Willis KJ, Araujo MB, Bennett KD, Figueroa-Rangel B, Froyd CA, Myers N (2007) How can a knowledge of the past help to conserve the future? Biodiversity conservation and the relevance of long-term ecological studies. Philos Trans R Soc B 362:175–186.
WHO (2008) World malaria report 2008. World Health Organization/UNICEF, Geneva/New York.
Wilson RJ, Gutierrez D, Gutierrez J, Martinez D, Agudo R, Monserrat VJ (2005) Changes to the elevational limits and extent of species ranges associated with climate change. Ecol Lett 8:1138–1146.
Woodward FI (1987) Climate and plant distribution. Cambridge University Press, Cambridge.
Zhou SR, Zhang DY (2008) A nearly neutral model of biodiversity. Ecology 89:248–258.

Chapter 3
The State of Spatial and Spatio-Temporal Statistical Modeling

Mevin B. Hooten

3.1 Introduction

The purpose of this chapter is to provide an overview of how statistical analyses have been used for studying ecological processes on landscapes and where the field of statistics is headed in general. Various approaches to the statistical analysis of spatial and spatio-temporal problems are presented and discussed; also, references for several suggested readings, containing further information and examples, are provided at the end of each section.

3.1.1 Why Statistics?

Scientific endeavor owes a great debt of gratitude to pioneers in the field of statistics, a relatively young area of study that has undergone significant change (including paradigm shifts, as well as both splits and merges in philosophical underpinnings) since its inception (Stigler 1990; Brown 2000). Historical flux in the field of statistics aside, one thing can be said with certainty: much of the scientific progress made in the past century would not have been possible without it (Salsburg 2001). Contemporary statistical analyses now encompass an incredibly wide range of methods, some of which can be used to study very complicated natural systems while still following the original basic tenet of statistics; that is, formally addressing and characterizing uncertainty when using data to learn about natural phenomena in an inverse fashion. Inverse modeling is the act of using data explicitly to learn about the underlying causal process; this is in contrast to

M.B. Hooten (✉)
USGS Colorado Cooperative Fish and Wildlife Research Unit, Department of Fish, Wildlife, and Conservation Biology, Colorado State University, 201 JVK Wagar Bldg, 1484 Campus Delivery, Fort Collins, CO 80523-1484
e-mail: mevin.hooten@colostate.edu

C.A. Drew et al. (eds.), *Predictive Species and Habitat Modeling in Landscape Ecology: Concepts and Applications*, DOI 10.1007/978-1-4419-7390-0_3,

rd modeling, where models are constructed to simulate possible future vations of the process under study. In light of this distinction, the act of ; statistical models to data is inherently stochastic inverse modeling. So, one ısk, "where do statistical prediction and forecasting fit in?" These are often ht of in a forward modeling context, but chronologically, predictions and ısts depend on the model fit.

ndscape ecology is a field concerned with the study of natural processes arge spatial extents. Many introductory statistical methods require strong ıptions that can be difficult for scientists and managers to justify in practice. ıf the most commonly required assumptions is that of independence among ɔservations, however, most spatial, temporal, and spatio-temporal data are :itly dependent because of latent spatial and temporal autocorrelation. The 'autocorrelation" refers to data or residuals (depending on the context) that ɔrrelated with themselve rather than independent. This can present signifi- hallenges in the development and implementation of appropriate statistical ɜes. On the other hand, these forms of dependence are invaluable for making ifically meaningful predictions and forecasts. Rigorous statistical approaches ·posed to *ad hoc* approaches) then, allow one to formally quantify the inher- ıcertainty in predictions and forecasts. The difference between statistical *d hoc* approaches is often described as "optimality;" that is, rather than ɔorate stochasticity in a haphazard manner, statistical methods provide the stimates and predictions, using the available data. This is accomplished by ng that the estimates have good properties (e.g., unbiasedness, minimum ıce among all other estimates). Therefore, statistical estimates and predic- with high quality properties lead to the best possible scientific inference the available data.

