 67.5°, etc.:

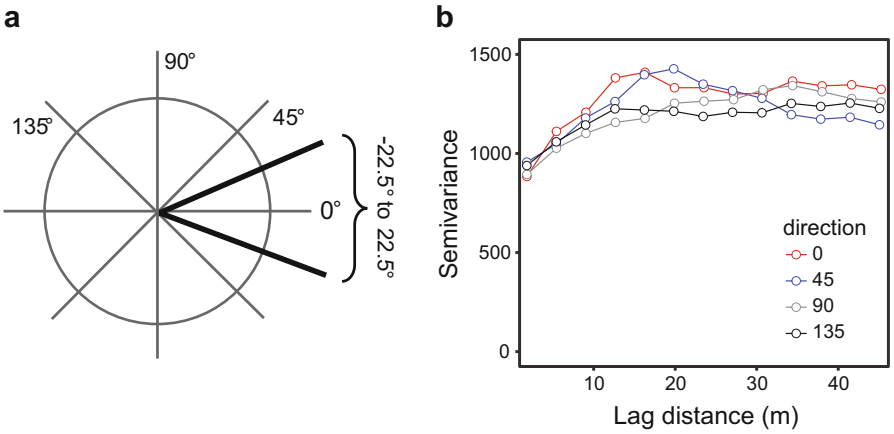

Fig. 5.8 Directional variography subsets the data to interpret variation in spatial dependence in different directions. (**a**) 0°, 45°, 90°, and 135° are typically considered (with windows ±22.5°). Larger values (between 180° and 360°) provide the same patterns because the semivariance formula is symmetric. (**b**) Directional variograms for vegetation height on the plot

```
#in geoR, variogram in each direction
> emp4.geoR <- variog4(geo.veg, max.dist = maxdist)
> plot(emp4.geoR)

#in gstat
> emp4.gstat <- variogram(Height ~ 1, cutoff = maxdist, alpha =
  c(0, 45, 90, 135), gstat.veg)
> plot(emp4.gstat)
```

In this plot (Fig. 5.8b), strong differences in the empirical variograms would suggest that anisotropy might be occurring in the data. Note, however, that to calculate these directional variograms, `geoR` and `gstat` are subsetting the data into four subsets, such that less data are used in each individual variogram. Consequently, the directional variograms may bounce around more than the variogram considered with all of the data. Why are only 0°, 45°, 90°, 135° considered? If directions between 180° and 360° were considered, for example, it would result in the same variogram patterns, because the calculation of the variogram is symmetric (the squared term in Eq. (5.7); $[z(x_i) - z(x_i + d)]^2 = [z(x_i + d) - z(x_i)]^2$).

We can fit theoretical variograms to the data using maximum likelihood techniques with the `likfit` function in `geoR` and contrast different variogram models using model selection criteria (e.g., Akaike's Information Criterion, AIC) (Oliver and Webster 2014). To do so, we must provide initial values for the partial sill (i.e., sill—nugget; Fig. 5.2) and the range, for which we can make an educated guess based on the empirical variogram. To fit exponential and spherical variogram models (Fig. 5.3):

```
#exponential variogram
> mlexp <- likfit(geo.veg, cov.model = "exp", ini = c(700, 10))

#spherical variogram
> mlsph <- likfit(geo.veg, cov.model = "sph", ini = c(700, 10))
> summary(mlexp)

##
Summary of the parameter estimation
-----------------------------------
Estimation method: maximum likelihood

Parameters of the mean component (trend):
 beta
43.0708

Parameters of the spatial component:
 correlation function: exponential
 (estimated) variance parameter sigmasq (partial sill) = 504.7
 (estimated) cor. fct. parameter phi (range parameter) = 5.884
 anisotropy parameters:
 (fixed) anisotropy angle = 0 ( 0 degrees )
 (fixed) anisotropy ratio = 1

Parameter of the error component:
 (estimated) nugget = 732

Transformation parameter:
 (fixed) Box-Cox parameter = 1 (no transformation)

Practical Range with cor=0.05 for asymptotic range: 17.62812

Maximised Likelihood:
 log.L n.params AIC BIC
 "-3298" "4" "6603" "6621"

non spatial model:
 log.L n.params AIC BIC
 "-3368" "2" "6739" "6748"

Call:
likfit(geodata = geoR.veg, ini.cov.pars = c(500, 15), cov.model = "exp")

> AIC(mlexp, mlsph)

##
 df AIC
mlexp 4 6603.375
mlsph 4 6603.830
```

The output from these models provides several key insights. For our purposes, we will focus on two types of important output. First, for each model, the log-likelihood, AIC and BIC (Bayesian Information Criterion) are provided for interpreting model fit and model selection. These values are provided for the spatial model considered, as well as a "non-spatial" model, which would assume a constant variance (i.e., variance does not change with lag distance). The output also provides estimates of the range, nugget, and partial sill for the model under consideration. For some theoretical variograms, the "practical range" is also provided. The practical range uses an approximation (which varies, depending on the theoretical variogram model) to determine the effective range distance when the variogram function shows a smoothed asymptotic relationship to the sill (e.g., exponential models; Fig. 5.3). For instance, in an exponential variogram it is typically defined as the distance where the variance reaches 95% of the estimated sill. In this example, the exponential variogram fits the data slightly better than the spherical variogram, based on AIC. Both of these models fit the data substantially better than a non-spatial model.

We can fit an exponential variogram in `gstat` as:

```
> exp.gstat <- fit.variogram(emp.gstat, vgm("Exp"))
```

Note that while `gstat` does not implement model selection based on likelihood techniques, it does provide a wider variety of model-based variograms that `geoR`. These alternatives can be perused with the `vgm()` and `show.vgm()` functions.

