

Chapter 6

Accounting for Spatial Dependence in Ecological Data



6.1 Introduction

Inference and prediction are fundamental to all aspects of ecology and conservation. Yet the presence of dependency in the data due to either phylogeny, space, or time can impair the statistical inference and subsequent ecological interpretation of the pattern(s) observed (Sokal and Oden 1978; Swihart and Slade 1985; Garland et al. 1992; Lennon 2000; Miller 2012). In this chapter, we will focus specifically on how the presence of spatial dependency complicates our ability to make statistical inferences and prediction (Legendre 1993), as the principles due to space are analogous to those due to time and phylogeny (Bauman et al. 2018). It is important to understand how statistical biases due to spatially structured data can affect answering a wide array of ecological questions ranging from species–environment relationships to predicting the spread of invasive species. Consequently, there is an increasing emphasis on formally accounting for spatial dependence in inferential problems in ecology and conservation (Segurado et al. 2006; Dormann et al. 2007; Hooten et al. 2007; Carroll and Johnson 2008; Beale et al. 2010; Crase et al. 2014).

Accounting for spatial dependence in modeling is, however, very challenging. This challenge arises because spatial dependence in data can emerge for a variety of reasons (see Chap. 5). In particular, when modeling spatial data, spatial dependence can occur simply due to model mis-specification, such as an important covariate not being included in the model or that its functional relationship is mis-specified (e.g., effects may be non-linear). Spatial dependence could also occur through processes such as localized dispersal or social behavior (Koenig 1999). In these cases, adding environmental covariates will likely not be sufficient for appropriate inferences.

Here, we provide an overview regarding several ways in which space has been addressed in regression-like models of species–environment relationships. Regression models are frequently used in ecology and conservation to address a variety of problems, ranging from interpreting habitat suitability to forecasting the effects of climate change (Guisan and Zimmermann 2000; Algar et al. 2009). Our overview is

largely guided by some comprehensive reviews and syntheses on the topic (Keitt et al. 2002; Dormann et al. 2007; Miller et al. 2007; Diniz et al. 2009; Bini et al. 2009; Beale et al. 2010), but we update these syntheses with more recent advances (Crase et al. 2012; Rousset and Ferdy 2014; Bardos et al. 2015; Blangiardo and Cameletti 2015; Ver Hoef et al. 2018). Our goals are threefold. We first describe the problem of spatial dependence on inferences in ecology and conservation. Then, we discuss how to diagnose problems of spatial dependence. Finally, we illustrate common ways to address these statistical problems using a variety of approaches aimed at accounting for spatial dependence in statistical analyses.

6.2 Key Concepts and Approaches

6.2.1 *The Problem of Spatial Dependence in Ecology and Conservation*

Bivand (1980) was one of the first to explore the importance of spatial dependence on statistical inference from correlation coefficients, a problem that Legendre (1993) later clearly illustrated for ecology. These articles highlight how spatial correlations may create spurious inference and ecological interpretation when spatial dependency of the data is ignored (Fig. 6.1). Depending of the magnitude of spatial autocorrelation (see Chap. 5), parameter estimation can be erroneous and hence our subsequent understanding of ecological patterns and processes: at small values of spatial autocorrelation (e.g., <0.2) the effect tends to be negligible, whereas when the value of spatial autocorrelation is high (e.g., >0.2) then the effect tends to be important and will affect statistical inferences (Bivand 1980). The reason for this problem generally lies in the estimation of uncertainty, where standard errors and confidence intervals around point estimates of correlation coefficients (and other parameters) tend to be artificially narrow. This issue can be considered from the point of degrees of freedom (df), where one df is counted for each independent observation. Yet spatial dependence causes observations to not be independent, such that each observation should not be counted as one df. In effect, this issue essentially leads to “pseudo-replication” in space, a well-known problem for ecology (Hurlbert 1984).

This problem has several practical consequences for conservation. For example, Crase et al. (2014) illustrated that ignoring spatial dependence in forecasts of species response to climate change leads to greater estimated effects of climate change. Ignoring spatial dependence has also been shown to affect conservation planning and understanding habitat suitability for a wide range of species of conservation concern (Carroll and Johnson 2008; Lichstein et al. 2002; Carroll et al. 2010).

Several approaches have been proposed to account for spatial dependence in statistical analyses and modeling (Keitt et al. 2002; Dray et al. 2006; Carl and Kuhn 2010). In the simplest approaches, we might subset data such that sample points are

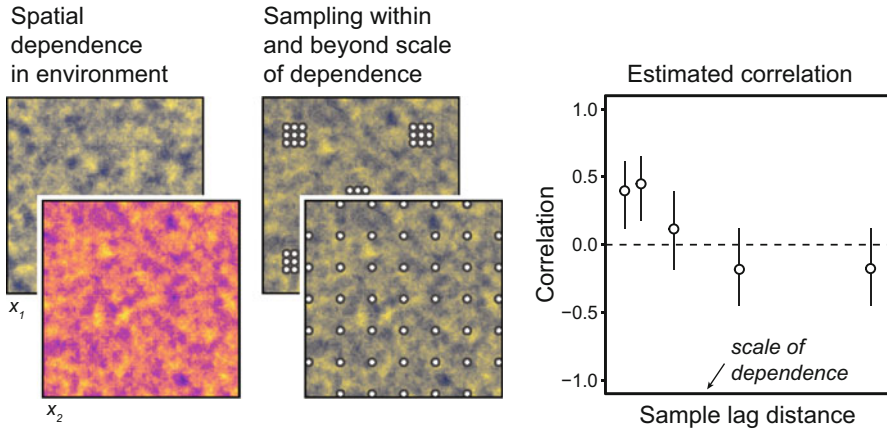


Fig. 6.1 The problem of spatial dependence for ecological inferences. When spatial dependence occurs and is ignored, type I error rates increase. Shown are two, independently derived, environmental variables, x_1 and x_2 , that have spatial dependence (generated from a Gaussian random field; see Chap. 5). If sampling occurs within the range of spatial dependence, spurious inferences can occur when such dependence is ignored. In contrast, if sampling is implemented beyond the range of spatial dependence, reliable inference is obtained. Shown are results from Pearson correlation coefficient between the environmental variables taken from five sampling designs that vary in their spatial distribution based on lag distance (each has the same number of samples). Correlations are high when sampling is implemented within the range of spatial dependence, but declines as the lag distance between samples increases

greater than the range of estimated spatial autocorrelation (Chap. 5) (Hawkins et al. 2007), or perhaps just adjust α levels in statistical tests to be more conservative (Dale and Fortin 2014). Some of the most common approaches focus on extending linear regression models to accommodate spatial dependence by either using *autocovariate* variables (Table 6.1) to account for spatial dependence (Augustin et al. 1996; Wagner and Fortin 2005; Betts et al. 2006; Melles et al. 2011) or geostatistical models (see Chap. 5; Cressie 1993). Ordination techniques for community data can also be used to account for spatial structure in the data (see Chap. 11; Wagner 2003, 2004; Dray et al. 2012). Below we explain some of the most common approaches in detail. To do so, we first reintroduce the generalized linear model, which was briefly described in Chap. 2, and use this model framework to build from for accounting for spatial dependence.

6.2.2 The Generalized Linear Model and Its Extensions

Before jumping into approaches aimed at dealing with spatial dependence, we briefly discuss some critical background material. As a reminder, linear regression

Table 6.1 Common terms for spatial regression analysis in ecology

Term	Description
Aerial data	Spatial polygon data that are typically exhaustive tessellations of an area.
Autocovariate	A predictor variable that quantifies the frequency (or related metrics) of the response variable in the surrounding neighborhood.
Autoregressive model	Models that use information on the neighborhood matrix to account for spatial dependence based on deviations from the expected values.
Fixed effect	Deterministic effects that are constant across samples.
Lattice data	Spatial data indexed over a regularly spaced set of points.
Multilevel model	A type of mixed model, where random effects are used to capture hierarchies in the system.
Neighborhood matrix	A square matrix (dimensions are the number of sample points) that quantifies relationships between sampling points, such as binary neighbor connections or distance-weighted linkages.
Random effect	Effects that come from a distribution and vary across samples.
Residual	The difference between the observed value of the dependent variable and the predicted value.
Spatial filtering	When fixed effect covariates in a regression are added that attempt capture the spatial signal through the inclusion of functions of x - y coordinates or related distance metrics.
Tessellation	An arrangement of polygons closely fitted together without gaps or overlapping boundaries.
Trend surface analysis	Analyses where variation in the response variable is expressed as a function of the geographic coordinates of the sampling locations.

and ANOVA are types of linear models (Nelder and Wedderburn 1972). A linear model can be described as:

$$y_i = \alpha + \beta_1 x_i + \varepsilon_i, \quad (6.1)$$

where y_i is the response variable for sampling unit i (e.g., density of a species at a location), α is the intercept, β_1 is the slope (coefficient), x_i is the explanatory variable measured at i , and ε_i is the error, which is assumed to come from a normal distribution and be iid = independent and identically distributed. That is, each residual i is not dependent on other residuals and each comes from the same underlying distribution. This error distribution is assumed to come from a normal distribution with a mean of zero and an unknown finite variance, written as $\varepsilon_i \sim N(0, \sigma^2)$. Plotting the residuals of the model, or the deviation of the predictions to the observed data for a given value of x (Fig. 6.2), helps understand whether this assumption is met. Note that the equivalence of linear regression and ANOVA in this framework can be seen by considering categorical treatments (x_i) in an ANOVA as “dummy” variables (e.g., 0, 1 variables) in a regression model.

Linear models can be extended in two very useful ways. The first major extension, the *generalized linear model* (GLM), allows for alternative distributions for the response variable other than the normal distribution. These other distributions specifically come from the exponential family of distributions, which includes

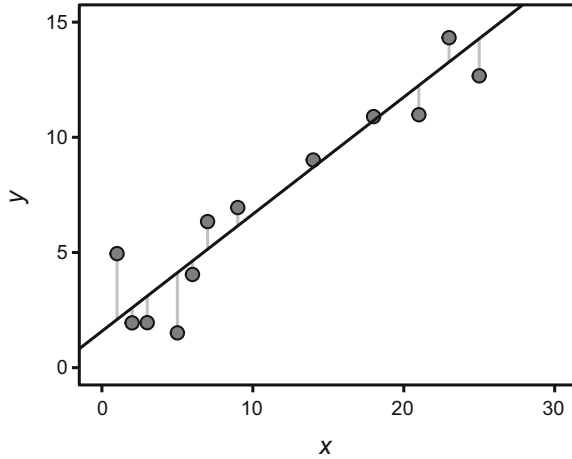


Fig. 6.2 A linear regression model and the residuals from that model. In standard regression techniques, residuals (the difference of the observed value from the predicted value of the response variable for a given value of the explanatory variable) are assumed to be independent and identically distributed. When spatial autocorrelation occurs in the residuals of models, such autocorrelation can impact inference if ignored. Dots represent observed data, black line is the prediction from the linear model, and the vertical gray lines represent residuals

distributions such as the Poisson, binomial, Bernoulli, and gamma distributions. This extension greatly increases the flexibility of these models, allowing for responses such as the presence/absence of a species at a sampling location (a Bernoulli distribution). The classic text for generalized linear models is McCullagh and Nelder (1989). In GLMs, we specify a link function and a distribution for the errors (ε).

Perhaps the two most common types of GLMs are logistic regression and Poisson (or log-linear) regression. For logistic regression we have:

$$\text{logit}(p_i) = \alpha + \beta_1 x_i, \quad (6.2)$$

where p_i is the expected probability of a “success” and a “logit” link function is used (i.e., $\log(p_i/(1 - p_i))$). In this case, we assume a binomial error distribution. A binomial distribution can be thought of as a distribution that arises from a series of coin tosses. If there is only one toss, it is called a Bernoulli distribution; if there is more than one toss (sometimes referred to as “trials”), then the distribution is called a binomial distribution. In the latter case, we are interested in the frequency or proportion of “successes” out of the total number of trials.

For a Poisson regression, we have:

$$\log(\lambda_i) = \alpha + \beta_1 x_i, \quad (6.3)$$

where λ_i is the expected count for sample i and we use a “log” link function and assume a Poisson error distribution. The Poisson distribution is a discrete

distribution were values are integers greater than or equal to zero (i.e., negative values are not allowed). The Poisson distribution assumes that the mean equals the variance, which is often a restrictive assumption. A related distribution that relaxes this assumption is the negative binomial distribution. There are several other types of GLMs; however, we will focus on only a few in this book. Interested readers should see Bolker (2008) and Bolker et al. (2009) for the use of GLMs in ecology.