Main Types of Data

.ta will be discussed in detail in the following chapter, only a general ıew is presented here. Statistical analyses of all kinds require quantitative ırements of the process under study. Qualitative observations certainly heir place in the scientific process (usually in the development of hypoth- nd interpretation of inference), but statistical methods are not currently ɔed to use them directly. Therefore, those measurements useful for ics come in two broad varieties: discrete, most commonly in the form of ; or quantified categories, and continuous, often measurements of mass in form. Adequately accommodating various types of data is one of the chief rns in statistics. To guarantee important properties of statistical quantities ust pay careful attention to the many characteristics of data (e.g., orienta- ·cale, measurement error, dependence) as well as modeling assumptions ɔrn and Mangel 1997).

3.2 Statistical Models

3.2.1 Parameters: Fixed or Random?

In general, data are considered as observed random variables; the probability distribution they arise from is generally of interest for making inference. When the form of the distribution is specified as part of the statistical model, the parameters in the distribution often become the subject of interest and thus the approach is labeled "parametric" statistical modeling. Conversely, non-parametric statistical modeling seeks to loosen the distributional assumptions made *a priori*. Non-parametric models are discussed in more detail in Sect. 3.2.5.

In parametric modeling, parameters are predominantly treated as fixed but unknown population quantities to be estimated using data. It should be noted, however, that some of the earliest statistical models [e.g., Laplace's model for astronomical quantities and Bayes' model for billiard balls (Stigler 1990)] considered parameters to be random variables, the probability distribution of which was to be estimated using the data (Carlin and Louis 2000; Salsburg 2001). The difference between the two views leads to a subtle but fundamentally different implementation and interpretation of the results (Clark 2007; Cressie et al. 2009). That is, both forms of modeling are still considered statistical because they serve to formally help learn about natural phenomena in an inverse fashion while accommodating uncertainty (that is, they work backwards from the data toward the parameters, rather than forward from the parameters to the data) (Clark 2007); the primary difference then, is in the resulting inference. If one believes that the true parameters governing the process are indeed fixed quantities then a carefully designed experiment and accompanying frequentist statistical analysis is in order. In this case, inference will be made in terms of long-run frequencies, and thus statements such as, "if the experiment were conducted a large number of times, we would expect to make the same decision approximately 95% of the time," are used to convey the results of the analysis.

On the other hand, the treatment of model parameters as random variables can be useful if data are observational or if measurements were obtained in such a manner that does not guarantee the assumption of fixed parameters holds; in this case, either a frequentist or Bayesian approach (i.e., methods with specification and inference based on conditional probability) may be taken. For example, if a set of measurements is collected over a period of time (which is nearly always the case in ecological studies) then one has to ask themselves if the unknown "population parameter" varied during that time. The same analogy could be applied to measurements collected over space, and thus it is often most appropriate to treat such parameters as random and employ methods that characterize the manner in which they are random (e.g., estimate the inherent stochasticity via a probability distribution). Another situation where random terms can be useful in a statistical model commonly arises in analysis of variance (Neter et al. 1996). To illustrate this, consider the situation where, out of 100 study sites, a random sample of 10 is selected. At each of the selected study sites, a sample of stationed field technicians collects data at their site. In this situation, if the researcher wishes to make general inference about the whole set of one hundred study sites, rather than each of the ten selected sites individually, they could let the ten sites constitute ten levels of a random factor in their analysis.

Frequentist and Bayesian statistical methods can both be employed to help learn about random parameters. Common frequentist approaches to dealing with random parameters are often in conjunction with fixed parameters and take the form of mixed models (Neter et al. 1996, pgs. 978–981) and state-space models (Chap. 6 of Shumway and Stoffer 2006). Many times Bayesian methods are preferred for modeling random parameters (Cressie et al. 2009) because of their flexibility, ease in specification and implementation for complex models, and the ability to directly incorporate prior scientific information (e.g., conclusions resulting from different data or historically documented quantities in the literature). The frequentist approach to parameter estimation is still preferred when frequency-based inference (e.g., confidence intervals) or objectivity in the parameters is desired (Lele et al. 2007), though numerous objective Bayesian methods exist for fitting various statistical models to data (Gelman et al. 2004). Given that statistical analysis can proceed in either fashion, an important question in the model construction phase is whether model parameters should be treated as fixed, random, or some combination of both (i.e., a mixed model).