Finally, we can overlay plots of the theoretical variograms with the empirical variograms (Fig. 5.7):

```
> plot(emp.geoR)
> lines(mlexp, col = "blue")
> lines(mlsph, col = "red")
```

We can use model selection to contrast spatial and non-spatial models provided in the output of the `likfit` function. Another useful approach is to determine confidence envelopes of spatial randomness (analogous to envelopes calculated in Chap. 4). Null envelopes can then be overlaid with the empirical and theoretical variograms. In `geoR`, we can obtain null envelopes Monte Carlo permutations. The code below executes 99 permutations, where vegetation height is shuffled among x–y coordinates and plots the maximum and minimum values at each distance lag relative to the empirical variogram:

```
> emp.env <- variog.mc.env(geo.veg, obj.var = emp.geoR)
> plot(emp, envelope = emp.env)
> lines(mlexp, col = "blue")
```

These envelopes describe the variance as a function of lag distance under spatial randomness, given the underlying data. Consequently, when our observed variogram falls outside of this envelope, there is some signature of significant spatial

dependence in the data. In this case, we observe that the variogram only falls outside of the null envelope at distances <10 m, loosely similar to our conclusions on spatial dependence using correlograms (Fig. 5.5).

5.3.5 *Kriging*

With our theoretical variogram model, we can create an interpolated map using *kriging*. We could potentially make kriged predictions onto our observed sampling locations or onto a grid that covers the entire plot. We will illustrate the latter, where we use the `expand.grid` function to create a new set of locations. Note that distances between this expanded grid will provide the resolution of the map that we create. First, we krige with `geoR` (Fig. 5.9a).

```
#grid with 1-unit intervals (1-m)
> new.grid.1m <- expand.grid(0:max(matrix$x), 0:max(matrix$y))

#kriging: krige.control, cov.pars: partial sill, range
> krig.geoR.exp <- krige.conv(geoR.veg, locations = new.grid.1m,
  krige = krige.control(cov.pars = c(mlexp$cov.pars[1],
  mlexp$cov.pars[2]), nugget = mlexp$nugget,
  cov.model = "exp", type.krige = "OK"))

#get the prediction values for the kriged surface
> image(krig.geoR.exp, main = "kriged estimates")
```

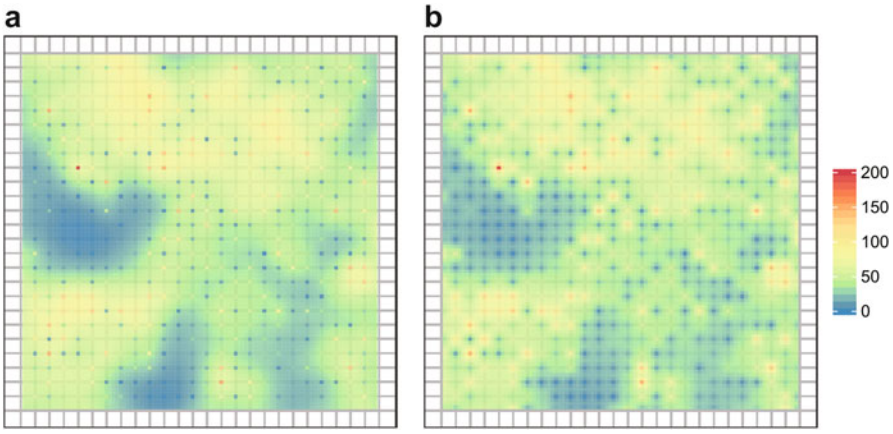

Fig. 5.9 Kriged map of vegetation height based on the exponential model (**a**). Also shown is the background sampling grid (2×2 m) for reference. (**b**) Inverse distance weighting interpolation (IDW)

In the above code, we take the estimates from the exponential variogram and use them in ordinary kriging (`type.krige = "OK"`) for spatial interpolation. The output includes predictions that we can use for mapping, as well as uncertainty in those predictions. In our case, the initial sampling grid was detailed at a high resolution, such that the uncertainty is very low. Note also that the kriged image uses our z values at the observed samples and only makes predictions for unsampled areas on our new grid. We can also plot the uncertainty in the predictions:

```
> image(krig.geoR.exp, val = sqrt(krig.geoR.exp $krige.var),
  main = "kriging SE")
```

Here, the model does not estimate variance for the sampled points; it fixes the variance to zero for those locations. Consequently, one could remove those sampled points from the mapping of the uncertainty in kriged predictions. This kriged map could then be used as a raster image for other purposes.

We can also implement kriging in `gstat` with the `krige` function.

```
> new.grid.1m <- expand.grid(x = 0:max(matrix$x), y =
  0:max(matrix$y))
> gridded(new.grid.1m) <- ~x + y
> krig.gstat <- krige(Height ~ 1, gstat.veg, new.grid.1m, model
  = exp.gstat)

#plot
> image(krig.gstat, main = "kriging-gstat")
```

In `gstat`, we need to have labels for the x–y coordinates in the new grid (unlike `geoR`). As an aside, inverse distance weighting interpolation is also straightforward in `gstat` with the `idw` function (Fig. 5.9b):

```
> idw.gstat <- idw(Height ~ 1, gstat.veg, new.grid.1m)
```

We can check the similarity in the kriged predictions from `geoR` and `gstat` and inverse distance weighting by calculating the correlation between predictions as:

```
> cor(cbind(geoR.exp = krig.geoR.exp$predict,
  gstat.exp = krig.gstat$var1.pred,
  gstat.idw = idw.gstat$var1.pred))

##
geoR.exp gstat.exp gstat.idw
geoR.exp 1.000 1.000 0.984
gstat.exp 1.000 1.000 0.984
gstat.idw 0.984 0.984 1.000
```

The two packages provide identical predictions based on kriging. In this case, the inverse distance weighting also provides nearly identical predictions to kriging. This

is not surprising, given the dense, regular sampling in the plot. With sparse and/or irregularly spaced sampling, we might expect these approaches to be less correlated.