The second major extension of a linear model is to allow for random effects, what is frequently termed a random-effects model, or if fixed effects are considered alongside random effects, a mixed model. *Random effects* can be contrasted with *fixed effects* (the β above) in several ways. Random effects are extremely flexible in how they can accommodate complex data structures and provide inference unattainable with fixed effects. Some uses for random effects include: (1) conditional inference—when you would like to make inferences on a particular sampling unit, location, etc. (e.g., a particular watershed contained within the study area); (2) accommodating block, split-plot, Latin-square, and other treatment structures in experiments; (3) more generally accounting for both temporal and spatial dependencies in data, such as temporal repeated measures or spatial autocorrelation; (4) when one thinks treatment effects may vary in space or time (similar to including an “interaction” term in a linear model); and (5) “broad-sense” inference: making inferences for an entire region/population from a sample (in contrast to “narrow-sense” inference, where we make inferences only for the specific samples or locations being considered) (Littell et al. 2006; Gelman and Hill 2007; Zuur et al. 2009).

There has been some confusion in ecology regarding when an effect should be considered random versus fixed, and how inferences may change depending on whether a variable is considered random or fixed. Gelman and Hill (2007) discussed how random effects have been loosely described and used in the literature, and the resulting problems that have arisen. We do not focus on this issue; rather we will simply consider mixed models as one means to accommodate spatial dependence.

We can formally describe a linear mixed model as:

$$y_i = \alpha + \beta_1 x_i + \gamma + \varepsilon_i, \quad (6.4)$$

where γ is a random effect and is typically assumed to be distributed $\sim N(0, \sigma^2)$. When we put these two extensions together, we have *generalized linear mixed models* (GLMMs), which are very powerful models that are seeing increasing use in ecology, evolution, and conservation (Bolker et al. 2009; Thorson and Minto 2015). Note that we can also model the variance, σ^2 , as a variance–covariance matrix, which is how we specifically extend this model to explicitly account for spatial dependence, as we will see below.

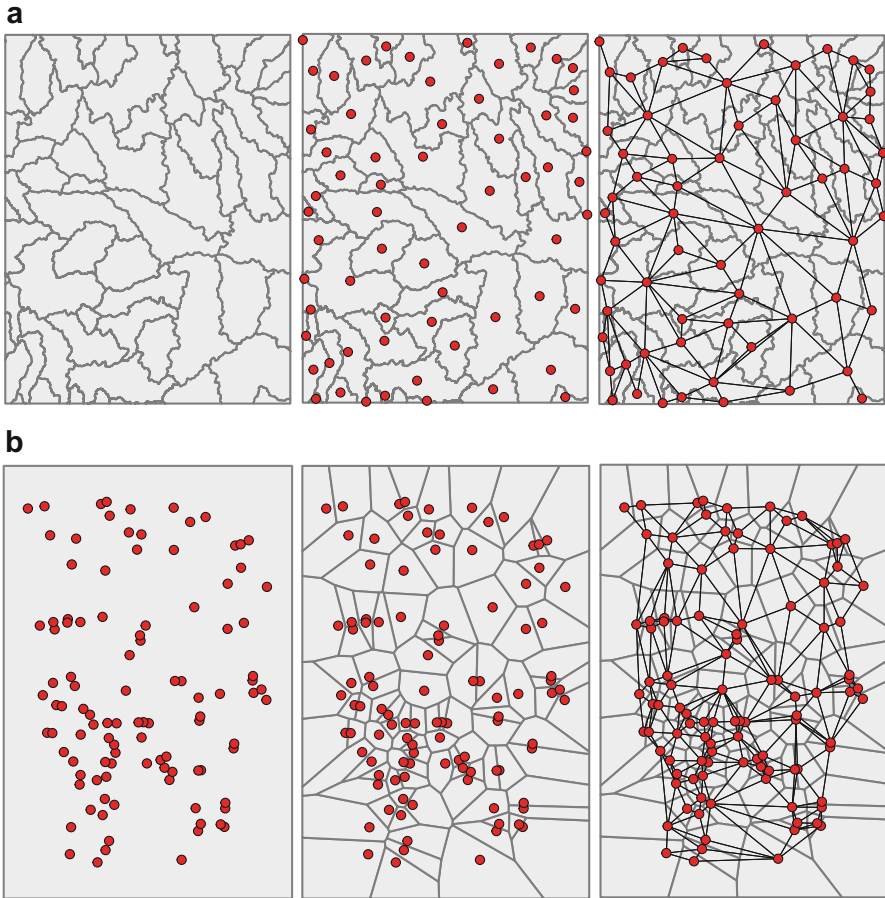


Fig. 6.3 Examples of aerial data used in spatial modeling. Aerial data can come from (a) polygon-based information (e.g., maps of counties, watersheds, etc.) or can be generated (b) from point or line data using Voronoi tessellation. In either approach, we can describe spatial dependence through the links among locations (right panel) with a spatial neighborhood (weights) matrix

6.2.3 General Types of Spatial Models

The vast diversity of spatial regression-like models can be organized in several ways. Three important properties include: (1) the type of response data (quantitative, count, presence–absence); (2) whether samples are irregularly spaced samples across continuous space or lattice/gridded data that are discrete in nature (Fig. 6.3); and (3) the way in which spatial dependence is considered.

The type of response data used will ultimately affect the type of regression model being fit. Different types of response data lend themselves to different distributions

used in GLM-like models. Overall, most of the approaches to spatial dependence have been better developed for normally distributed response variables than non-normally distributed response variables (Beale et al. 2010). Dealing with non-normal response data is generally more challenging than normally distributed data. For instance, data such as presence–absence data (0/1 data) have much less information content than normally distributed response variables, which impact the ability to identify, interpret, and account for spatial dependence in models.

Samples of data frequently come from *aerial data* (or *lattice data*) where neighborhoods are considered, such as samples arising from counties or watersheds. In such cases, spatial dependence is frequently considered based on neighboring polygons or related neighbors through the use of a *neighborhood matrix* (or spatial weights matrix). In contrast to aerial data, samples can also come from points across a study region. In this case, information on x – y coordinates are used either directly (e.g., using an x -coordinate as a predictor) or indirectly (e.g., by calculating distances between pairs of points).

Models can also be categorized based on how spatial structure is considered. For some models, often referred to as *spatial filtering* models (Getis and Griffith 2002), space is considered as predictor variables in a regression, where we attempt to “filter out” the spatial signal through the inclusion of functions of x – y coordinates or related distance metrics. In these cases, spatial dependence is thought to be largely dominated by exogenous drivers such as spatial dependence in environmental gradients, and often (but not always) occurs at relatively large scales (Fortin et al. 2012). In contrast, other models focus specifically on accounting for spatial dependence in the error terms of regression models. These models frequently assume spatial dependence is more localized and dominated mostly by endogenous processes (e.g., localized dispersal, species interactions) (Fortin et al. 2012; Teng et al. 2018).

6.2.4 Common Models that Account for Spatial Dependence

6.2.4.1 Trend Surface Analyses

Trend surface analyses use x – y coordinates in an attempt to capture large-scale spatial dependence in a region. There have been two common ways in which coordinates are added to regression models: polynomial regression (Haining 2003) and generalized additive models (GAMs) (Zuur et al. 2009).

The idea of trend surface analysis with polynomial regression is simply to include x – y coordinates and their polynomials (e.g., x^2 , x^3 , etc.) in the regression as covariates (Legendre 1993). Incorporating coordinates in this way is thought to be useful to deal with large-scale dependencies arising from exogenous processes (e.g., climate gradients across a geographic range), but it may be more limited in accounting of local autocorrelation. Legendre (1993) suggested simply adding quadratic and cubic terms for x – y coordinates to the regression model (Fig. 6.4). Adding quadratic and cubic terms allows for some potential non-linear responses across geographic

Fig. 6.4 Incorporating polynomial terms into a regression model to account for non-linearity in environmental relationships. Shown is an example of a linear model, contrasted with a model that adds a quadratic term, and a model that includes both a quadratic and cubic term

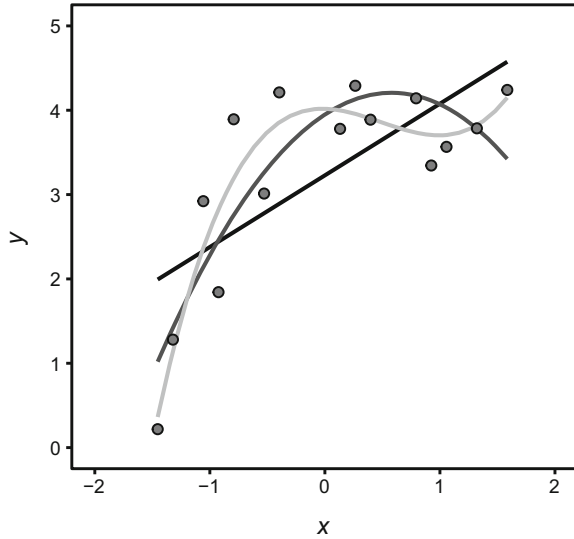
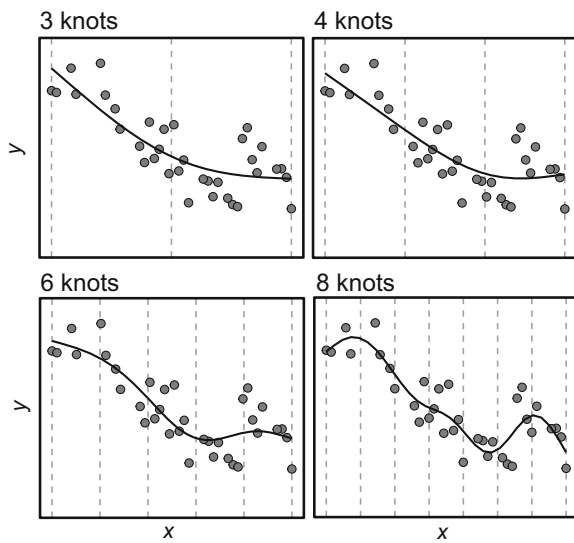


Fig. 6.5 The generalized additive model (GAM) and the concept of smoothers. Shown are GAMs fit to the data based on different numbers of knots (vertical lines; ranging from 3 to 8 knots). Within each knot, a simple spline (e.g., a cubic spline; see Fig. 6.4) is fit, with the constraint that splines must connect at the knots. As the number of knots increases, the complexity of the smoother function increases. Modified from Zuur et al. (2009)



space. Note that trend surface analysis will not formally adjust estimates of fixed effects for uncertainty due to spatial dependence (unlike mixed models, see below), but they may account for dependence in model residuals.

Generalized additive models (GAMs) (Hastie and Tibshirani 1986; Wood 2006) can be used in a similar way to trend surface analysis based on polynomial regression. GAMs use a class of equations called “smoothers” that attempt to generalize data into smooth curves by local fitting to subsections of the data (Fig. 6.5). This approach allows for more flexibility in capturing non-linearity in responses across

geographic space and GAMs have frequently been used in species distribution modeling more broadly (see Chap. 7). The simplest example of a smoother that is likely to be familiar to scientists is the running average, where one calculates the average value of data in a “window” across values of a covariate. While the running average is an example of a smoother, much more efficient smoothers have been developed. LOWESS (i.e., locally weighted regression; Cleveland 1979) is one example of a more efficient smoother used in some GAMs. The idea is to plot the value of the dependent variables (e.g., occurrences) along a single environmental variable, and then to calculate a smooth curve that fits the data as closely as possible while being parsimonious based on some sort of criterion. The algorithm fits a smooth curve to each variable and then combines the results additively. The approach generally employed with GAMs is to divide the data into some number of sections, using “knots” at the ends of the sections. Then a low-order polynomial or spline function (a spline is a function of polynomials relationships stitched together) is fit to the data in the section, with the added constraint that the second derivative of the function at the knots must be the same for both sections sharing that knot. This latter criterion eliminates kinks in the curve, and ensures that it is smooth and continuous (Fig. 6.5).

6.2.4.2 Eigenvector Mapping

Eigenvector mapping extends the general eigenvector approach described in Chap. 5 by using eigenvectors that describe different scales of spatial structure as predictors in regression models (Dray et al. 2006; Griffith and Peres-Neto 2006). In effect, this is somewhat similar to a trend surface model, but where eigenvector values, rather than x - y coordinates, are used as predictors. The ability of this approach to capture multiple scales of potential spatial structure is a relatively unique benefit in contrast to other approaches. Because each eigenvector captures spatial patterns at different scales, the combined use of several eigenvectors can potentially address problems of anisotropy and non-stationarity in spatial autocorrelated data. However, this approach and related techniques can sometimes lead to bias in coefficients of fixed effects and may not improve Type I error rates (Beale et al. 2010; Emerson et al. 2015).

Spatial eigenvectors are derived from a distance matrix from sample points, typically through the use of principal coordinates analysis (PCoA) on distance matrices (see Chap. 5; Dray et al. 2006). In this approach, a pairwise distance matrix is first calculated between all sampling points. This distance matrix is converted to a binary connectivity (or weights) matrix based on some distance threshold that allows for a minimum representation of connectivity among all points. For instance, a “minimum spanning tree,” which is the minimum set of links that ensures all points being considered are connected, is often used as a parsimonious way to guarantee connectivity among all points considered (see below). With this binary connectivity matrix, PCoA (also known as classic multidimensional scaling) is performed. PCoA generates new components that capture the variation in the distance matrix, which

are summarized with eigenvalues and eigenvectors, similar to Principal Components Analysis (Legendre and Legendre 1998). The set of eigenvectors that reduce or eliminate spatial autocorrelation in the residuals of the models is then identified. This can be assessed through the use of Moran's I on the residuals of models that include eigenvectors as predictors (Dray et al. 2006). Those eigenvectors that reduce autocorrelation the most are then used as predictors in a standard regression model to "filter out" spatial dependence.