3.2.2 *Naïve Models*

Conventionally, the dominant type of statistical model used to study natural processes is specified in such a manner that its form facilitates implementation and capitalizes on the rigor of study design in controlled experiments. For example, the linear regression model is often a model of choice for linearly linking a response variable to a set of covariates (Neter et al. 1996), such as in linking coyote abundance to a set of environmental variables like canopy openness, distance to nearest house, or distance to paved road, as discussed by Kays et al. (2008). If specified with independent additive Gaussian error, numerous beneficial properties of the estimated regression coefficients and predictions can be exploited for inference (Christensen 2002). In fact, rather than specify a more scientifically meaningful model (e.g., a nonlinear model with multiplicative error for example), it is a common practice to perform various transformations of the response and/or covariate data to justify necessary assumptions and the use of a linear model. In other words: when you've got a hammer, every problem starts to look like a nail.

It should be noted, however, that a simplified analysis does not imply a useless analysis. That is, naïve models such as linear models with continuous response variables and additive error can be fit easily, and in situations where statistical assumptions hold, they can be readily used to make valid inference about underlying natural processes of interest. Such models can be thought of as "structurally parsimonious," in the sense that they are sparse on model structure, but still useful under certain circumstances. Also, linear models are not always trivial. Consider the common situation in landscape ecological studies where a continuous response variable is measured over a landscape (e.g., soil moisture) and the researcher wishes to investigate its relationship to other important environmental features (i.e., covariates) of the landscape (e.g., slope, aspect, elevation, percent vegetation cover) and also possibly utilize that information to make predictions. In order to appropriately employ multiple linear regression analysis for making inference about the natural process, several model assumptions must be justified. A critical assumption that is often overlooked in such analyses is that the additive model errors are independent. In fact, assessing and accommodating possible spatial dependence in the errors is the premise of geostatistics (Cressie 1993; Diggle and Ribeiro 2007). Erroneous inference is one consequence of failing to account for dependence in the error when it is present (Chap. 9 in Waller and Gotway 2004). This fact is easily shown, and often used as an early exercise in a course on spatial statistics. In short, if residual spatial dependence exists, parameter estimates can be both biased and have incorrect precision (Chap. 6 in Schabenberger and Gotway 2005).

Generalizing the linear model specification used in regression analysis to accommodate residual spatial dependence is relatively simple yet adds significant complexity to the fitting procedure. That is, rather than assume observations can be modeled by a large-scale trend (involving spatial covariates and an associated set of regression coefficients) plus some independent measurement error, we wish to allow for possible dependence in these additive errors. In this way, any potential residual autocorrelation beyond what can be explained by the large-scale trend may be accounted for. In most cases, such autocorrelation in the errors can be characterized via variogram estimation and modeling (Chap. 2 in Cressie 1993), and then incorporated into the linear model for parameter estimation and prediction. On the surface, the regression model still looks the same (i.e., response = covariate effects + error), though the incorporation of correlated error necessitates a slightly more complicated estimation procedure for the regression coefficients (i.e., generalized least squares rather than ordinary least squares).

Once the residual spatial autocorrelation is taken into account, the prediction of continuous spatial processes is referred to as Kriging (Chap. 3 in Cressie 1993) and can be employed, under certain distributional assumptions, with relative ease; in fact, this can often be accomplished at the click of a button in many geographic information systems (GIS) and statistical software. Many types of Kriging have been developed for spatial prediction in various circumstances. For example, Ver Hoef et al. (2006) provide a method for extending Kriging from the standard Euclidean setting to stream networks.