5.3.6 Simulating Spatially Autocorrelated Data

Once the parameters of a theoretical variogram have been estimated, one can use these values to generate simulated spatially autocorrelated data having the same statistical properties of the observed spatial pattern using an annealing algorithm (Cressie 1993) or a Gaussian random fields algorithm, both of which are stochastic distribution functions (Lantuéjoul 2002). These procedures, and others, are often used to generate null reference distributions to test significance of observed spatial patterns in ecological data (e.g., Remmel and Fortin 2013). Note that when simulating Gaussian random fields, the simulations by default have a mean of zero.

We can use the `gstat` or `RandomFields` package (Schlather et al. 2015) to simulate spatial patterns based on the variogram parameters. Here, we show the use of `RandomFields`, which has more flexibility in this regard than `gstat`. We will also provide include the mean value of observed vegetation height (otherwise, the mean value of the simulated random field would be approximately zero).

```
#variogram models to simulate
> library(RandomFields)
> model.exp <- RMexp(var = mlexp$cov.pars[1], scale =
  mlexp$cov.pars[2]) + RMnugget(mlexp$nugget) + RMtrend(mean =
  mean(matrix$Height))

> dimx <- 1:50
> dimy <- 1:50

#simulate
> sim.exp <- RFsimulate(model = model.exp, x = dimx, y = dimy)
> data.sim <- as.matrix(sim.exp)

#plot with image
> image(dimx, dimy, data.sim,xlab = "x", ylab = "y")

#plot with raster package
> library(raster)
> RMexp.grid <- raster(data.sim)
> plot(RMexp.grid)
```

These simulated maps (Fig. 5.10) are called "unconditional Gaussian random fields." If we provide sample values for mapping (as in kriging), then the maps would be considered "conditional Gaussian random fields." Note that even when adjusting for the mean value of vegetation height with the `RMtrend` function, this

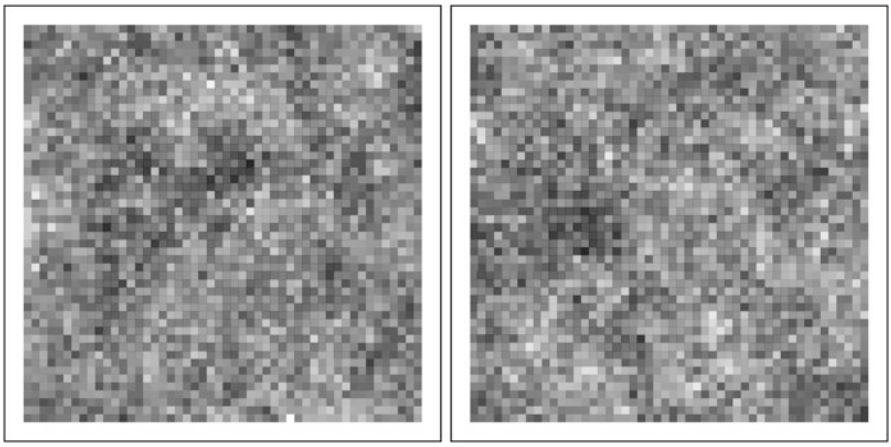

Fig. 5.10 Two realizations of simulating spatial dependence under the exponential variogram model fit to the vegetation data using unconditional Gaussian random fields

approach does end up simulating some values less than zero, which is not biologically plausible, because it is a Gaussian (normally distributed) random field. The `RandomFields` package has options for simulating fields that can circumvent this problem, but it is beyond the scope of this book.

The approach above can also be used to more generally create spatially autocorrelated maps of different degrees, similar in function to the neutral landscape approaches described in Chap. 3. For example, we illustrate how altering the partial sill and range parameters can generate different types of neutral landscape maps (Fig. 5.11) with the following alternative scenarios:

```
> model.exp.ps2r5 <- RMexp(var = 20, scale = 5) + RMnugget(var = 2)
> model.exp.ps8r5 <- RMexp(var = 80, scale = 5) + RMnugget(var = 2)
> model.exp.ps2r20 <- RMexp(var = 20, scale = 20) + RMnugget(var = 2)
```

The above scenarios take a base model, where the range = 5, the nugget = 2 and the partial sill = 20 and then increased the partial sill 4× and the range 4×. When plotting realizations of these models (similar to above), it is clear that changes in the partial sill increases the magnitude of variation and increases in the range parameter makes the map smoother. When truncating these maps similar to a neutral landscape scenario where we alter the proportion of habitat or land cover on the landscape (Chap. 3), however, changing the partial sill has negligible effects on the map while increasing the range leads to much greater aggregation of the land cover (Fig. 5.11).