6.2.4.3 Autocovariate Models

In these and related models, we typically work with "areal" or "lattice" data, rather than point-based samples. Autocovariate regression is similar to linear regression, but an *autocovariate* is included into the regression model. This autocovariate can be defined in various ways, such as a weighted mean of the response variable in surrounding locations (Augustin et al. 1996):

$$\text{auto}_i = \frac{\sum_{j=1}^{k_i} w_j y_j}{\sum_{j=1}^{k_i} w_j}, \quad (6.5)$$

where auto_i is the spatially weighted mean of the response variable, y , in the neighborhood (with a neighbor set k_i , reflecting the size of neighborhood considered) around sample i . This autocovariate is frequently calculated based on first-order neighbors (e.g., adjacent polygons or surrounding eight cells in a lattice), but the idea can be extended to account for further away points, typically weighting points based on the inverse of the distance (samples farther away get less weight than those closer to the sample). This approach can be used in a generalized linear model context; for instance, applying autocovariates in logistic regression, termed autologistic regression, is a common approach in ecology (Augustin et al. 1996; Wintle and Bardos 2006).

In effect, this approach assumes that if nearby locations are occupied, there should be a greater likelihood that the focal point is occupied. This is a relatively simple approach, although in practice, it was shown to not perform well because it can cause bias in coefficients of fixed effects for environmental predictor variables (Dormann et al. 2007; Beale et al. 2010). In these cases, autocovariate models tended to de-emphasize the effect of the environmental covariates, while overemphasizing the effects of autocovariates, leading to Type II error in inferences on environmental covariates. This issue is at least partly driven by the fact that the autocovariate is calculated on the raw data before fitting the explanatory variables, even though explanatory variables may contain spatial dependence that can reduce spatial dependence in the residuals of models (Crane et al. 2012). There are also difficulties with using these models to interpolate (or extrapolate) to new locations (see below).

Crane et al. (2012) developed a related approach in which autocovariates are quantified from the residuals of models, rather than the raw data, termed the residual

autocorrelation approach (RAC). This approach replaces the use of the raw data (y_i in Eq. (6.5)), with $y_i - q_i$, where q_i is the fitted value from an environment-only model that ignores autocorrelation. This leads to an autocovariate that captures only the variance not explained by explanatory variables. Crase et al. (2012) argued that this approach better captures spatial dependence than using standard autocovariates because explanatory variables are fitted first to the data.

Bardos et al. (2015) raised concerns regarding the validity of prior analyses (Dormann et al. 2007; Beale et al. 2010) that emphasized bias in auto-models. They show that a weighting scheme based on weighted means (Eq. 6.5) is not valid for autocovariate models. Rather a weighted sums scheme should be used instead:

$$\text{auto}_i = \sum_{j=1}^{k_i} w_j y_j. \quad (6.6)$$

This weighting scheme has not been evaluated as thoroughly as a weighted means approach described above, but Bardos et al. (2015) illustrated that it may perform better, in terms of capturing autocorrelation and providing unbiased estimates of fixed effects.

6.2.4.4 Autoregressive Models

Autoregressive models work with aerial or lattice data, similar to autocovariate models. The difference lies in how spatial dependence is captured with these model formulations. Two common *autoregressive models* are simultaneous autoregressive models (SAR) and conditional autoregressive models (CAR) (Lichstein et al. 2002; Ver Hoef et al. 2018). In both SAR and CAR, spatial dependence is captured through the use of a spatial neighborhood weights matrix akin to autocovariate models, but dependence is described based on deviations from the expected value given the covariates (Keitt et al. 2002).

SAR and CAR models share several similar features. In practice, a primary difference is that SAR can accommodate anisotropic spatial dependence, while CAR cannot. Nonetheless, the CAR is often used. Also, note that some work suggests that both CAR and SAR perform well on regular lattices, but suffer diminished performance on irregular lattices (e.g., county or watershed data) (Wall 2004). Both of these models use a spatial weights matrix, \mathbf{W} , that captures the neighborhood surrounding sampling locations. Typically, \mathbf{W} is a binary matrix that identifies neighbors, but it could also include non-binary data.

The general CAR model can be written in matrix notation as:

$$\mathbf{y} = \boldsymbol{\beta}\mathbf{X} + \rho\mathbf{W}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) + \boldsymbol{\varepsilon}, \quad (6.7)$$

where ρ is the first-order autocorrelation between neighbors, $\boldsymbol{\beta}$ is a vector of coefficients (i.e., slopes) related to the explanatory variables \mathbf{X} described through the “design matrix” (i.e., a $N \times K$ matrix of explanatory variable values for each sample of data used in model fitting, where N is the total number of samples and K is the total number of explanatory variables). In this equation, $\boldsymbol{\beta}\mathbf{X}$ is the same as a standard regression (Eq. 6.1) written in matrix form (i.e., Eq. 6.1 can be rewritten in matrix form as $\mathbf{y} = \boldsymbol{\beta}\mathbf{X} + \boldsymbol{\varepsilon}$), such that the only difference in this equation and a standard GLM is that the $\boldsymbol{\varepsilon}$ in the standard GLM is now broken into $\rho\mathbf{W}(\mathbf{y} - \boldsymbol{\beta}\mathbf{X}) + \boldsymbol{\varepsilon}$. The $(\mathbf{y} - \boldsymbol{\beta}\mathbf{X})$ captures the deviation of the observed data from that expected from the covariates and this is multiplied by the correlation for the neighbors ($\rho\mathbf{W}$; note that this only captures the neighbors because \mathbf{W} is 0 for all non-neighbors).

There are several types of SAR models that capture different kinds of spatial dependence, which assume that the dependence occurs in the response variable, predictor variables, or the error (Dormann et al. 2007). The general SAR model can be written in matrix notation as:

$$\mathbf{y} = \boldsymbol{\beta}\mathbf{X} + \rho\mathbf{W}\mathbf{y} + \boldsymbol{\varepsilon}. \quad (6.8)$$

While there are several types of SAR models, Ver Hoef et al. (2018) did not recommend the use of certain specifications of SAR models for ecological data, such as the use of “lag” or “SAR mixed models.” See Kissling and Carl (2008), Dale and Fortin (2014), and Ver Hoef et al. (2018) for more details.

6.2.4.5 Multilevel Models

The effects of potential spatial dependence can be also handled by using “multilevel” or “hierarchical” modeling. This type of modeling is a natural extension of generalized linear models, where we specify random effects to account for dependencies (correlations and hierarchical structure) in the data. Thus, multilevel models can be considered one type GLMM. An excellent text on this approach is Gelman and Hill (2007). Keitt et al. (2002) also touched on this approach when they contrasted “blocking” with other approaches to addressing spatial dependence.

Multilevel models are relevant when there is a natural hierarchical structure to the data being used (Fortin et al. 2012). For example, point samples may be collected in a grid or along a transect (with replicate grids or transects across a region), samples may be nested with counties or watersheds nested within larger regions (e.g., states), etc. In the absence of such sampling structure, this framework may not be helpful for accounting for spatial dependences. Some reasons to consider multilevel models with spatial data: (1) it can accommodate using all the data to perform inferences when some groups or blocks have small sample size; (2) it provides more efficient inference for regression parameters; (3) it can appropriately include predictors at >1 level in a hierarchy (e.g., within patch, patch, and landscape predictors); and (4) it can provide correct estimates of uncertainty (standard error, confidence interval, etc.) (Gelman and Hill 2007). For example, if we collect multiple samples within patches

and sample across different landscapes or region, we could specify a multilevel regression as follows:

$$y_{i,p,l} = \alpha + \beta_1 x_i + \gamma_p + \delta_l + \varepsilon_i, \quad (6.9)$$

where γ_p is a random effect of a patch, and δ_l is a random effect of the landscape or region. In doing so, this formulation acknowledges that observations within each region have some correlation/similarity.

6.2.4.6 Generalized Least Squares and Spatial Mixed Models

Generalized least squares models (GLS) and spatial mixed models are similar in scope to a multilevel model. The main conceptual difference is that we specify spatial correlation structures explicitly in the random effects (GLMMs) or residuals (GLS) by modeling the variance–covariance matrix over space.

In a GLS spatial model, we take a typical regression, $y_i = \alpha + \beta_1 x_i + \varepsilon_i$, where ε is $\sim N(0, \sigma^2)$ and replace the variance on the error term with a variance–covariance matrix: $\varepsilon \sim N(0, \Sigma)$ (Keitt et al. 2002). In a GLMM spatial model, a similar approach is taken, but a variance–covariance matrix is added for the random effect: $\gamma \sim N(0, \Sigma)$ rather than the residuals (Littell et al. 2006). In both cases, parametric correlation functions are fit to explain the variance–covariance matrix by specifying model-based correlation structures, akin to model-based variogram structures we described in Chap. 5. These correlation structures are sometimes referred to as Gaussian random fields (Thorson and Minto 2015). For example, in the GLS we will consider below, we will fit a spatial exponential covariance (see Chap. 5):

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \exp\left(-\frac{d_{ij}}{\alpha}\right) \\ \exp\left(-\frac{d_{ij}}{\alpha}\right) & 1 \end{bmatrix}, \quad (6.10)$$

where σ^2 is the non-spatial variance estimated, d_{ij} is the distance between two observations i and j , and α is a parameter to be estimated (related to the range). With mixed effects, one can specify models that only account for spatial autocorrelation within the regions/groups specified by the random effect. For instance, Fletcher (2005) used this general approach to account for within-patch spatial dependence of species distribution while assuming that among-patch dependence was negligible. Similar to CAR and SAR, GLS has a strong foundation for normally distributed response variables, but the application of these models to non-normal data is more challenging (Rousset and Ferdy 2014). Note the utility of GLS may depend upon the scale of environmental relationships being considered. For instance, Diniz et al. (2003) found that GLS tended to de-emphasize covariates

operating across large spatial scales while overemphasizing covariates operating at more local scales.

6.2.5 *Inference Versus Prediction*

An implicit but pervasive issue regarding spatial regression and other modeling approaches considered in this book regards whether the goal of the work is for inference or prediction. When our goal is inference, we are interested in estimating factors influencing response variables (Stephens et al. 2007). In contrast, if our goal is prediction, we aim to build models that can make accurate predictions or projections across space and time (Boyce et al. 2002), including both interpolating between sample locations and predicting to new areas (i.e., model transferability; see Chap. 7 for more). Ecologists and conservation biologists often use models in both ways, but ultimately these are very different goals and model building and evaluation will be (or should be) different depending on the goal.

Spatial regression models can be helpful in problems of inference, where we are interested in understanding spatial or environmental relationships, such as factors related to species distribution and abundance. These approaches can potentially provide more reliable inference in regard to parameter estimates and their uncertainty, as well as more reliable statistical hypothesis tests. However, the use of these models for prediction, projection, or interpolation can sometimes be difficult, depending on the type of model considered. For example, with autocovariate models, prediction requires information on the response variable (e.g., occurrence) across the region being predicted, because the regression model includes this information in the form of the autocovariate (Augustin et al. 1996). In contrast, trend surface and related spatial filtering models are straightforward to use in prediction because only the physical locations are used as predictors. In some cases, spatial regression models are used for prediction where the dependence term is ignored (e.g., using only the fixed effects from a mixed effects model). Depending on the goal of spatial modeling, the utility of the above approaches may vary.

6.3 Examples in R

6.3.1 *Packages in R*

In R, there are a few libraries that can be used for spatial regression models. We use the `mgcv` package for fitting generalized additive models (Wood 2006), `lme4` for fitting multilevel models (Bates et al. 2015), and `vegan` (Oksanen et al. 2018) and `spdep` (Bivand and Piras 2015) for fitting eigenvector maps. We use the `spdep` package for models requiring lattice data (autocovariate, SAR, CAR) and interpreting autocorrelation in the residuals of models. We fit spatial GLS and mixed models with

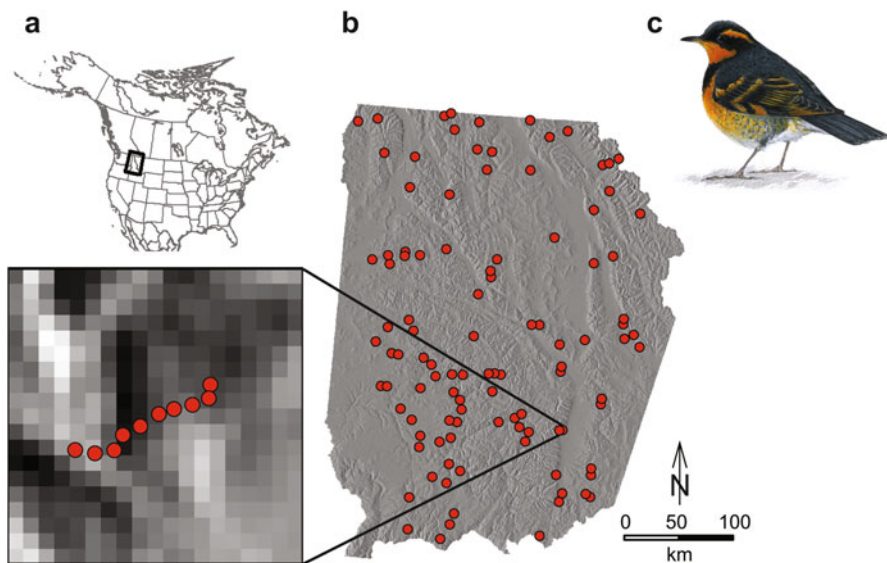


Fig. 6.6 The Northern Region Landbird Monitoring Program applies a hierarchical sampling design for surveying bird communities. This monitoring program covered (a) northern Idaho and western Montana, where (b) transects were distributed across different watersheds, with typically 10 points per transect. Here we focus on the occurrence of (c) varied thrush (picture courtesy of Matthew Dodder at <http://www.birdguy.net/>)

MASS (Venables and Ripley 2002) and `spaMM` (Rousset and Ferdy 2014), but other packages can be used, particularly Bayesian modeling packages (e.g., `spBayes`) (Finley et al. 2015).