When the response variable of interest has discrete support (e.g., presence/absence or counts of an organism at various locations across a landscape) similar naïve models can be useful for linking the observed natural process to a set of covariates. These models are referred to as generalized linear models and many specifications require similar assumptions about independence of errors, but can also be modified to accommodate correlated errors if necessary (Chap. 6 in Schabenberger nd Gotway 2005). Some of the most common generalized linear models are for inary data (i.e., logistic and probit regression) and can be used for studying presnce/absence or occupancy (Royle and Dorazio 2008). For example, Hooten et al. 2003) and Gelfand et al. (2006) present similar approaches for modeling vegetation bundance on a landscape using generalized linear models and binary data. In the ormer, Hooten et al. (2003) use presence/absence data on forest understory legumes i.e., *Desmodium glutinossum* and *D. nudiflorum*) collected over a large number of lots spread across a Southern Missouri watershed as part of the Missouri Ozark 'orest Ecosystem Project. In this study, large-scale spatial predictions of these plant istributions were desired. Thus, a generalized linear mixed model was specified to xplicitly accommodate the binary data while characterizing the underlying probaility of presence in terms of a set of spatial covariates (i.e., aspect, elevation, land ype, and soil depth) and latent spatial autocorrelation. This model allowed for the rediction of probability of presence across the entire study area as well as provided naps of prediction standard deviation as a measure of uncertainty in the redictions.

In situations where boundless counts of organisms are the response variable of nterest, a Poisson regression approach can be taken (see Royle and Dorazio 2006 or an example of avian abundance modeling). Another study, by Royle and Wikle 2005), discusses a generalized linear model for predicting avian abundance across he Eastern United States using North American Breeding Bird Survey (BBS) data. n their study, they assumed that BBS route counts of species followed a Poisson istribution. They incorporated covariate effects and spatial autocorrelation in the nuch same manner as Hooten et al. (2003), but in this case, rather than probability f presence; they linked these spatial effects to the log of the Poisson intensity arameter. This allowed Royle and Wikle to make large-scale predictive maps for ird abundance (specifically for Carolina Wren in this study) as well as maps of rediction uncertainty.

Numerous significant scientific findings have benefited from the use of a naïve nodel structure such as the linear model, but new tools have come to light with dvancements in statistical theory and the advent of high performance personal omputers.

.2.3 *Scientific Models*

n this section, "scientific" is used to describe those statistical models that explicitly ncorporate mathematical and/or physical processes. Such specifications are often nost useful for studying time-evolving natural processes because they can incorpoate explicit dynamic behavior (Hilborn and Mangel 1997). Because landscape cology involves the study of spatial systems, relevant statistical models with a emporal component are termed spatio-temporal.

The study of dynamical systems has a long history in both pure and applied nathematics but only recently has it become prominent in statistics. As with static

systems such as the spatial-only examples of the last section, naïve statistical models can be employed for studying temporal systems. In these cases, the "dynamics" (i.e., the components of the model controlling the change in the system being studied over time) are expressed in a general form that may be flexible but lacks a direct scientific interpretation. The temporal autoregressive specification is an example of a naïve time-series model where the form contains a distinct dynamic component but in most cases is over-simplified (Chap. 9 in Clark 2007). Employed in an ecological setting, such models may capture dynamic behavior and can often be useful for making inference but are not built on formal principles of ecological theory. Hooten and Wikle (2007) provide an example of a naïve spatio-temporal model that was used for studying the changes in dynamics of forest growth. In this study, a vector autoregressive model is used to analyze the differences in a reduced dimensional dynamical system (representing the spatio-temporal growth in shortleaf pine forests) before and after an anthropogenically created change-point and in response to climatic fluctuation. Their findings included a notable acceleration in the temporal evolution of shortleaf pine growth after a massive clearing of forest at the turn of the twentieth century and also in response to periods of drought. This model can be considered naïve because, although it is dynamic, the dynamics are represented by a simple autoregressive evolution equation where the estimated parameters have no inherent scientific meaning or interpretation.