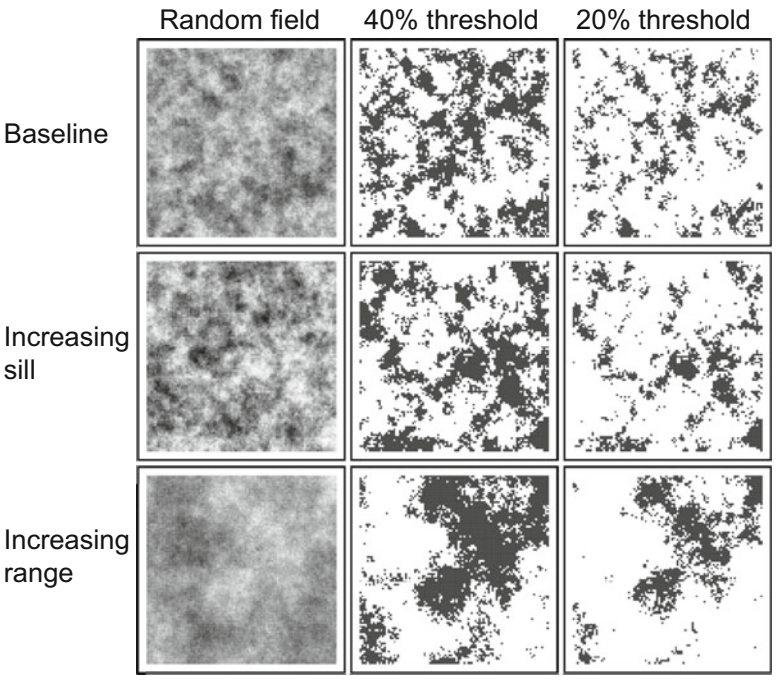

Fig. 5.11 Using unconditional Gaussian random fields to simulate neutral landscapes with varying spatial dependence based on an exponential variogram model. The first row is taken from a model with partial sill = 20, range = 5, and nugget = 2. The second row increases the partial sill 4×, while the third row increases the range 4×. Shown is the continuous random field and two thresholds of this map, similar to the use of neutral landscape models described in Chap. 3

5.3.7 Multiscale Analysis

With the availability of remotely sensed data and increasingly large databases that span broad extents, areas under investigation are usually large enough to include the effects of several process acting at various spatial scales that generate observed spatial patterns. With the potential of such multiscale effects, the first step consists therefore to identify the key spatial scales of the patterns. There are two multiscale analysis methods that can be used to decompose the key spatial scales from remotely sensed data or other data from a study area: Fourier spectral decomposition and hierarchical wavelet decomposition analysis (Keitt and Urban 2005).

5.3.7.1 Wavelets and Fourier Series

Fourier's technique and wavelets are related (Dale et al. 2002). Fourier's technique assumes the data have been generated by stationary processes that occur across the entire area of interest. With this approach, processes are envisioned as a series of sine and cosine waves operating at different scales that sum together to drive observed

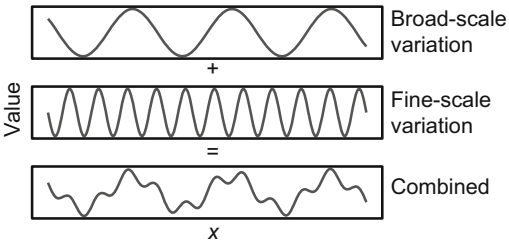

Fig. 5.12 How Fourier transforms work. Fourier transforms assume that observed variation comes from multiple processes operating at different scales. This approach uses sine/cosine waves of different amplitudes and periods to capture such variation. These broad and fine-scale waves are combined to interpret observed variation. This approach assumes stationarity, where waves operate across the entire extent of interest, unlike wavelet transforms

variation (Fig. 5.12). This formulation can be helpful in many situations when broad extents are considered; indeed, linear trends across broad extents are unlikely in many situations (Austin 2002).

When the stationary assumption is not fulfilled, a wavelet discrete transform can be used, as we illustrate in the example below. Wavelets transforms use a similar approach to Fourier decomposition but with two key differences. First, wavelet transforms come in a variety of shapes (the "haar," "Mexican hat," etc.; Dale et al. 2002), such that users need not assume only sine/cosine functions to describe spatial variation. Second, wavelet transforms do not assume stationarity, but instead allow for local variation in wavelet templates at different resolutions to be fit to observed data.

These techniques tend to be applied to dyadic grids of data, where the dimensions of the grid are of a power 2 (e.g., 32×32, 128×128), similar to our use of fractal algorithms in Chap. 3. The reason for this constraint is that it allows us to recursively decompose the spatial variation on the map. For instance, if we have a map of dimensions 64×64, this map can be broken into successive blocks representing different spatial resolutions, such as four blocks of 32×32 units, to interpret spatial variation. Note that wavelet analysis has been extended to work with maps that do not comply with this constraint, but that is beyond the scope of our application here.

Wavelets can be discrete or continuous. Here, we focus on the simplest, discrete transform, the Haar (Fig. 5.13). Using observed data, the discrete wavelet transform can be computed for each sampling location, or pixel, for a series of wavelet template scales that are as a power of two. Then, the wavelet values can be mapped and all scales analyzed. The plot of wavelet variance against the scaling factor is called a *scalogram* (Dale and Fortin 2014). The highest wavelet variance values indicate the spatial scales that fit the data best.

We calculate wavelets using the `waveslim` package (Whitcher 2015). We use the Haar wavelet, which is a common type of wavelet used in spatial analysis (Fig. 5.13). To calculate wavelets, we need to pass the maximum scale being considered, which should be a power of 2 (8, 16, 32, 64, etc.). First, we reformat

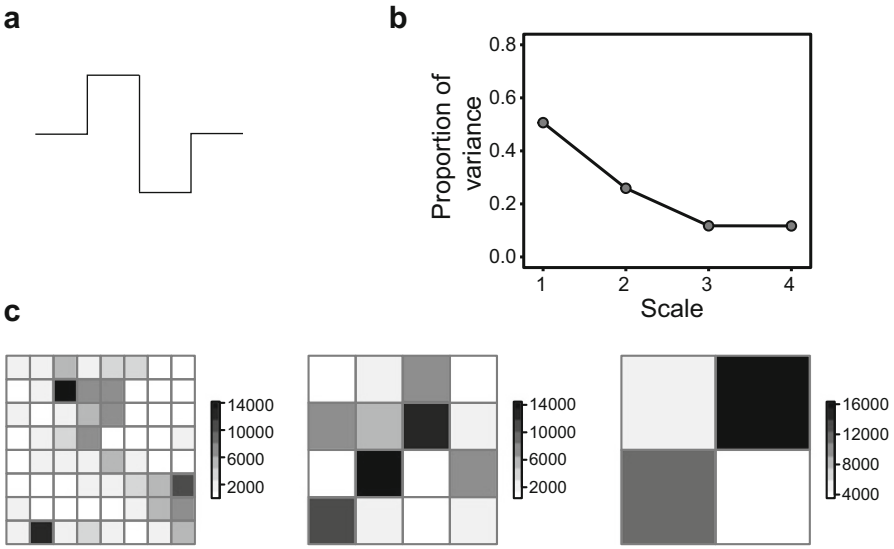

Fig. 5.13 Wavelets and scalograms. (**a**) The Haar template, which is a common discrete wavelet template. (**b**) The scalogram from the fitting the Haar template to a portion of the observed data, where variance is plotted as a function of scale. (**c**) Maps of the spatial variation identified at each scale (scales 1–3) from wavelet analysis