6.3.2 The Data

Monitoring programs are often hierarchically structured and filled with spatio-temporal dependence in the data. The Northern Region Landbird Monitoring Program is one such example (Hutto and Young 2002). Sampling locations consisted of point counts (100-m radius), along a transect (typically 10 points/transect; transects are approximately 3 km long), with transects randomly selected within USFS Forest Regions across Montana and Idaho (Fig. 6.6). Ten-minute point counts were conducted by trained observers, where all birds seen or heard were recorded. Here we only consider birds detected within 100-m of the point. These points were also resampled over time, although we will not consider these temporal repeated measures here.

To interpret spatial regression models, we consider a simple environmental relationship of species occurrence along an elevation gradient. Elevation is

frequently considered to be an important, though often indirect, factor correlated with species distribution. We focus on the occurrence of the varied thrush (*Ixoreus naevius*) (Fig. 6.6), a migratory bird that breeds in the western USA. Varied thrush have declined in the western USA over the past several decades, based on Breeding Bird Survey data (Sauer et al. 2017), with the annual decline of approximately 2–3% per year (1966–2015: -2.47 , 95% CI: $-3.19, -1.79$; 2005–2015: -3.32 , 95% CI: $-5.14, -1.56$). Furthermore, they are often considered an old-growth, interior species (Brand and George 2001; Betts et al. 2018). Consequently, this species has been of some interest for conservation.

We fit logistic regression models and their spatial extensions to infer and predict the distribution of varied thrush as a function of elevation in this mountainous region. Here we focus on modeling detection/non-detection of thrushes (0/1 data). Elevation was derived from a 30-m resolution Digital Elevation Model (DEM). Prior to analysis, all GIS layers were aggregated to a common 200-m resolution, reflecting the grain of the sampling unit (100-m-radius point counts).

With this sampling design, there are likely observation errors in detecting varied thrushes, such that models that explicitly account for imperfect detection would be useful (McCarthy et al. 2012). Rota et al. (2011) estimated that detection probabilities of varied thrushes with this sampling design was relatively high ($p = 0.87$ /count), which is likely driven by their distinctive and loud song. We do not consider that sampling error here to focus specifically on the problem of spatial dependence. See sect. 6.4 for further discussion on sampling errors.

6.3.3 Models that Ignore Spatial Dependence

To begin, we import a raster layer of elevation with the `raster` package and use this layer to also derive other key variables related to elevation, such as slope and aspect (Fig. 6.7).

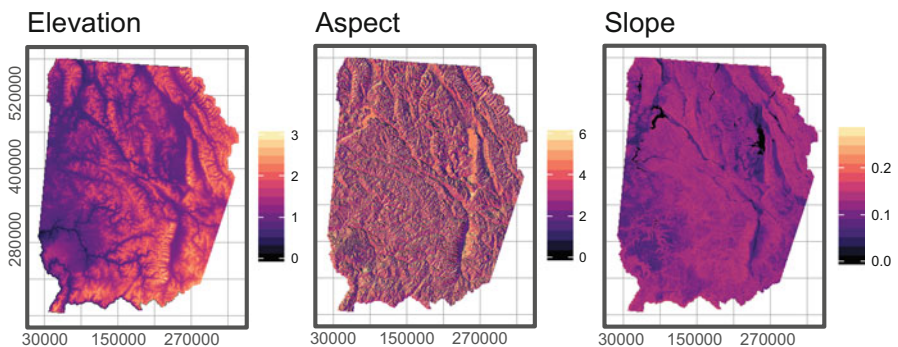


Fig. 6.7 The raster data considered come from a digital elevation model, including elevation (in km), aspect (in radians), and slope. Note that slope is double square-root transformed for visualization

Table 6.2 Terrain metrics that the `raster` package can calculate based on elevation data

Metric	Description
Aspect ^a	The compass direction that a slope faces.
Flow direction	The direction of the greatest drop in elevation (or smallest rise) from focal cell and its eight neighboring cells, coded as integer values in powers of 2 (starting east of focal cell and moving in a clockwise direction; 1, 2, 4, 8, 16, 32, 64, and 128).
Roughness	Difference between the maximum and minimum value of the focal cell and its eight neighboring cells.
Slope	The change in elevation, described as the difference in elevation between two points divided by the distance.
Terrain ruggedness index	Absolute difference between the value of a cell and the value of its eight neighboring cells (eight-neighbor rule).
Topographic position index	Difference between the value of a cell and the mean value of its eight neighboring cells (eight-neighbor rule).

^aCan be measured in degrees or radians

```
> library(raster)
> elev <- raster("elev.gri")

#create aspect and slope layers from the elevation layer
> elev.terr <- terrain(elev, opt = c("slope", "aspect"), unit =
  "radians")
```

The `terrain` function in the `raster` package takes an elevation layer (e.g., DEM) and returns raster layers that are calculated from elevation, including slope, aspect, topographic position index, terrain ruggedness index (TRI), roughness, and flow direction (Wilson et al. 2007) (Table 6.2). Here, we just calculate slope and aspect (Fig. 6.7). Note that for this function, the projection must be set on the raster layer for implementation. This function defaults to aspect being calculated in radians, using the algorithm in Horn (1981).

We can merge the slope and aspect layers into a single raster stack that holds all raster layers. We create a single object that holds all of the data with the `stack` function:

```
#makes a multilayered file for extraction
> layers <- stack(elev, elev.terr)
> names(layers) <- c("elev", "slope", "aspect")
```

We first consider a non-spatial logistic regression model. To do so, we use the `extract` function to grab covariate values from `layers` at the sample locations from the survey data and we then merge the covariates with our data on thrush occurrence using `cbind`.

```
> point.data <- read.csv("vath_2004.csv", header=T)
> coords <- cbind(point.data$EASTING, point.data$NORTHING)
> land.cov <- extract(x = layers, y = coords)
> point.data <- cbind(point.data, land.cov)
```

We consider a simple set of logistic regression models. We expect that elevation may help explain varied thrush occurrence, where thrushes may be most likely to occur at either low or moderate elevations. Consequently, we consider quadratic terms in the logistic regression model to account for potential non-linearities (Fig. 6.4) in occurrence as a function of elevation. We also consider slope and aspect as proxies for local variation in environmental conditions. First, we transform the explanatory variables to a mean of 0 and a variance of 1 (sometimes referred to as a z -transformation or “centering and scaling”). Centering and scaling can help improve model convergence and facilitates comparing coefficients for different parameters.

```
> point.data$elevs <- scale(point.data$elev, center = T, scale = T)
> point.data$slopes <- scale(point.data$slope, center = T, scale = T)
> point.data$aspects <- scale(point.data$aspect, center = T, scale = T)
```

Note that the default for the `scale` function is to both center and scale, but we explicitly request this here to illustrate. Now we can fit logistic regression models of varying complexity.

```
> VATH.elev <- glm(VATH ~ elevs, family = "binomial", data =
  point.data)
> VATH.all <- glm(VATH ~ elevs + slopes + aspects, family =
  "binomial", data = point.data)
> VATH.elev2 <- glm(VATH ~ elev + I(elev^2), family = "binomial", data
  = point.data)
```

Note that to specify a quadratic term in R, we write `I(elev^2)`. This could also be accomplished through the `poly()` (see below). We can contrast model fit using AIC:

```
> round(AIC(VATH.elev, VATH.all, VATH.elev2), 2)

##
df AIC
VATH.elev 2 583.10
VATH.all 4 584.84
VATH.elev2 3 566.54

> summary(VATH.elev2)

##
Call:
glm(formula = VATH ~ elev + I(elev^2), family = "binomial", data = point.
data)

Deviance Residuals:
  Min 1Q Median 3Q Max
-0.6088 -0.5787 -0.5032 -0.3231  3.0804
```

Coefficients:

```
Estimate Std. Error z value Pr(>|z|)
(Intercept) -7.984 1.990 -4.012 6.01e-05 ***
elev 10.698 3.227 3.316 0.000915 ***
I(elev^2) -4.476 1.281 -3.494 0.000475 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(Dispersion parameter for binomial family taken to be 1)

```
Null deviance: 584.34 on 804 degrees of freedom
Residual deviance: 560.54 on 802 degrees of freedom
AIC: 566.54
```

Number of Fisher Scoring iterations: 6

For each of these models, we can use the `summary` function to view the coefficients estimated from the model and related diagnostics. While this small set of candidate models is far from a complete set, from this comparison there is some evidence of thrush occurrence increasing at moderate elevations. This can be concluded because the linear elevation term is positive while the quadratic term is negative (both of which are significant based on p -values, or $\Pr(>|z|)$), which will cause a humped-shaped relationship with elevation. We can plot this relationship by first generating a new data set to predict onto and then use the `predict` function (Fig. 6.8):

```
> elev <- seq(min(point.data$elev), max(point.data$elev), length = 15)
> newdata <- data.frame(elev = elev)

> glm.pred <- predict(VATH.elev2, newdata = newdata, type = "link", se = T)
> ucl <- glm.pred$fit + 1.96*glm.pred$se.fit
> lcl <- glm.pred$fit - 1.96*glm.pred$se.fit
```

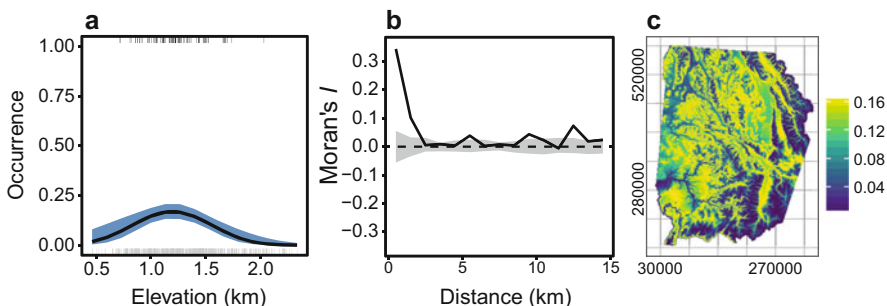


Fig. 6.8 (a) Predicted relationship (with 95% prediction intervals) of varied thrush occurrence with elevation based on a standard logistic regression model. (b) Correlogram using the raw response data, where gray region shows the 99% null envelope from a permutation test. (c) Mapping predictions from model

```
#create data frame and back-transform to probability scale
> glm.newdata <- data.frame(newdata, pred = plogis(glm.pred$fit),
  lcl = plogis(lcl), ucl = plogis(ucl))

> plot(glm.newdata$elev, glm.newdata$pred, ylim = c(0, 0.5))
> lines(glm.newdata$elev, glm.newdata$lcl)
> lines(glm.newdata$elev, glm.newdata$ucl)
```

We can also plot predictions of this model across the study region by predicting onto the raster stack layer. To do so, the `raster` package defaults to making predictions on the link scale, but we can then back-transform the predictions on the raster to the probability scale (Fig. 6.8c).

```
> glm.raster <- predict(model = VATH.elev2, object = layers)
> glm.raster <- exp(glm.raster) / (1 + exp(glm.raster))
> plot(glm.raster, xlab = "Longitude", ylab = "Latitude")
```

In this model and subsequent models, we will generally focus on two issues. First, is there evidence for spatial autocorrelation in the residuals of the models? Second, how do estimated relationships, that is, the coefficients and standard errors (SEs), change depending on the model?

We can determine if spatial dependence might be problematic for inferences by considering if there is evidence for spatial dependence in the residuals of the model (Dormann et al. 2007; Beale et al. 2010). First, we consider if there is spatial autocorrelation in the response variable. For interpreting spatial autocorrelation, we will use the correlogram function described in Chap. 5 when we used the `spdep` package. The function in Chap. 5 was altered to allow specification of different bins for lag distances and the maximum distance considered. This function is useful because it can be readily used for both binary data (0/1 response data) and for other response variable distributions (e.g., residuals), although other functions, such as the `correlog` function in the `ncf` package (Bjørnstad and Falck 2001) could also do the trick. We call this function `icorrelogram` and add it to our data frame with the `source` function. We then plot the resulting the correlogram (Fig. 6.8b).