In contrast to naïve models, scientific models for studying ecological systems on a landscape over time explicitly incorporate meaningful physical processes. For example, the diffusion (i.e., dispersal in ecological terms) of a natural phenomenon through a medium can be expressed using a number of different mathematical models such as:

- *Integro-Difference Equations (IDE)*: Wikle (2001) models a dynamic atmospheric process by integrating the product of two functions; one describing the increase in cloud intensity over time and the other represented by a spatial redistribution kernel that describes cloud spread over time. The distinguishing characteristic in IDE models is that the dynamical component operates via integration.
- *Partial Differential Equations (PDE)*: The mathematical opposite of integration is differentiation and this can also be a reasonable way to describe some natural dynamic processes. For example, Wikle (2003) places a spatio-temporal PDE model into a statistical framework for describing the spread of an invasive species over the North American continent using BBS count data. These models are distinguished from IDE models by the fact that the change in the underlying process of interest (e.g., bird abundance) is expressed in terms of spatial and temporal derivatives.
- *Markov Matrix Equations*: Another approach to modeling both spreading and growing phenomena is through the use of matrix models (Caswell 2001). Though typically employed to study changes in population demographics, matrix models can also be placed in a statistical spatio-temporal context. For example, Hooten et al. (2007) use a spatio-temporal matrix model to characterize and forecast the invasion of the Eurasian Collared-Dove in North America. While PDE and IDE models are inherently continuous in time and space, matrix models are derived explicitly in a discrete spatio-temporal setting.

- *Agent-Based Models*: Though the class of agent-based models is quite large, many consider them to include individual-based models. In general, agent-based models can be thought of as bottom-up models (as opposed to top-down), and are constructed by specifying how a small scale process behaves and then scaling them up to examine their larger-scale properties (Grimm and Railsback 2005). Given that this is how many believe all natural systems work, agent-based modeling has great potential. Hooten and Wikle (2010) construct a bottom-up statistical model to describe a spreading epidemiological process. Specifically, they specify a spatio-temporal cellular automata model that is capable of characterizing the complex dynamical behavior of the rabies epidemic as it spreads through raccoon populations in Connecticut.

3.2.4 *Hierarchical Models*

Many of the naïve models discussed earlier can be specified hierarchically (and are in many of the references provided), though complicated scientific models (including spatial and spatio-temporal dynamic models) can often be formulated with ease using a hierarchical framework. It is important to note here that the term "hierarchical," though always similar in spirit, is used differently across disciplines. In statistics, a hierarchical model is one that specifies a complicated joint probability distribution in terms of a set of simpler conditional distributions using well-known results from probability (see Cressie et al. 2009 for an excellent overview). In essence, this allows the modeler to break up a large intractable problem into a set of simpler problems that can be readily solved. Though the details are technical in nature, the basic premise is intuitive. Bayesian methods are particularly useful for specifying and fitting hierarchical models and thus have become very popular recently in complicated statistical analyses. When specifying a Bayesian hierarchical model, one can generally consider three main components (Berliner 1996): The data model (i.e., likelihood), the process model, and the parameter model (i.e., prior distribution). The product of these three appropriately scaled models (i.e., probability distributions) yields the "posterior distribution," a joint distribution of the model parameters and process given the data. This distribution is generally unknown and analytically intractable, hence the need to specify it in terms of a set of simpler conditional models. The first and second components (i.e., the likelihood and process model) by themselves have been considered from a traditional perspective in statistics, where it is the third component (i.e., the prior) that is both necessary and useful in a Bayesian implementation. In principle, the prior distribution contains all of the information about the model parameters that is available before the current data were collected. In practice, it is sometimes the case that little or no prior information exists, or if so, cannot be specified in terms of a probability distribution; thus, in such cases, a non-informative or vague distribution is then used for the parameter model so that any statistical learning is forced to come from the data rather than from an exogenous source. The posterior distribution is then used to make inference and can

be thought of as the distribution of the process and parameters that has been updated (from the prior) using the data.