the data into a square matrix using the `acast` function from the `reshape2` package (Wickham 2007):

```
> library(reshape2)
> matrix.mat <- acast(matrix, x ~ y, value.var = "Height")
> dim(matrix.mat)

##
[1] 26 26
```

This reformatting emphasizes that the sampling grid used is a 26 × 26 grid. For the purposes here, we will subset the grid to become a dyadic grid that is 16 × 16.

```
> max.scale <- 4

#DWT: Discrete Wavelet Transform
> library(waveslim)
> x.dwt <- dwt.2d(matrix.mat[1:16, 1:16], 'haar', J = max.scale)
```

This function creates new matrices that describe the wavelet variance at different scales. To do so, it creates three bands of variation, labeled LH, HL, and HH. We sum the squared values of these bands to quantify a total measure of wavelet

variance. We can then calculate the proportion of variance at each scale and plot the scalogram (Fig. 5.13b).

```
#Sum the wavelet spectrums
> t.var <- (sum(x.dwt$LH1^2 + x.dwt$HL1^2 + x.dwt$HH1^2)
  + sum(x.dwt$LH2^2 + x.dwt$HL2^2 + x.dwt$HH2^2)
  + sum(x.dwt$LH3^2 + x.dwt$HL3^2 + x.dwt$HH3^2)
  + sum(x.dwt$LH4^2 + x.dwt$HL4^2 + x.dwt$HH4^2))

#proportional variance
> x.lev.1 <- (sum(x.dwt$LH1^2 + x.dwt$HL1^2 + x.dwt$HH1^2)) / t.var
> x.lev.2 <- (sum(x.dwt$LH2^2 + x.dwt$HL2^2 + x.dwt$HH2^2)) / t.var
> x.lev.3 <- (sum(x.dwt$LH3^2 + x.dwt$HL3^2 + x.dwt$HH3^2)) / t.var
> x.lev.4 <- (sum(x.dwt$LH4^2 + x.dwt$HL4^2 + x.dwt$HH4^2)) / t.var

> var.all.dwt <- c(x.lev.1, x.lev.2, x.lev.3, x.lev.4)
> sum(var.all.dwt)

#Scalogram: plotting global Wavelet spectrum profiles
> plot(var.all.dwt, pch = 21, type = "b", lwd = 1, ylab = "Average
  Variance", xlab = "Scale")
```

The scalogram suggests that most of the spatial variation occurs at the finest distance considered. Finally, we can plot the wavelet images (Fig. 5.13c) at each scale using the `raster` package. Below we show an example of scale 1.

```
#Map Wavelet values according to scales
> wave.raster1 <- raster((x.dwt$LH1^2 + x.dwt$HL1^2 +
  x.dwt$HH1^2))
> plot(wave.raster1)
```

These measures of variation at different scales can then be used as predictor variables in regression or related analyses to account for spatial dependence arising at different scales (Keitt and Urban 2005). The key to do so is to link the submatrices to the appropriate response data in a hierarchical way.

5.3.7.2 Eigenvector Spectral Decomposition

When data are sampled from an irregular grid or layout in a contiguous fashion, an eigenvector spectral decomposition can be used to identify the key scales that match the data. Here, we use principal coordinates of neighborhood matrices, PCNM, which is a special case of the generalized Moran's Eigenvector Map (Dray et al. 2006, 2012). Unlike the wavelet analysis that is performed on the data given the spatial layout of contiguous pixels, the PCNM multiscale analysis is performed on the *x*–*y* coordinates of the sampling locations. In this approach, there are potentially as many PCNM spatial scales as there are sampling locations. PCNM uses a

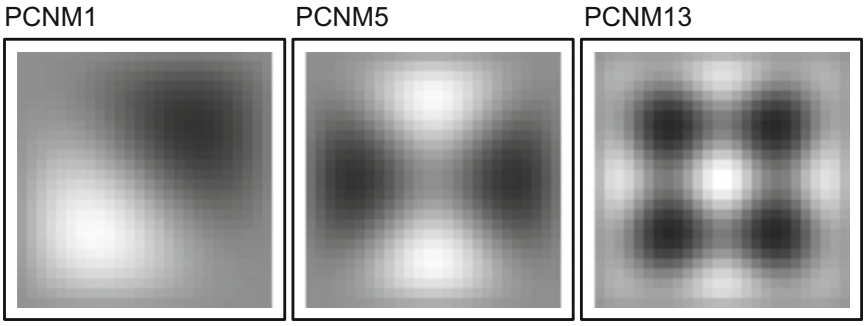

Fig. 5.14 Eigenvector mapping captures spatial variation at different scales. The first few eigenvectors capture broad-scale spatial variation, similar to what would be captured if a linear or polynomial term of x–y coordinates was considered (e.g., in trend surface analysis) whereas the latter eigenvectors capture increasingly fine scale variation. Shown are the first three eigenvectors selected with forward, stepwise regression on the observed vegetation data, where dark pixels indicate higher values of the eigenvectors

Principal Coordinates Analysis (PCoA; Gower 1966), which is also referred to as metric multidimensional scaling or classical scaling (Legendre and Legendre 2012). PCoA shares some similarities to the more common principal components analysis (PCA), but it focuses on using distance or similarity matrices (rather than the original data, as in PCA) to position objects in a space of lower dimensionality than the original data, with a focus on Euclidean distance space. PCNM provides eigenvectors that can capture spatial structure: as the eigenvectors increase, they capture increasingly finer scales of space that resemble sin waves of smaller and smaller periods (Fig. 5.14), analogous to Fourier's technique.