```
#import function
> source('icorrelogram.r')
```

To inspect this function, simply type:

```
> icorrelogram

##
function(locations, z, binsize, maxdist) {

distbin <- seq(0, maxdist, by = binsize)
Nbin <- length(distbin) - 1
```

```

moran.results <- data.frame(dist = rep(NA,Nbin) ,
Morans.i =NA,null.lcl=NA, "null.ucl"=NA)

for (i in 1:Nbin) {
  d.start <- distbin[i]
  d.end <- distbin[i+1]
  neigh <- dnearneigh(x=locations, d1=d.start, d.end, longlat=F)
  wts <- nb2listw(neighbours=neigh, style='B', zero.policy=T)
  mor.i <- moran.mc(x=z, listw=wts, nsim=200, alternative="greater",
                    zero.policy=T)

  moran.results[i, "dist"] <- (d.end+d.start)/2
  moran.results[i, "Morans.i"] <- mor.i$statistic
  moran.results[i, "null.lcl"] <- quantile(mor.i$res, probs = 0.025, na.rm = T)
  moran.results[i, "null.ucl"] <- quantile(mor.i$res, probs = 0.975, na.rm = T)
}
return(moran.results)
}

```

This function identifies neighbors between points using the `dnearneigh` function for different distance classes. It then takes the object created, reformats it to a list of relevance to the **W** spatial neighbor matrix, and uses a `moran.mc` function to run a permutation-based Moran's *I*. The distance classes, Moran's *I* and the null envelope from the permutations are then stored in a data frame. We can run the function on the observed data and plot (Fig. 6.8):

```

#run correlogram function
> VATH.cor <- icorrelogram(locations = coords, z
  = point.data$VATH, binsize = 1000, maxdist = 15000)
> head(VATH.cor)

##
Dist Morans.i Null.lcl Null.ucl
1 500 0.34 -0.06 0.06
2 1500 0.10 -0.03 0.03
3 2500 0.01 -0.02 0.03

#plot correlogram
> plot(VATH.cor$Dist, VATH.cor$Morans.i, ylim = c(-0.5, 0.5))
> abline(h=0, lty = "dashed")
> lines(VATH.cor$Dist, VATH.cor$Null.lcl)
> lines(VATH.cor$Dist, VATH.cor$Null.ucl)

```

Now we consider if there is spatial autocorrelation in the residuals of the logistic regression model.

```

#residuals from quadratic elevation model
> VATH.elev2.res <- residuals(VATH.elev2, type = "deviance")

```

Note that we request the deviance-based residuals. For GLM-type models, there are several related residuals that could be calculated, the default being a deviance-based residual. For a binomial or Bernoulli GLM, this type of residual is calculated as:

$$d_i = s_i \sqrt{-2(y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i))}, \quad (6.11)$$

where d_i is the deviance of observation i , y_i is the observation, \hat{y} is the predicted value, and $s_i = 1$ if $y_i = 1$ and -1 if $y_i = 0$. The deviance residuals are potentially more useful in GLMs in comparison to others because they are directly related to the overall deviance (and likelihood) of the model, where the sum of the deviance residuals equals the deviance of the model ($-2\log$ -likelihood). We can visualize spatially the residuals by mapping them. More formally, we can assess this using the `icorrelogram` function:

```
#correlogram on residuals
> corr.res <- icorrelogram(locations = coords, z =
  VATH.elev2.res, binsize = 1000, maxdist = 15000)
```

Here, we find evidence for spatial autocorrelation in the residuals of the model (Fig. 6.9). Note that rather than using correlograms, we could have used semivariograms on the residuals to interpret spatial autocorrelation in the residuals (Beguin et al. 2012).

It is important to understand the interpretation of the use of residuals in this analysis in comparison to the raw data. For instance, if we fit an intercept-only (mean) model and contrast correlograms from the raw data and the residuals of the mean model:

```
> VATH.int <- glm(VATH ~ 1, family = "binomial", data = point.data)
> VATH.int.res <- residuals(VATH.int, type = "deviance")

> corr.int.res <- icorrelogram(locations = coords, z =
  VATH.int.res, binsize = 1000, maxdist = 15000)
> cor(VATH.cor$Morans.i, corr.int.res$Morans.i)

##
[1] 1
```

We find that the Moran's I is identical ($r = 1$). This illustrates the equivalence of considering residuals from regression models in correlograms when no predictors are considered to that of the raw data (Bivand et al. 2013).

Because of the spatial dependence in the residuals, we consider either subsetting the data based on the approximate range of spatial autocorrelation or regression-like models that attempt to account for spatial autocorrelation. First, we subset the data.

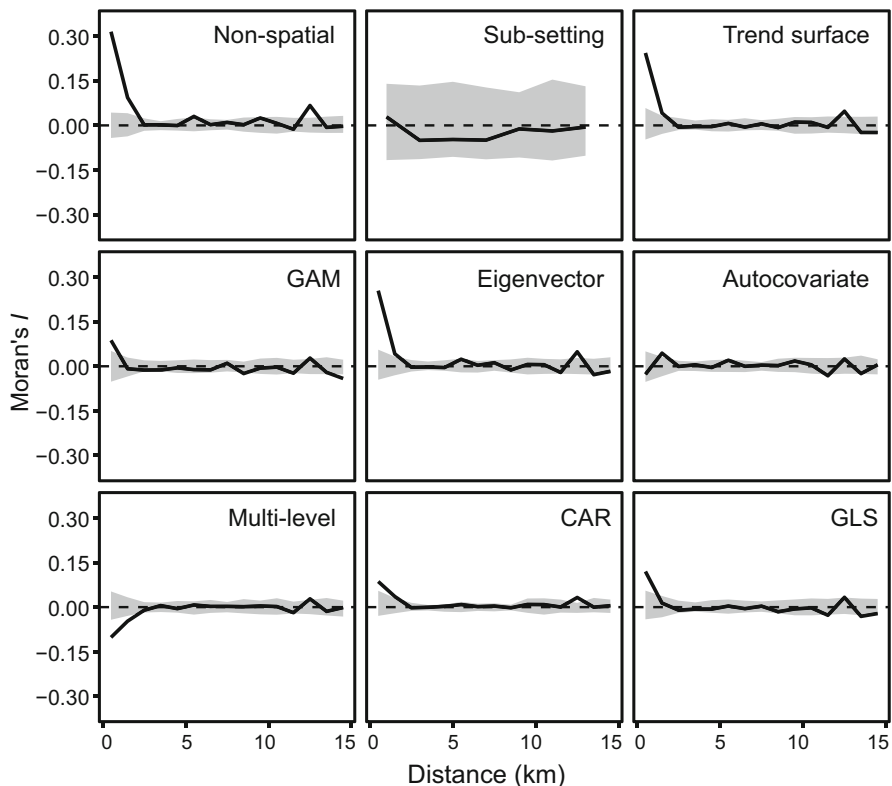


Fig. 6.9 Correlograms on the residuals of the models considered. Note that for subsetting the data, correlograms were calculated for wider lag distance bins because of less data being used

Given the sampling design and the shape of the correlogram (Fig. 6.8b), it would be natural to only consider one point per transect. Note we could also potentially pool across all points on each transect, however, such an approach would increase the spatial grain of the analysis, which might not be ideal. Below we use a function to pick one random point from each transect.

```
#shuffle points on transects
> rand.vector <- with(point.data, ave(POINT, as.factor(TRANSECT),
  FUN=function(x) sample(length(x))))

#pick one random point on transect and remove rest
> point.datasub <- point.data[rand.vector == 1,]

#coordinates from subset data
> coords.sub <- cbind(point.datasub$EASTING, point.datasub$NORTHING)
```


With this data subset, we then refit the logistic regression model.

```
> VATH.sub <- glm(VATH ~ elev + I(elev^2), family = "binomial", data =
  point.datasub)
> summary(VATH.sub)

##
Call:
glm(formula = VATH ~ elev + I(elev^2), family = "binomial", data = point.
  datasub)

Deviance Residuals:
Min 1Q Median 3Q Max
-0.5673 -0.5408 -0.4677 -0.3022 2.6507

Coefficients:
Estimate Std. Error z value Pr(>|z|)
(Intercept) -6.158 4.498 -1.369 0.171
elev 8.254 7.519 1.098 0.272
I(elev^2) -3.860 3.076 -1.255 0.209

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 109.89 on 166 degrees of freedom
Residual deviance: 105.18 on 164 degrees of freedom
AIC: 111.18

Number of Fisher Scoring iterations: 6
```

When we subset the data, our sample size decreases substantially, from 805 to 167 points. Not surprisingly, the SEs on the parameter estimates increase substantially and there is no longer strong evidence for an elevation effect. We can interpret whether this subsetting removed the spatial autocorrelations in the residuals of the model. Note that for this subset, we need to use a larger lag distance than 1-km because we no longer have data points <1 km (or alternatively, one could just increase the first few bin sizes). We calculate the correlogram using a 2-km lag distance.

```
> VATH.sub.res <- residuals(VATH.sub)

#correlogram on residuals
> corr.sub.res <- icorrelogram(locations = coords.sub, z =
  VATH.sub.res, binsize = 2000, maxdist = 15000)
```

This subsetting suggests that spatial autocorrelation is no longer problematic (Fig. 6.9), but there is a cost in terms of reduced power. Regression models that use all of the data but account for spatial dependence might be a useful alternative in this case.

6.3.4 Models that Account for Spatial Dependence

We consider several types of models that account for spatial dependence. These include: trend surface models, eigenvector-based models, autocovariate models (autologistic regression), autoregressive models (a CAR model), a multilevel model, generalized least squares, and spatial GLMMs.

6.3.4.1 Trend Surface Models

We consider two types of trend surface models. In the first model, we simply extend our logistic regression model to include x - y coordinates, along with their quadratic and cubic polynomial terms with the $I()$ function.

```
> VATH.trend <- glm(VATH ~ elev + I(elev^2) + EASTING + NORTHING +
  I(EASTING^2) + I(EASTING^3) + I(NORTHING^2) + I(NORTHING^3),
  family = "binomial", data = point.data)
> summary(VATH.trend)
```

```
##
```

```
Call:
```

```
glm(formula = VATH ~ elev + I(elev^2) + EASTING + NORTHING +
  I(EASTING^2) + I(EASTING^3) + I(NORTHING^2) + I(NORTHING^3),
  family = "binomial", data = point.data)
```

```
Deviance Residuals:
```

```
Min 1Q Median 3Q Max
```

```
-1.0301 -0.5425 -0.2959 -0.1743 2.8718
```

```
Coefficients:
```

```
Estimate Std. Error z value Pr(>|z|)
```

```
(Intercept) -9.861e+00 6.332e+00 -1.557 0.11943
```

```
elev 8.795e+00 3.138e+00 2.803 0.00507 **
```

```
I(elev^2) -3.195e+00 1.248e+00 -2.559 0.01049 *
```

```
EASTING 2.208e-04 4.769e-05 4.631 3.65e-06 ***
```

```
NORTHING -8.018e-05 5.076e-05 -1.580 0.11420
```

```
I(EASTING^2) -1.263e-09 2.806e-10 -4.502 6.72e-06 ***
```

```
I(EASTING^3) 2.090e-15 5.122e-16 4.081 4.48e-05 ***
```

```
I(NORTHING^2) 2.296e-10 1.366e-10 1.681 0.09275 .
```

```
I(NORTHING^3) -2.049e-16 1.179e-16 -1.738 0.08216 .
```

```
---
```

```
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
(Dispersion parameter for binomial family taken to be 1)
```

```
Null deviance: 584.34 on 804 degrees of freedom
```

```
Residual deviance: 486.89 on 796 degrees of freedom
```

```
AIC: 504.89
```

```
Number of Fisher Scoring iterations: 6
```

In the above model, we manually added quadratic and cubic terms. A more automated way to do this is with the `poly` function, where specifying `poly(EASTING, 3)` would add the linear, quadratic and cubic terms. Note that the `poly` function also standardizes polynomials to be orthogonal, removing the correlation between terms (which in many situations would be preferred). While the above model is straightforward to implement, it is limited in the spatial variation it can capture. An alternative to this model is to consider a generalized additive model (GAM), where we allow spline functions to capture spatial variation. The `mgcv` package provides a means to automate the selection of spline variation through the use of generalized cross-validation procedures. This model can be run as:

```
> library(mgcv)
> VATH.gam <- gam(VATH ~ elev + I(elev^2) + s(EASTING, NORTHING),
  family = "binomial", data = point.data)
```

In this model formulation, elevation is considered in a similar way as above but splines are considered for both Easting (x) and Northing (y) coordinates with the `s` command. This syntax defaults to automated selection of the number of knots being considered. We can manually adjust the number of knots (Fig. 6.5) by adding some syntax to the `s` command. We will look at GAMs in more detail in Chap. 7. In this case, the use of the `gam` formulation reduces spatial autocorrelation in the residuals (Fig. 6.9); however, it does not appear to fully remove the spatial dependence.

6.3.4.2 Eigenvector Mapping

To account for spatial dependence with eigenvector mapping, there are three steps. First, we create a neighborhood weights matrix with the `spdep` package. We can do this in several ways. Here we calculate a neighborhood weights matrix by using the maximum distance needed for a minimum spanning tree—the minimum set of connections needed to fully connect points across the landscape. The distance needed for a minimum spanning tree can be determined with the `vegan` package using the `spantree` function (note: this distance could also be determined using the `pcnm` function and finding the threshold, as discussed in Chap. 5).