For further detail and examples of statistical hierarchical modeling in ecology see Banerjee et al. (2004), Zhu et al. (2005), Clark and Gelfand (2006), Arab et al. (2007), and Royle and Kery (2007). Note that although the methodological details may be technical and custom software is often necessary, many statistical packages, tutorials, and open-source code for fitting such models are readily available. One caveat is that, in high-dimensional settings, the algorithms used to implement hierarchical Bayesian models can be computationally cumbersome. Ecological science is currently transitioning from being data-poor to data-rich. With GIS layers, automated monitoring devices, and remotely sensed data becoming more prevalent, the field of statistics is rapidly adapting to meet such challenges. New methods for using statistical models with massive datasets are being regularly developed for the purposes of obtaining predictions in high-dimensional settings. Furrer et al. (2006) and Shi and Cressie (2007) have developed methods for modeling high-dimensional spatial datasets and illustrate their utility using examples pertaining to satellite data. Specifically, Shi and Cressie (2007) employed fixed rank kriging (Cressie and Johannesson 2006) to obtain global predictions of atmospheric aerosols (an important forcing component for climate models) using massive remotely sensed datasets resulting from the MISR sensor on NASAs Terra satellite. These rigorous statistical methods are shown by Shi and Cressie (2007) to be superior over the previously used *ad hoc* approaches.

3.2.5 *Semi- and Non-parametric Models*

In the current scientific era, statistics is transforming from a field that sought to get as much information as possible out of a small amount of data into a field that needs to reduce the dimensionality of the data before gleaning any useful information. Where scientific modeling seeks to explicitly introduce information about the underlying physical process into the model (which could be thought of as supervised model building), the statistical sub-discipline of machine learning (i.e., data mining) seeks to uncover naturally occurring relationships in data rather than build in predefined ones (Hastie et al. 2001).

Non-parametric statistical methods generally take a distribution-free approach to modeling. That is, they seek to make as few a priori assumptions about the data as possible. This can be incredibly valuable when data do not conform to conventional modeling assumptions and/or when no scientific modeling approach is obvious or available. Many non-parametric methods are able to fit data and make predictions extremely well, often better than any other technique; the only downside is that they sacrifice scientific interpretability. That is, because they do not explicitly build scientific information into the model itself, they must rely on *post hoc* scientific interpretations and inference. In fact, even though methods for machine learning involve elegant mathematical theory, they are still often treated as mere "black box" ıodels by their users; however, the same could be said about many conventional atistical methods. See Cutler et al. (2007) and Chap. 8 in this volume for an ‹ample of a promising new non-parametric method, called "random forests," ɔplied to invasive plant classification, rare lichen presence, and nesting site prefer-ıce by cavity nesting birds.

Semi-parametric statistical models include both parametric and non-parametric ɔmponents (Chap. 9 in Hastie et al. 2001). Various smoothing splines or wavelets ·e often used to implement such models, and in some specific cases, spatial predic-ons using these methods end up being equivalent to those via Kriging (Nychka ℑ00). Additive semi-parametric models are generally specified in a regression-yle framework and can be useful for accommodating more complicated nonlinear ːlationships between response variables and covariates (Efron and Tibshirani 991). For example, Holan et al. (2008) present an approach to modeling site-specific ·op response to varying treatments (e.g., irrigation) in a spatial setting using ːmi-parametric relationships between the response and predictors.