Once the PCNMs are computed they can be used as spatial predictors either in a multiple regression (Dormann et al. 2007) or other analyses, such as redundancy analysis for community data (see Chap. 11). Because there as many PCNM eigenvectors as the number of locations, model selection is necessary to reduce the number of eigenvectors considered. Narrowing the number of spatial scales can also be guided using knowledge about the scales of processes that may have generated the data; for example large-scale trend, intermediate patchiness, and small-scale patchiness. If we were to only consider the first few eigenvectors, this may be functionally similar to what has been referred to a trend surface analysis, where x or y coordinates are used as predictors (and potentially their polynomial terms, such as x^2 or x^3) in regression models to allow for large-scale variation in spatial dependence (see Chap. 6). In contrast, the larger eigenvectors capture fine-scale variation in spatial dependence.

We illustrate the eigenvector method with the grid used for kriging above. We can use the `vegan` package (Oksanen et al. 2018) to first determine the PCNM based on a distance matrix calculated from the sample locations. Eigenvectors for each sampling location are calculated.

```
> library(vegan)

#PCNM on distance matrix based on coords
> xypcnm <- pcnm(dist(coords))

#eigenvectors:
> xypcnm$vectors
```

We can visualize the eigenvectors in a variety of ways. Here we create a raster of the eigenvectors with the `raster` package and plot (Fig. 5.14).

```
#create raster
pcnm1.raster <- rasterFromXYZ(data.frame(x = matrix$x, y =
matrix$y, z = xypcnm$vectors[,1]))
plot(pcnm1.raster)
```

Once we have generated the eigenvectors, we can use them as predictors for vegetation height. Given the large number of eigenvectors that are generated, a common approach is to use a procedure to select a subset of the eigenvectors for further inclusion (Bauman et al. 2018). One approach is forward selection procedure to determine the eigenvectors that best explain the response variable (Dray et al. 2006; Blanchet et al. 2008). Other options are possible, such as using the reduction in spatial autocorrelation in the residuals of models (see Chap. 6; Dray et al. 2006; Dormann et al. 2007). Here, we simply focus on a forward selection procedure suggested by Blanchet et al. (2008), which has been shown to reliably capture multiscale spatial dependence in some situations (Bauman et al. 2018). This approach can be implemented in the `adespatial` package (Dray et al. 2018). In this approach, we first fit a full (global) model with all eigenvectors as covariates. From this model, we extract the adjusted R^2 value, which provides a measure of the variation explained from the eigenvectors (adjusted for the number of variables in the model). Blanchet et al. (2008) then proposed a forward selection approach with a "double-stopping" rule; that is, the forward selection terminates either when the current model reaches the adjusted R^2 of the full model, or when new eigenvectors are no longer significant, based on a prespecified α. Below, we implement this approach, using a conservative $\alpha = 0.005$, given the large number of potential covariates included.

```
> library(adespatial)
> height <- matrix$Height
> xypcnm.df <- data.frame(xypcnm$vectors)

#fit full model
> xypcnm.full <- lm(height ~ ., data = xypcnm.df)
> R2adj <- summary(xypcnm.full)$adj.r.squared
```

```
#forward selection with adespatial
> xypcnm.for <- forward.sel(height, xypcnm$vectors, adjR2thresh =
  R2adj, alpha = 0.005, nperm = 999)
```

In this case, we find that ten eigenvectors are retained in the linear regression model to explain the spatial variation in vegetation height (the total number retained might vary slightly due to the stochastic nature of the permutation test used; see Fig. 5.14 for plots of three of the selected eigenvectors). In Chap. 6, we dive deeper into spatial regression and how this method and others can account for spatial dependence while interpreting environmental relationships.

5.4 Next Steps and Advanced Issues

5.4.1 Local Spatial Dependence

Throughout this chapter, we have focused on "global" spatial statistics that assume stationarity. However, the intensity of spatial dependence can often vary across a region. When the stationary assumption is not valid, other methods can be used. There are two general ways in which uncovering local spatial dependence can occur (Brunsdon and Comber 2015). First, some approaches take global indices and use decomposition methods to understand the role of individual locations on the global statistic. For example, local indicators of spatial association, or LISA (Anselin 1995; Boots 2002), take indices like Moran's *I* to determine the contribution of each observation to the global statistic, which can subsequently be mapped to understand spatial variation in spatial dependence. The examples of using Moran's *I* in this chapter can be extended to interpret local Moran's *I* in both the `ncf` and `spdep` packages. Second, moving window analyses can be done to understand spatial dependencies in the chosen windows. In this approach, global statistics are applied, but only to the neighborhood (window) under consideration. See Chap. 3 for examples of moving window analyses.

5.4.2 Multivariate Spatial Dependence

In ecology, data are frequently multivariate. For instance, in community ecology we often work with matrices of species occurrence or abundance. In these cases, we may be interested in understand and accounting for multivariate spatial dependence (Dray et al. 2012). Many of the methods described in this chapter can be extended to multivariate data (Wackernagel 2003). When two types of data occur, cross-correlograms and cross-variograms can be used to understand the spatial dependence between variables (Wagner 2003). See Chap. 11 for some discussion of these approaches in the context of spatially structured communities.