```
> library(vegan)
> spantree.em <- spantree(dist(coords), toolong = 0)
> max(spantree.em$dist)

##
[1] 41351.09
```

We then identify neighborhoods with the `dnearneigh` function using this distance. With these neighbors, we extract the distances between neighbors with the `nbdists` function. Finally, we transform distances as suggested in Dormann et al. (2007) with the `lapply` function (because `nbdists` object is in list form),

and then create a list in the format relevant to the **W** matrix with the `nb2listw` function:

```
> dnn <- dnearneigh(coords, 0, max(spantree.em$dist))
> dnn_dists <- nbdists(dnn, coords)
> dnn_sims <- lapply(dnn_dists, function(x) (1 - ((x / 4)^2)))
> ME.weight <- nb2listw(dnn, glist = dnn_sims, style = "B",
  zero.policy = T)
```

With this **W** matrix, we use the `ME` function in the `spdep` package to identify the most important eigenvectors that reduce spatial dependence, based on a permutation bootstrap test on Moran's *I* for the residuals (Griffith and Peres-Neto 2006). In this function, we include the relevant covariates in the model formula, but we also add the neighborhood matrix (in list form):

```
> VATH.ME <- ME(VATH ~ elev + I(elev^2), family = "binomial", listw =
  ME.weight, data = point.data)
> VATH.ME$selection
```

```
##
Eigenvector ZI pr (ZI)
0 NA NA 0.01
1 796 NA 0.01
2 804 NA 0.03
3 805 NA 0.20
```

```
> head(fitted(VATH.ME), 2)
```

```
##
vec796 vec804 vec805
[1,] 0.003042641 -0.008629250 0.01187249
[2,] 0.003088077 -0.008737222 0.01196633
```

The `ME` function provides output regarding the eigenvectors selected but we need to then refit the logistic regression model with this eigenvectors included as covariates.

```
#new glm with ME covariates
> VATH.evm <- glm(VATH ~ elev + I(elev^2) + fitted(VATH.ME), family =
  "binomial", data = point.data)
> summary(VATH.evm)
```

```
##
Call:
glm(formula = VATH ~ elev + I(elev^2) + fitted(VATH.ME), family =
"binomial",
data = point.data)
```

```

Deviance Residuals:
Min 1Q Median 3Q Max
-1.5359 -0.5175 -0.4027 -0.1416 2.7454

Coefficients:
Estimate Std. Error z value Pr(>|z|)
(Intercept) -8.401 1.948 -4.312 1.62e-05 ***
elev 8.776 3.029 2.898 0.00376 **
I(elev^2) -3.112 1.168 -2.664 0.00773 **
fitted(VATH.ME)vec796 -14.742 3.198 -4.610 4.03e-06 ***
fitted(VATH.ME)vec804 -8.644 3.242 -2.666 0.00767 **
fitted(VATH.ME)vec805 38.110 8.789 4.336 1.45e-05 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 584.34 on 804 degrees of freedom
Residual deviance: 499.73 on 799 degrees of freedom
AIC: 511.73

Number of Fisher Scoring iterations: 7

```

In this case, the approach identifies three eigenvectors to include, each of which explains occurrence to some degree. However, the inclusion of these eigenvectors does not remove spatial autocorrelation in the residuals of the model (Fig. 6.9). Overall, the main difference in this approach relative to the trend surface model described above is the creation of the eigenvector covariates and determining which of these covariates to include in the final logistic regression model.

6.3.4.3 Autocovariate Models

To fit autocovariate models, we calculate new autocovariates and then use these covariates in a standard logistic regression model. We will calculate these autocovariates with the `autocov_dist` function in the `spdep` package. Because most of the significant autocorrelation in the residuals occurs < 1 km (Fig. 6.8b), we will calculate the autocovariates at this scale.

```
> auto1km <- autocov_dist(point.data$VATH, coords, nbs = 1000, type =
  "one", style = "B", zero.policy = T)
```

The `type=` provides the weighting scheme. When `inverse` is specified, points are weighted by the inverse of the distance between the focal point and the neighboring point. If `"one"` is specified, all points within the distance (`nbs`) are given equal weight. `style` describes how the covariate will be calculated, with `"B"` reflecting a binary coding. Bardos et al. (2015) stated that using `style = "B"` provides a valid weighting scheme for autocovariate models.

We then fit standard logistic regression models with these covariates included.

```
> VATH.autolkm <- glm(VATH ~ elev + I(elev^2) + autolkm, family =
  "binomial", data = point.data)

> summary(VATH.autolkm)

##
Call:
glm(formula = VATH ~ elev + I(elev^2) + autolkm, family = "binomial",
    data = point.data)

Deviance Residuals:
    Min 1Q Median 3Q Max
-2.0314 -0.4131 -0.3809 -0.2902  2.9077

Coefficients:
Estimate Std. Error z value Pr(>|z|)
(Intercept) -6.6518  1.9660 -3.383 0.000716 ***
elev  6.9046  3.1222  2.211 0.027006 *
I(elev^2) -2.8061  1.2106 -2.318 0.020450 *
autolkm 0.8665  0.1008  8.596 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 584.34 on 804 degrees of freedom
Residual deviance: 470.99 on 801 degrees of freedom
AIC: 478.99

Number of Fisher Scoring iterations: 6
```

In this case, the inclusion of the autocovariate in the model is very significant, while the coefficients for the elevation effect decrease. In addition, the inclusion of the autocovariate removes the spatial autocorrelation in the residuals (Fig. 6.9).

6.3.4.4 Autoregressive Models

Fitting autoregressive models to non-normal data is challenging. One approach is to use Bayesian modeling. While there are some packages for fitting spatial autoregressive models with Bayesian modeling (e.g., see the `spBayes` package; Finley et al. 2015), using Bayesian methods for spatial regression is often computationally demanding. A new alternative is using “Integrated Nested Laplace Approximation” or INLA (Blangiardo and Cameletti 2015). The value of this approach is that it greatly reduces the computational demands of Bayesian modeling. However, it does only apply to certain types of Bayesian models. For example,

INLA can be used to fit CAR models for binary data. To do so, we need to create a neighborhood weights matrix of the "dgTMatrix" form, which is a type of a sparse matrix (sparse matrices are those that have very few observations and are largely filled with zeros. There are efficient ways to store and manipulate these types of matrices in R). We first create a neighborhood matrix by creating Thiessen polygons from the point data with the `deldir` and `dismo` packages (Fig. 6.3). Thiessen polygons, also known as Voronoi polygons, are based on Delaunay triangulation. These polygons partition a region into convex polygons such that each polygon contains exactly one point.

```
> library(INLA)
> library(deldir)
> library(dismo)

> thiessen <- voronoi(coords)

#plot thiessen polygons
> plot(thiessen)
> points(coords, col = "red")

> point.poly <- poly2nb(thiessen)

#plot neighborhood matrix
> plot(point.poly, coords, col = "red", add = T)

#format neighborhood matrix
> adj <- nb2mat(point.poly, style = "B")
> adj <- as(adj, "dgTMatrix")
```

With this neighborhood matrix, we can then fit the CAR model. To do so, for INLA we need to first specify the type of the model fitting, including the covariates being considered. For the CAR model, we add an observation-level covariate to the data frame (`id`) and then specify "besag" for the CAR model. We then fit the model with the `inla` function:

```
> point.data$id <- 1:nrow(point.data)
> VATH.inla <- inla(VATH ~ elev + I(elev^2) + f(id, model = "besag",
  graph = adj), family = "binomial", data = point.data,
  control.predictor = list(compute = TRUE))

> summary(VATH.inla)

##
Call:
c("inla(formula = form, family = \"binomial\", data = point.data, ", "
control.predictor = list(compute = TRUE))")
```

Time used:

```
Pre-processing Running inla Post-processing Total
2.8085 3.2775 0.5343 6.6203
```

Fixed effects:

```
mean sd 0.025quant 0.5quant 0.975quant mode kld
(Intercept) -8.0537 1.9683 -12.1721 -7.9648 -4.4301 -7.7824 0
elev 10.8093 3.1908 4.9396 10.6618 17.4881 10.3574 0
I(elev^2) -4.5148 1.2666 -7.1800 -4.4519 -2.1971 -4.3222 0
```

Random effects:

```
Name Model
ID Besags ICAR model
```

Model hyperparameters:

```
mean sd 0.025quant 0.5quant 0.975quant mode
Precision for ID 18537.90 18336.86 1248.75 13131.81 66833.34 3386.31
```

Expected number of effective parameters (std dev) : 2.993 (0.0029)

Number of equivalent replicates : 268.99

Marginal log-Likelihood: -899.24

Posterior marginals for linear predictor and fitted values computed

This approach allows for a binomial CAR model (note that if our response data were normally distributed, we could use the `spautolm` function in the `spdep` package). The `spaMM` package can also fit CAR models to binomial data, but the above model in that package takes $>50\times$ longer to fit than with INLA. Also, the INLA approach is computationally much faster than using other Bayesian modeling approaches, which is a major benefit of this package. With the `inla` package, we must manually calculate residuals to interpret spatial autocorrelation:

```
#manual deviance residual calculation:
> VATH.inla.fit <- VATH.inla$summary.fitted.values$mean
> si <- ifelse(point.data$VATH==1, 1, -1)
> VATH.inla.res <- si * (-2 * (point.data$VATH * log(VATH.inla.fit) + (1 -
  point.data$VATH) * log(1 - VATH.inla.fit)))^0.5

#correlogram on residuals
> cor.inla.res <- icorrelogram(locations = coords, z = VATH.inla.res,
  binsize = 1000, maxdist = 15000)
```

In this case, we find that the CAR model removes most, but not all, of the autocorrelation in the residuals (Fig. 6.9). This may be due to the fact that the CAR model is only using first-order neighbors, such that only dependence between neighboring points (~ 300 m apart) is captured. As the observed spatial dependence in the residuals extends out to 1–2 km (Fig. 6.8b), this smaller scale is not sufficient in this case.

6.3.4.5 Multilevel Models

A simple multilevel model can also be fit to these data by considering transects as a random effect in the regression model. In doing so, we effectively “block” with transects, treating points within transects as having potential spatial dependence (Keitt et al. 2002). Because this structure is not spatially explicit, we effectively assume that dependence is constant within transects (e.g., neighboring points have the same dependence as points located along the ends of the transects). These models can be fit using the `lme4` package. Prior to the model fitting, we need to make sure that transect is considered a factor. Then we can fit the model with `glmer` function.

```
> library(lme4)

#random effects should be a factor
> str(point.data)
> point.data$TRANSECT <- as.factor(point.data$TRANSECT)

#glmm using lme4
> VATH.glmm <- glmer(VATH ~ elev + I(elev^2) + (1|TRANSECT), family =
  "binomial", data = point.data)

> summary(VATH.glmm)

##
Generalized linear mixed model fit by maximum likelihood (Laplace
Approximation) ['glmerMod']
Family: binomial (logit)
Formula: VATH ~ elev + I(elev^2) + (1 | TRANSECT)
Data: point.data

AIC BIC logLik deviance df.resid
498.4 517.2 -245.2 490.4 801

Scaled residuals:
Min 1Q Median 3Q Max
-1.3520 -0.1755 -0.1541 -0.1129 5.7688

Random effects:
Groups Name Variance Std.Dev.
TRANSECT (Intercept) 4.456 2.111
Number of obs: 805, groups: TRANSECT, 167

Fixed effects:
Estimate Std. Error z value Pr(>|z|)
(Intercept) -8.470 3.262 -2.596 0.00942 **
elev 9.459 5.155 1.835 0.06653 .
I(elev^2) -4.043 1.979 -2.043 0.04106 *
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

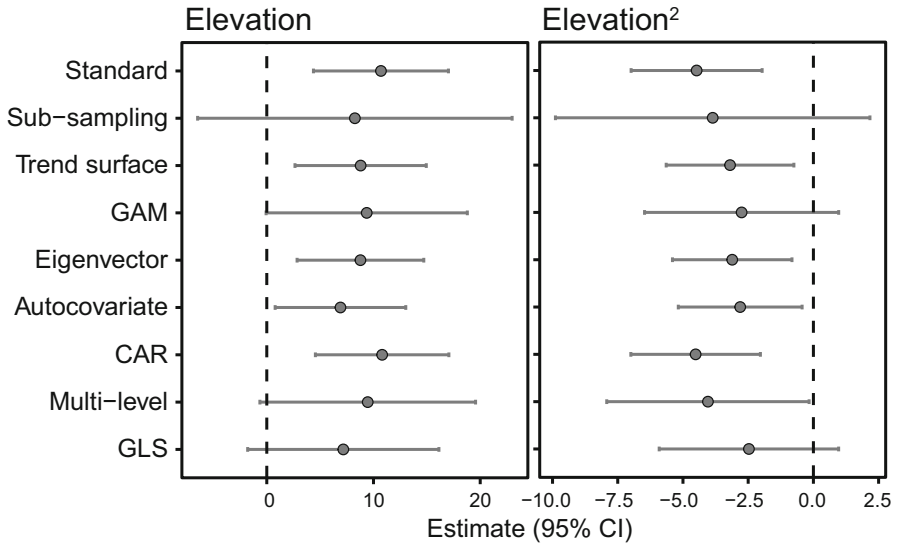


Fig. 6.10 Estimates of elevation relationships based on the spatial models considered

```
Correlation of Fixed Effects:
(Intr) elev
elev -0.981
I(elev^2) 0.946 -0.988
```

When fitting random effects, we specify (1|TRANSECT), which signals that our transect is being considered as a random intercept. We will see more uses of random effects and their specification in Chap. 11. In this case, we find that by adding a random transect effect to the model structure, positive spatial dependence in the residuals vanishes (Fig. 6.9), although now there is some slight negative autocorrelation in the residuals at short distances. Also, note that the SEs increase and that the elevation effect is only weakly significant (Fig. 6.10).