.3 Optimal Design

he notion of optimal sampling is not new; however, extensions of this concept to ıe spatial and spatio-temporal setting are being proposed with more regularity ℑlea 1984; Cressie 1993). In this setting, the term "optimal" implies that the ːlected sampling design performs the best with regard to some design objective (Le nd Zidek 2006). The choice of the design objective can vary depending on the goals f the study. For example, if the study involves the collection of spatial data at a nite set of locations across a landscape with the hopes of learning about a natural rocess at unobserved locations within the study area, then a design criterion based n spatial prediction error may be reasonable. In that case, one would examine all ossible sampling designs to find the one design that minimizes prediction error ariance (Stevens and Olsen 2004). In other situations with more vague study goals ː.g., data are being collected for many purposes and/or future modeling efforts), one ıay wish to reduce uncertainty in general; entropy-based design methods have been roposed and successfully used in these cases (Le and Zidek 2006). Regardless of ıe design criterion chosen, optimality allows one to get the most "bang for their uck" out of the data collected. That is, by exploiting properties pertaining to the ıanner in which the natural process is observed, it is possible to reduce uncertainty ı the information gleaned from the data in a more efficient manner rather than just ɔllecting more data. A simple example of how an optimal design could be useful is ·hen more efficient estimation of the spatial structure in the process is desired. ı such a case, both a regularly spaced sampling scheme and a fully random sampling ːheme are inefficient (i.e., sub-optimal); however, a combination of the two can erform substantially better by providing better coverage of the study area (and thus eing more representative) while still capturing small scale spatial structure in the rocess (Stevens and Olsen 2004; Zhu and Stein 2006; Zimmerman 2006).

Optimal design methods have also been proposed for spatio-temporal settings (Wikle and Royle 1999). That is, if one seeks to characterize a dynamically evolv-ing system, specific forms of uncertainty can be reduced if care is taken in the construction of the sampling or monitoring design over time. In the current age of data collection, with remote wireless sensors constantly measuring features of the environment and mobile roving sensors capable of adaptively sampling a spatial domain, dynamic optimal monitoring designs will become more prevalent and use-ful. As an example, in a dynamic setting where a process is being monitored repeat-edly over time and the data collected are to be used for various purposes including estimation of the dynamics of the system and statistical forecasting, the design objective may be to reduce uncertainty in the forecast (possibly via prediction error variance). In this case, several monitoring approaches could be taken to observe the system. If the system is evolving dynamically (i.e., changing over time) it is sen-sible to allow the design to change over time as well. This way, the monitoring scheme can adapt to capture important aspects of the behavior in the system. Allowing for roving monitors over time can help to avoid re-sampling redundant behavior and instead move to where the action is occurring. Ultimately this can maximize the power of the collected data given the available resources. An example of this kind of optimal design implemented to study plant community dynamics is presented by Hooten et al. (2009).

3.4 Conclusion

Statistics is an ever-changing field, constantly adapting to the new developments and needs of the scientific disciplines that depend on it. In this new and exciting com-puter era, we are witnessing a blurring of the lines between previously distinct areas of study. Methods for statistically analyzing spatial processes are built into GIS software, where spatial data manipulation has been occurring for decades; GIS tools are also being built into statistical software to aid in exploratory analyses and visu-alization of modeling results on spatial and spatio-temporal domains. In the face of mountains of data, machine learning and data mining methods are becoming more prominent, as is the need for formal data management skills. We are also seeing the direct integration of information formerly restricted to applied mathematical and physical studies into rigorous statistical analyses. Likewise, it can be said that new statistical methods (e.g., hierarchical Bayesian models) are being readily used in the applied mathematics literature. Statistics is growing and changing and we are rapidly approaching a time where every scientific problem, no matter how complex, can be considered naturally in a statistical framework. Landscape ecology, as a field con-cerned with multidimensional systems and an abundance of data, stands to benefit greatly from new statistical methods for analyzing spatial and spatio-temporal data.

Acknowledgments The author would like to thank A. Arab, C. Wikle, and two anonymous reviewers of this chapter for their helpful comments and suggestions.

References

Arab A, Hooten MB, Wikle CK (2007) Hierarchical spatial models. In: Encyclopedia of geographical information science. Springer, New York.

Banerjee S, Carlin BP, Gelfand AE (2004) Hierarchical modeling and analysis for spatial data. Chapman & Hall/CRC, Boca Raton, FL.

3erliner LM (1996) Hierarchical Bayesian time series models. In: Hanson K, Silver R (eds), Maximum entropy and Bayesian methods, pp. 15–22. Kluwer Academic Publishers, New York.