5.5 Conclusions

The detection, characterization, and significance testing of spatial pattern is the first step to understand spatial ecological data and the processes that generated them. Spatial dependence commonly occurs in ecological data and it is often argued that failure to account for spatial dependence can impact inferences in ecology (Legendre 1993; Dormann et al. 2007; Beale et al. 2010). Here, we illustrate ways in which spatial dependence can be diagnosed in ecological data. This type of diagnosis can be useful to provide insights into why patterns occur in data and whether spatial dependence can be problematic for inferences on ecological patterns and processes.

Both correlograms and variograms provide useful insights for interpreting the magnitude and extent of spatial dependence in data. These approaches provide much richer information than single tests of spatial dependence, such as using a single Moran's *I* test statistic (Eq. 5.3). Correlograms have the benefit of providing a standardized metric (i.e., correlation coefficient) that can be compared across variables, while variograms provide a means to formally estimate the scale of spatial dependence through the use of model-based variograms and estimating the spatial range. Variogram modeling can also be used for interpolating spatial data via kriging, providing a formal means for predicting ecological patterns across space. A variety of techniques have been used for inferring the significance of spatial dependence, with Monte Carlo permutations providing perhaps the most flexibility in their applications. Multiscale analyses can also provide useful insights, in particular in situations where questions and/or data come from broad extents where multiple scales of spatial dependence may operate.

References

Acevedo MA, Fletcher RJ Jr (2017) The proximate causes of asymmetric movement across heterogeneous landscapes. Landsc Ecol 32:1285–1297

Anselin L (1995) Local indicators of spatial association: LISA. Geogr Anal 27(2):93–115

Austin MP (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecol Model 157(2–3):101–118

Bauman D, Drouet T, Dray S, Vleminckx J (2018) Disentangling good from bad practices in the selection of spatial or phylogenetic eigenvectors. Ecography 41:1–12

Beale CM, Lennon JJ, Yearsley JM, Brewer MJ, Elston DA (2010) Regression analysis of spatial data. Ecol Lett 13(2):246–264. https://doi.org/10.1111/j.1461-0248.2009.01422.x

Bivand R (2006) Implementing spatial data analysis software tools in R. Geogr Anal 38(1):23–40. https://doi.org/10.1111/j.0016-7363.2005.00672.x

Bivand R, Piras G (2015) Comparing implementations of estimation methods for spatial econometrics. J Stat Softw 63(18):1–36

Bivand RS, Pebesma EJ, Gomez-Rubio V (2013) Applied spatial data analysis with R. Use R! 2nd edn. Springer, New York

Bjørnstad ON, Falck W (2001) Nonparametric spatial covariance functions: estimation and testing. Environ Ecol Stat 8(1):53–70. https://doi.org/10.1023/a:1009601932481

Blanchet FG, Legendre P, Borcard D (2008) Forward selection of explanatory variables. Ecology 89(9):2623–2632. https://doi.org/10.1890/07-0986.1

Bolker BM (2003) Combining endogenous and exogenous spatial variability in analytical population models. Theor Popul Biol 64(3):255–270. https://doi.org/10.1016/s0040-5809(03)00090-x
Boots B (2002) Local measures of spatial association. Ecoscience 9(2):168–176
Brown JH, Mehlman DW, Stevens GC (1995) Spatial variation in abundance. Ecology 76 (7):2028–2043. https://doi.org/10.2307/1941678
Brunsdon C, Comber L (2015) An introduction to R for spatial analysis and mapping. Sage Publications, Inc, London
Burnham KP, Anderson DR (1998) Model selection and inference: a practical information-theoretic approach. Springer, New York
Carroll SS, Pearson DL (2000) Detecting and modeling spatial and temporal dependence in conservation biology. Conserv Biol 14(6):1893–1897. https://doi.org/10.1046/j.1523-1739.2000.99432.x
Cohen JM, Civitello DJ, Brace AJ, Feichtinger EM, Ortega CN, Richardson JC, Sauer EL, Liu X, Rohr JR (2016) Spatial scale modulates the strength of ecological processes driving disease distributions. Proc Natl Acad Sci U S A 113(24):E3359–E3364. https://doi.org/10.1073/pnas.1521657113
Cressie NAC (1993) Statistics for spatial data. Wiley, Chichester
Dale MRT, Fortin MJ (2014) Spatial analysis: a guide for ecologists, 2nd edn. Cambridge University Press, Cambridge
Dale MRT, Dixon P, Fortin MJ, Legendre P, Myers DE, Rosenberg MS (2002) Conceptual and mathematical relationships among methods for spatial analysis. Ecography 25(5):558–577. https://doi.org/10.1034/j.1600-0587.2002.250506.x
Dormann CF, McPherson JM, Araújo MB, Bivand R, Bolliger J, Carl G, Davies RG, Hirzel A, Jetz W, Kissling WD, Kuehn I, Ohlemueller R, Peres-Neto PR, Reineking B, Schroeder B, Schurr FM, Wilson R (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30(5):609–628. https://doi.org/10.1111/j.2007.0906-7590.05171.x
Dray S, Legendre P, Peres-Neto PR (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecol Model 196(3–4):483–493. https://doi.org/10.1016/j.ecolmodel.2006.02.015
Dray S, Pelissier R, Couteron P, Fortin MJ, Legendre P, Peres-Neto PR, Bellier E, Bivand R, Blanchet FG, De Caceres M, Dufour AB, Heegaard E, Jombart T, Munoz F, Oksanen J, Thioulouse J, Wagner HH (2012) Community ecology in the age of multivariate multiscale spatial analysis. Ecol Monogr 82(3):257–275. https://doi.org/10.1890/11-1183.1
Dray S, Bauman D, Blanchet G, Borcard D, Clappe S, Guenard G, Jombart T, Larocque G, Legendre P, Madi N, Wagner HH (2018) adespatial: multivariate spatial analysis. R package version 0.2-0
Efron B (1979) Bootstrap methods - another look at the jackknife. Ann Stat 7(1):1–26. https://doi.org/10.1214/aos/1176344552
Fletcher RJ Jr, Sieving KE (2010) Social-information use in heterogeneous landscapes: a prospectus. Condor 112:225–234
Fletcher RJ Jr, Acevedo MA, Robertson EP (2014) The matrix alters the role of path redundancy on patch colonization rates. Ecology 95(6):1444–1450
Fletcher RJ, Reichert BE, Holmes K (2018) The negative effects of habitat fragmentation operate at the scale of dispersal. Ecology 99(10):2176–2186
Giraudoux P (2018) pgirmess: spatial analysis and data mining for field ecologists. R package version 1.6.9
Goovaerts P (1994) Study of spatial relationships between two sets of variables using multivariate geostatistics. Geoderma 62(1–3):93–107. https://doi.org/10.1016/0016-7061(94)90030-2
Gower JC (1966) Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53(3–4):325–338
Haining R (2003) Spatial data analysis: theory and practice. Cambridge University Press, Cambridge
Hurlbert SH (1984) Pseudoreplication and the design of ecological field experiments. Ecol Monogr 54(2):187–211. https://doi.org/10.2307/1942661