6.3.4.6 GLS and Mixed Models

GLS and spatially explicit mixed models are difficult to implement for non-normal response data. For normal response data, the `nlme` package can accommodate spatial correlation structures in the model residuals (sometimes referred to as “R-side” correlation structures) or in the random effects (sometimes referred to as “G-side” correlation structures) (Littell et al. 2006).

Given the hierarchical structure of the data, we can fit spatial mixed models where spatial correlation is calculated within transects or across the entire region. The `glmmPQL` function in `MASS` package can be used for GLS and spatial mixed models.

However, this approach uses penalized “quasi-likelihood” and has been shown to have poor properties (Rousset and Ferdy 2014). Because maximum likelihood is not used, we cannot use model selection approaches with this function. Nonetheless, we can still estimate environmental relationships that account for spatial dependence. Here, we fit an exponential correlation function within transects, by identifying transect as a random effect and using the `corExp` command:

```
> library(MASS)
> library(nlme)
> VATH.pql <- glmmPQL(VATH ~ elev + I(elev^2), random = ~1|TRANSECT,
  correlation = corExp(form = ~ EASTING +
    NORTHING), family = "binomial", data = point.data)
```

A similar model can be fit that considers spatial dependence throughout the region (not just within transects). This model takes considerable time to run, but we illustrate it as an example. To do so, we create an observation-level factor. This factor is then fit into the model as a random effect (Dormann et al. 2007).

```
> GROUP <- factor(rep("obs", nrow(point.data)))
> VATH.gls <- glmmPQL(VATH ~ elev + I(elev^2), random = ~1|GROUP,
  correlation = corExp(form = ~ EASTING + NORTHING),
  family = "binomial", data = point.data)
```

Penalized quasi-likelihood has some limitations, including potential bias in estimating random and fixed effects, as well as an inability to use model selection. A recent development that fits similar models without use of penalized quasi-likelihood may overcome some of these limitations (Rousset and Ferdy 2014). The `spaMM` package uses maximum likelihood to estimate a spatial GLMM with Laplace approximation. A similar formulation to that above can be fit with this package using the `corrHLfit` function:

```
> library(spaMM)
> VATH.spamm.ml <- corrHLfit(VATH ~ elev + I(elev^2) +
  Matern(1|EASTING+NORTHING),
  HLmethod = "ML", data = point.data, family = binomial(), ranFix =
  list(nu=0.5))
```

In this function, we specify a general Matérn spatial correlation structure. The negative exponential function used above is a specific form of a Matérn correlation structure, which in this case is called by setting $\nu = 0.5$ in the `ranFix` statement (see Chap. 5). Overall, this model provides similar estimates and results to the `gls` model in this situation. For both, spatial autocorrelation is not removed in the residuals when fitting the spatial correlation function across the entire region (Fig. 6.9). However, when only fitting the function within transects with the `pql` model, we find that spatial autocorrelation is removed in the residuals.

6.4 Next Steps and Advanced Issues

6.4.1 *General Bayesian Models for Spatial Dependence*

Proper accounting for spatial dependence in non-normal data can be difficult. In this chapter, we have focused on approaches that aim to address this issue in a variety of ways, but each of these approaches has some limitations. Bayesian models that capture spatial dependence provide a flexible means to accommodate spatial dependence. The INLA package provides one straightforward approach to do so, but INLA is limited to only certain types of regression models. More flexibility can be achieved through modeling spatial dependence using the bugs language (via Winbugs or Jags) (Kery and Royle 2016). In these approaches, either CAR/SAR types of models can be fit or GLS and mixed model-like formulations can be fit. This is often accomplished through the hierarchical formulation of spatial dependence coming from a multivariate normal distribution. These types of models are often thought to be useful for accounting for spatial dependence, although they can be challenging to fit (Beale et al. 2010).

6.4.2 *Detection Errors and Spatial Dependence*

Throughout this chapter, we have ignored the problem of sampling error, such as imperfect detection of species, to focus more simply on the issue of accounting for spatial autocorrelation. However, observation errors are common in data sets and these errors frequently need to be accounted for (MacKenzie et al. 2002). Several models exist for accounting for imperfect detection, both in terms of false positive and false negative errors (Miller et al. 2011). False negative errors are more common, where a species or individual occurs in an area but we fail to detect it. Occupancy, N -mixture, and distance sampling models are common approaches to account for these issues (Kery and Royle 2016). False positive errors occur when we misidentify species: we record that a species occurs in an area when in fact it does not. False positive errors are more difficult to account for, but some models exist that do so (Miller et al. 2011). We do not focus on these models in this book, largely because there have been several excellent books that illustrate these models, including their implementation in R (Royle and Dorazio 2008; Kery and Royle 2016).

There has been recent interest in extending these models to account for spatial autocorrelation (Hines et al. 2010; Johnson et al. 2013). Initial attempts used autocovariates like those shown here to account for spatial dependence (Royle and Dorazio 2008). More recently, geostatistical models have been developed as well (Johnson et al. 2013; Broms et al. 2014). Most of these models require customized code and implementation with Winbugs or Jags interfaced through the use of R (Carroll and Johnson 2008; Rota et al. 2011). However, some specialized R

packages can also accommodate spatial dependence in this context. The `hSDM` and `stocc` packages provide occupancy implementations that can accommodate spatial dependence (Johnson et al. 2013).

6.5 Conclusions

Tobler's first law of geography emphasizes that spatial dependence is common in nature. Given that ignoring this fact can lead to spurious inferences (Bivand 1980; Legendre 1993), accounting for spatial dependence in ecological data is often needed. Doing so, however, can be challenging. Here, we illustrate a variety of approaches to accounting for spatial dependence, contrasting their utility when using binary response data. In this case, trend surface and related environmental filtering (GAMs, eigenvector mapping) did not remove the spatial dependence in the residuals. CAR models also did not remove the spatial dependence, presumably because of the small neighborhood considered. Autocovariate and multilevel models did remove spatial dependence in the residuals by appropriately capturing the spatial scale of dependence in the data. Similar to Beale et al. (2010) and Dormann et al. (2007), we find that autocovariate models tended to shrink the effects of the environment relative to other approaches. In general, we recommend the use of mixed models and CAR models that can account for local spatial dependence and adjust uncertainty (SEs/CIs) of environmental relationships. This example illustrates that appropriately capturing the scale of spatial dependence in model structure is important for well-specified spatial models.

Throughout this discussion, we have used geographic coordinates and distances to make inferences about spatial dependence and adjust for this issue in understanding environmental relationships. Yet in many situations, effective distances that capture the complexity of the environment (e.g., shopping malls as barriers to organism movement and resource acquisition) may be more relevant. Spatial weights matrices can capture such complexities when warranted (Dray et al. 2006). Ver Hoef et al. (2018) also emphasized how spatial neighborhoods used in autoregressive models capture similar ideas to the use of network modeling in connectivity assessments (see Chap. 9). This is an interesting and important linkage that we expect will receive more attention in the coming years.

Care should be taken when applying spatial models, particularly for non-normal response variables. There is ongoing debate regarding the utility of different modeling approaches to account for spatial dependence (Dormann et al. 2007; Betts et al. 2009; Dormann 2009). In addition, while several lines of evidence suggest that spatial autocorrelation is problematic for conventional regression modeling, counter examples have also been emphasized (Diniz et al. 2003; Hawkins et al. 2007). Further advances in this area will no doubt provide a useful set of tools for spatial ecologists and conservation biologists alike.

References

- Algar AC, Kharouba HM, Young ER, Kerr JT (2009) Predicting the future of species diversity: macroecological theory, climate change, and direct tests of alternative forecasting methods. *Ecography* 32(1):22–33. <https://doi.org/10.1111/j.1600-0587.2009.05832.x>
- Augustin NH, Muggleston MA, Buckland ST (1996) An autologistic model for the spatial distribution of wildlife. *J Appl Ecol* 33(2):339–347
- Bardos DC, Guillera-Aroita G, Wintle BA (2015) Valid auto-models for spatially autocorrelated occupancy and abundance data. *Methods Ecol Evol* 6(10):1137–1149. <https://doi.org/10.1111/2041-210x.12402>
- Bates D, Machler M, Bolker BM, Walker SC (2015) Fitting linear mixed-effects models using lme4. *J Stat Softw* 67(1):1–48
- 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>
- Beguín J, Martino S, Rue H, Cumming SG (2012) Hierarchical analysis of spatially autocorrelated ecological data using integrated nested Laplace approximation. *Methods Ecol Evol* 3(5):921–929. <https://doi.org/10.1111/j.2041-210X.2012.00211.x>
- Betts MG, Diamond AW, Forbes GJ, Villard MA, Gunn JS (2006) The importance of spatial autocorrelation, extent and resolution in predicting forest bird occurrence. *Ecol Model* 191(2):197–224
- Betts MG, Ganio LM, Huso MMP, Som NA, Huetmann F, Bowman J, Wintle BA (2009) Comment on “Methods to account for spatial autocorrelation in the analysis of species distributional data: a review”. *Ecography* 32(3):374–378. <https://doi.org/10.1111/j.1600-0587.2008.05562.x>
- Betts MG, Phalan B, Frey SJK, Rousseau JS, Yang ZQ (2018) Old-growth forests buffer climate-sensitive bird populations from warming. *Divers Distrib* 24(4):439–447. <https://doi.org/10.1111/ddi.12688>
- Bini LM, Diniz JAF, Rangel T, Akre TSB, Albaladejo RG, Albuquerque FS, Aparicio A, Araujo MB, Baselga A, Beck J, Bellocq MI, Bohning-Gaese K, Borges PAV, Castro-Parga I, Chey VK, Chown SL, de Marco P, Dobkin DS, Ferrer-Castan D, Field R, Filloy J, Fleishman E, Gomez JF, Hortal J, Iverson JB, Kerr JT, Kissling WD, Kitching IJ, Leon-Cortes JL, Lobo JM, Montoya D, Morales-Castilla I, Moreno JC, Oberdorff T, Olalla-Tarraga MA, Pausas JG, Qian H, Rahbek C, Rodriguez MA, Rueda M, Ruggiero A, Sackmann P, Sanders NJ, Terribile LC, Vetaas OR, Hawkins BA (2009) Coefficient shifts in geographical ecology: an empirical evaluation of spatial and non-spatial regression. *Ecography* 32(2):193–204. <https://doi.org/10.1111/j.1600-0587.2009.05717.x>
- Bivand R (1980) A Monte Carlo study of correlation coefficient estimation with spatially autocorrelated observations. *Quaestiones Geographicae* 6:5–10
- 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>
- Blangiardo M, Cameletti M (2015) Spatial and spatio-temporal Bayesian models with R-INLA. Wiley, Chichester
- Bolker B (2008) Ecological models and data in R. Princeton University Press, Princeton, NJ
- Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH, White JSS (2009) Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol Evol* 24(3):127–135. <https://doi.org/10.1016/j.tree.2008.10.008>

- Boyce MS, Vernier PR, Nielsen SE, Schmiegelow FKA (2002) Evaluating resource selection functions. *Ecol Model* 157(2–3):281–300. [https://doi.org/10.1016/s0304-3800\(02\)00200-4](https://doi.org/10.1016/s0304-3800(02)00200-4)
- Brand LA, George TL (2001) Response of passerine birds to forest edge in coast redwood forest fragments. *Auk* 118(3):678–686. [https://doi.org/10.1642/0004-8038\(2001\)118\[0678:Ropbf\]2.0.Co;2](https://doi.org/10.1642/0004-8038(2001)118[0678:Ropbf]2.0.Co;2)
- Broms KM, Johnson DS, Altwegg R, Conquest LL (2014) Spatial occupancy models applied to atlas data show Southern Ground Hornbills strongly depend on protected areas. *Ecol Appl* 24(2):363–374. <https://doi.org/10.1890/12-2151.1>
- Carl G, Kuhn I (2010) A wavelet-based extension of generalized linear models to remove the effect of spatial autocorrelation. *Geogr Anal* 42(3):323–337
- Carroll C, Johnson DS (2008) The importance of being spatial (and reserved): assessing Northern Spotted Owl habitat relationships with hierarchical Bayesian Models. *Conserv Biol* 22(4):1026–1036. <https://doi.org/10.1111/j.1523-1739.2008.00931.x>
- Carroll C, Johnson DS, Dunk JR, Zielinski WJ (2010) Hierarchical Bayesian spatial models for multispecies conservation planning and monitoring. *Conserv Biol* 24(6):1538–1548. <https://doi.org/10.1111/j.1523-1739.2010.01528.x>
- Cleveland WS (1979) Robust locally weighted regression and smoothing scatterplots. *J Am Stat Assoc* 74(368):829–836. <https://doi.org/10.2307/2286407>
- Cruse B, Liedloff AC, Wintle BA (2012) A new method for dealing with residual spatial autocorrelation in species distribution models. *Ecography* 35(10):879–888. <https://doi.org/10.1111/j.1600-0587.2011.07138.x>
- Cruse B, Liedloff A, Vesik PA, Fukuda Y, Wintle BA (2014) Incorporating spatial autocorrelation into species distribution models alters forecasts of climate-mediated range shifts. *Glob Chang Biol* 20(8):2566–2579. <https://doi.org/10.1111/gcb.12598>
- 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
- Diniz JAF, Bini LM, Hawkins BA (2003) Spatial autocorrelation and red herrings in geographical ecology. *Glob Ecol Biogeogr* 12(1):53–64. <https://doi.org/10.1046/j.1466-822X.2003.00322.x>
- Diniz JAF, Nabout JC, Telles MPD, Soares TN, Rangel T (2009) A review of techniques for spatial modeling in geographical, conservation and landscape genetics. *Genet Mol Biol* 32(2):203–211. <https://doi.org/10.1590/s1415-47572009000200001>
- Dormann CF (2009) Response to Comment on “Methods to account for spatial autocorrelation in the analysis of species distributional data: a review”. *Ecography* 32(3):379–381. <https://doi.org/10.1111/j.1600-0587.2009.05907.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>
- Emerson S, Wickham C, Ruzicka KJ (2015) Rethinking the linear regression model for spatial ecological data: comment. *Ecology* 96(7):2021–2025. <https://doi.org/10.1890/14-0879.1>
- Finley AO, Banerjee S, Gelfand AE (2015) spBayes for large univariate and multivariate point-referenced spatio-temporal data models. *J Stat Softw* 63(13):1–28
- Fletcher RJ Jr (2005) Multiple edge effects and their implications in fragmented landscapes. *J Anim Ecol* 74(2):342–352

- Fortin MJ, James PMA, MacKenzie A, Melles SJ, Rayfield B (2012) Spatial statistics, spatial regression, and graph theory in ecology. *Spatial Stat* 1:100–109. <https://doi.org/10.1016/j.spasta.2012.02.004>
- Garland T, Harvey PH, Ives AR (1992) Procedures for the analysis of comparative data using phylogenetically independent contrasts. *Syst Biol* 41(1):18–32. <https://doi.org/10.2307/2992503>
- Gelman A, Hill J (2007) *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press, New York
- Getis A, Griffith DA (2002) Comparative spatial filtering in regression analysis. *Geogr Anal* 34(2):130–140. <https://doi.org/10.1111/j.1538-4632.2002.tb01080.x>
- Griffith DA, Peres-Neto PR (2006) Spatial modeling in ecology: the flexibility of eigenfunction spatial analyses. *Ecology* 87(10):2603–2613. [https://doi.org/10.1890/0012-9658\(2006\)87\[2603:smietf\]2.0.co;2](https://doi.org/10.1890/0012-9658(2006)87[2603:smietf]2.0.co;2)
- Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecol Model* 135(2–3):147–186
- Haining R (2003) *Spatial data analysis: theory and practice*. Cambridge University Press, Cambridge
- Hastie T, Tibshirani R (1986) Generalized additive models. *Stat Sci* 1:297–310
- Hawkins BA, Diniz JAF, Bini LM, De Marco P, Blackburn TM (2007) Red herrings revisited: spatial autocorrelation and parameter estimation in geographical ecology. *Ecography* 30(3):375–384. <https://doi.org/10.1111/j.2007.0906-7590.05117.x>
- Hines JE, Nichols JD, Royle JA, MacKenzie DI, Gopalaswamy AM, Kumar NS, Karanth KU (2010) Tigers on trails: occupancy modeling for cluster sampling. *Ecol Appl* 20(5):1456–1466. <https://doi.org/10.1890/09-0321.1>
- Hooten MB, Wikle CK, Dorazio RM, Royle JA (2007) Hierarchical spatiotemporal matrix models for characterizing invasions. *Biometrics* 63(2):558–567. <https://doi.org/10.1111/j.1541-0420.2006.00725.x>
- Horn BKP (1981) Hill shading and the reflectance map. *Proc IEEE* 69:14–47
- Hurlbert SH (1984) Pseudoreplication and the design of ecological field experiments. *Ecol Monogr* 54(2):187–211. <https://doi.org/10.2307/1942661>
- Hutto RL, Young JS (2002) Regional landbird monitoring: perspectives from the Northern Rocky Mountains. *Wildl Soc Bull* 30(3):738–750
- Johnson DS, Conn PB, Hooten MB, Ray JC, Pond BA (2013) Spatial occupancy models for large data sets. *Ecology* 94(4):801–808
- Keitt TH, Bjørnstad ON, Dixon PM, Citron-Pousty S (2002) Accounting for spatial pattern when modeling organism-environment interactions. *Ecography* 25(5):616–625
- Kery M, Royle JA (2016) *Applied hierarchical modeling in ecology: analysis of distribution, abundance and species richness in R and BUGS*. Academic, San Diego
- Kissling WD, Carl G (2008) Spatial autocorrelation and the selection of simultaneous autoregressive models. *Glob Ecol Biogeogr* 17(1):59–71. <https://doi.org/10.1111/j.1466-8238.2007.00334.x>
- Koenig WD (1999) Spatial autocorrelation of ecological phenomena. *Trends Ecol Evol* 14(1):22–26. [https://doi.org/10.1016/s0169-5347\(98\)01533-x](https://doi.org/10.1016/s0169-5347(98)01533-x)
- Legendre P (1993) Spatial autocorrelation: trouble or new paradigm? *Ecology* 74(6):1659–1673
- Legendre P, Legendre L (1998) *Numerical ecology*. Elsevier, Amsterdam
- Lennon JJ (2000) Red-shifts and red herrings in geographical ecology. *Ecography* 23(1):101–113. <https://doi.org/10.1034/j.1600-0587.2000.230111.x>
- Lichstein JW, Simons TR, Shriner SA, Franzreb KE (2002) Spatial autocorrelation and autoregressive models in ecology. *Ecol Monogr* 72(3):445–463
- Littell RC, Millien GA, Stroup RD, Schabenberger O (2006) *SAS for mixed models*. SAS Institute Inc., Cary, NC
- MacKenzie DI, Nichols JD, Lachman GB, Droege S, Royle JA, Langtimm CA (2002) Estimating site occupancy rates when detection probabilities are less than one. *Ecology* 83(8):2248–2255

- McCarthy KP, Fletcher RJ, Rota CT, Hutto RL (2012) Predicting species distributions from samples collected along roadsides. *Conserv Biol* 26(1):68–77. <https://doi.org/10.1111/j.1523-1739.2011.01754.x>
- McCullagh P, Nelder JA (1989) *Generalized linear models*. Chapman and Hall, London
- Melles SJ, Fortin MJ, Lindsay K, Badzinski D (2011) Expanding northward: influence of climate change, forest connectivity, and population processes on a threatened species' range shift. *Glob Chang Biol* 17(1):17–31. <https://doi.org/10.1111/j.1365-2486.2010.02214.x>
- Miller JA (2012) Species distribution models: spatial autocorrelation and non-stationarity. *Prog Phys Geogr* 36(5):681–692. <https://doi.org/10.1177/0309133312442522>
- Miller J, Franklin J, Aspinall R (2007) Incorporating spatial dependence in predictive vegetation models. *Ecol Model* 202(3–4):225–242. <https://doi.org/10.1016/j.ecolmodel.2006.12.012>
- Miller DA, Nichols JD, McClintock BT, Grant EHC, Bailey LL, Weir LA (2011) Improving occupancy estimation when two types of observational error occur: non-detection and species misidentification. *Ecology* 92(7):1422–1428. <https://doi.org/10.1890/10-1396.1>
- Nelder JA, Wedderburn RW (1972) Generalized linear models. *J R Stat Soc Series A General* 135(3):370. <https://doi.org/10.2307/2344614>
- Oksanen J, Guillaume B, Friendly M, Kindt R, Legendre P, McGlenn 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
- Rota CT, Fletcher RJ Jr, Evans JM, Hutto RL (2011) Does accounting for detectability improve species distribution models. *Ecography* 34:659–670
- Rousset F, Ferdy JB (2014) Testing environmental and genetic effects in the presence of spatial autocorrelation. *Ecography* 37(8):781–790. <https://doi.org/10.1111/ecog.00566>
- Royle JA, Dorazio RM (2008) *Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations, and communities*. Academic, San Diego
- Sauer JR, Niven DK, Hines JE, Ziolkowski DJ, Pardieck KL, Fallon JE, Link WA (2017) *The North American Breeding Bird Survey, results and analysis 1966–2015, Version 2.07.2017*. USGS Patuxent Wildlife Research Center, Laurel, MD
- Segurado P, Araújo MB, Kunin WE (2006) Consequences of spatial autocorrelation for niche-based models. *J Appl Ecol* 43(3):433–444. <https://doi.org/10.1111/j.1365-2664.2006.01162.x>
- Sokal RR, Oden NL (1978) Spatial autocorrelation in biology II: Some biological implications and four applications of evolutionary and ecological interest. *Biol J Linn Soc* 10(2):229–249. <https://doi.org/10.1111/j.1095-8312.1978.tb00014.x>
- Stephens PA, Buskirk SW, del Rio CM (2007) Inference in ecology and evolution. *Trends Ecol Evol* 22(4):192–197. <https://doi.org/10.1016/j.tree.2006.12.003>
- Swihart RK, Slade NA (1985) Testing for independence of observations in animal movements. *Ecology* 66(4):1176–1184. <https://doi.org/10.2307/1939170>
- Teng SN, Xu C, Sandel B, Svenning JC (2018) Effects of intrinsic sources of spatial autocorrelation on spatial regression modelling. *Methods Ecol Evol* 9(2):363–372
- Thorson JT, Minto C (2015) Mixed effects: a unifying framework for statistical modelling in fisheries biology. *ICES J Mar Sci* 72(5):1245–1256. <https://doi.org/10.1093/icesjms/fsu213>
- Venables WN, Ripley BD (2002) *Modern applied statistics with S*, 4th edn. Springer, New York
- Ver Hoef JM, Peterson EE, Hooten MB, Hanks EM, Fortin MJ (2018) Spatial autoregressive models for statistical inference from ecological data. *Ecol Monogr* 88(1):36–59. <https://doi.org/10.1002/ecm.1283>
- 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\[1045:scipcj\]2.0.co;2](https://doi.org/10.1890/0012-9658(2003)084[1045:scipcj]2.0.co;2)
- Wagner HH (2004) Direct multi-scale ordination with canonical correspondence analysis. *Ecology* 85(2):342–351. <https://doi.org/10.1890/02-0738>
- Wagner HH, Fortin MJ (2005) Spatial analysis of landscapes: concepts and statistics. *Ecology* 86(8):1975–1987. <https://doi.org/10.1890/04-0914>

- Wall MM (2004) A close look at the spatial structure implied by the CAR and SAR models. *J Stat Plan Inference* 121(2):311–324. [https://doi.org/10.1016/s0378-3758\(03\)00111-3](https://doi.org/10.1016/s0378-3758(03)00111-3)
- Wilson MFJ, O’Connell B, Brown C, Guinan JC, Grehan AJ (2007) Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Mar Geod* 30:3–35
- Wintle BA, Bardos DC (2006) Modeling species-habitat relationships with spatially autocorrelated observation data. *Ecol Appl* 16(5):1945–1958. [https://doi.org/10.1890/1051-0761\(2006\)016\[1945:msrwsa\]2.0.co;2](https://doi.org/10.1890/1051-0761(2006)016[1945:msrwsa]2.0.co;2)
- Wood SN (2006) *Generalized additive models: an introduction with R*. Chapman and Hall and CRC, Boca Raton, FL
- Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM (2009) *Mixed effects models and extensions in ecology with R*. Springer, New York