Keitt TH, Urban DL (2005) Scale-specific inference using wavelets. Ecology 86(9):2497–2504. https://doi.org/10.1890/04-1016
Keitt TH, Bjørnstad ON, Dixon PM, Citron-Pousty S (2002) Accounting for spatial pattern when modeling organism-environment interactions. Ecography 25(5):616–625
Koenig WD (1998) Spatial autocorrelation in California land birds. Conserv Biol 12(3):612–620. https://doi.org/10.1046/j.1523-1739.1998.97034.x
Koenig WD, Liebhold AM (2016) Temporally increasing spatial synchrony of North American temperature and bird populations. Nat Clim Chang 6(6):614. https://doi.org/10.1038/nclimate2933
Landeiro VL, Magnusson WE (2011) The geometry of spatial analyses: implications for conservation biologists. Natureza & Conservacao 9(1):7–19. https://doi.org/10.4322/natcon.2011.002
Lantuéjoul C (2002) Geostatistical simulation, models, and algorithms. Springer, Berlin
Legendre P (1993) Spatial autocorrelation: trouble or new paradigm? Ecology 74(6):1659–1673
Legendre P, Legendre L (2012) Numerical Ecology, 3rd Edition. Elsevier, Amsterdam
Nelson TA, Boots B (2008) Detecting spatial hot spots in landscape ecology. Ecography 31(5):556–566. https://doi.org/10.1111/j.0906-7590.2008.05548.x
Oksanen J, Guillaume B, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens HH, Szoecs E, Wagner H (2018) Vegan: community ecology package. R version 2.4-6
Oliver MA, Webster R (1991) How geostatistics can help you. Soil Use Manag 7(4):206–217. https://doi.org/10.1111/j.1475-2743.1991.tb00876.x
Oliver MA, Webster R (2014) A tutorial guide to geostatistics: computing and modelling variograms and kriging. Catena 113:56–69. https://doi.org/10.1016/j.catena.2013.09.006
Pebesma EJ (2004) Multivariable geostatistics in S: the gstat package. Comput Geosci 30(7):683–691. https://doi.org/10.1016/j.cargo.2004.03.012
Peres-Neto PR, Legendre P (2010) Estimating and controlling for spatial structure in the study of ecological communities. Glob Ecol Biogeogr 19(2):174–184. https://doi.org/10.1111/j.1466-8238.2009.00506.x
Remmel TK, Fortin MJ (2013) Categorical, class-focused map patterns: characterization and comparison. Landsc Ecol 28(8):1587–1599. https://doi.org/10.1007/s10980-013-9905-x
Ribeiro PJ, Jr., Diggle PJ (2016) geoR: analysis of geostatistical data. vol R package version 1.7-5.2
Rossi RE, Mulla DJ, Journel AG, Franz EH (1992) Geostatistical tools for modeling and interpreting ecological spatial dependence. Ecol Monogr 62(2):277–314. https://doi.org/10.2307/2937096
Schlather M, Malinowski A, Menck PJ, Oesting M, Strokorb K (2015) Analysis, simulation and prediction of multivariate random fields with package random fields. J Stat Softw 63(8):1–25
Schooley RL, Wiens JA (2004) Movements of cactus bugs: patch transfers, matrix resistance, and edge permeability. Landsc Ecol 19(7):801–810
Tobler WR (1970) Computer movie simulating urban growth in Detroit region. Econ Geogr 46(2):234–240. https://doi.org/10.2307/143141
Venables WN, Ripley BD (2002) Modern applied statistics with S, 4th edn. Springer, New York
Wackernagel H (2003) Multivariate geostatistics: an introduction with applications, 3rd edn. Springer, Berlin
Wagner HH (2003) Spatial covariance in plant communities: integrating ordination, geostatistics, and variance testing. Ecology 84(4):1045–1057. https://doi.org/10.1890/0012-9658(2003)084[1045:scipci]2.0.co;2
Wagner HH, Fortin MJ (2005) Spatial analysis of landscapes: concepts and statistics. Ecology 86(8):1975–1987. https://doi.org/10.1890/04-0914
Whitcher B (2015) waveslim: basic wavelet routines for one-, two- and three-dimensional signal processing. R package version 1.7.5
Wickham H (2007) Reshaping data with the reshape package. J Stat Softw 21(12):1–20
Yoo J, Ready R (2016) The impact of agricultural conservation easement on nearby house prices: incorporating spatial autocorrelation and spatial heterogeneity. J For Econ 25:78–93. https://doi.org/10.1016/j.jfe.2016.09.001