3rown LD (2000) An essay on statistical decision theory. J Am Stat Assoc. 95:1277–1281.

Carlin BP, Louis TA (2000) Bayes and empirical Bayes methods for data analysis, Second Edition. Chapman & Hall/CRC, Boca Raton, FL.

Caswell H (2001) Matrix population models: construction, analysis, and interpretation. Sinauer Associates, Inc., Sunderland, MA.

Christensen R (2002) Plane answers to complex questions. Springer-Verlag, New York.

Clark, JS (2007) Models for ecological data. an introduction. Princeton University Press, Princeton, NJ.

Clark JS, Gelfand AE (2006) Hierarchical modelling for the environmental sciences. Oxford University Press, New York.

Cressie NAC (1993) Statistics for spatial data, Revised Edition. John Wiley & Sons, New York.

Cressie NC, Calder C, Clark JS, Ver Hoef JM, Wikle C (2009). Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. Ecol Appl 19:553–570.

Cressie N, Johannesson G (2006) Fixed rank kriging for large spatial datasets. Technical Report No. 780, Department of Statistics, The Ohio State University, Columbus, OH.

Cutler DR, Edwards TC, Beard KH, Cutler A, Hess KT, Gibson J, Lawler JJ (2007) Random forests for classification in ecology. Ecology 88:2783–2792.

)iggle PJ, Ribeiro PJ, Jr (2007) Model-based geostatistics. Springer, New York.

ifron B, Tibshirani R (1991) Statistical analysis in the computer age. Science 253:390–395.

'urrer R, Genton MG, Nychka D (2006) Covariance tapering for interpolation of large spatial datasets. J Comput Graph Stat 15:502–523.

ielfand AE, Silander JA, Wu S, Latimer A, Lewis PO, Rebelo AG, Holder M (2006) Explaining species distribution patterns through hierarchical modeling. Bayesian Anal 1:41–92.

ielman A, Carlin JB, Stern HS, Rubin DB (2004) Bayesian data analysis: second edition, Chapman and Hall/CRC, Boca Raton, FL.

irimm V, Railsback SF (2005) Individual-based modeling and ecology. Princeton University Press, Princeton, NJ.

Iastie T, Tibshirani R, Friedman J (2001) The elements of statistical learning. Springer, New York.

Iilborn R, Mangel M (1997) The ecological detective, confronting models with data. Princeton University Press, Princeton, NJ.

olan S, Wang S, Arab A, Sadler J, Stone K (2008) Semiparametric geographically weighted response curves with application to site-specific agriculture. J Agric Biol Environ Stat 13:424–439.

ooten MB, Wikle CK (2007) Shifts in the spatio-temporal growth dynamics of shortleaf pine. Environ Ecol Stat 14:207–227.

ooten MB, Wikle CK (2010) Statistical agent-based models for discrete spatio-temporal systems. J Am Stat Assoc 105:236–248.

ooten, MB, Larsen DR, Wikle CK (2003) Predicting the spatial distribution of ground flora on large domains using a hierarchical Bayesian model. Landsc Ecol 18:487–502.

ooten MB, Wikle CK, Dorazio RM, Royle JA (2007) Hierarchical spatiotemporal matrix models for characterizing invasions. Biometrics 63:558–567.

ooten MB, Wikle CK, Sheriff S, Rushin J (2009) Optimal spatio-temporal hybrid sampling designs for monitoring ecological structure. J Veg Sci 20:639–649.

ays RW, Gompper ME, Ray JC (2008) Landscape ecology of eastern coyotes based on large-scale estimates of abundance. Ecol Appl 18:1014–1027.

Le ND, Zidek JV (2006) Statistical analysis of environmental space-time processes. Springer, New York.

Lele SR, Dennis B, Lutscher F (2007) Data cloning: easy maximum likelihood estimation for complex ecological models using Bayesian Markov chain Monte Carlo methods. Ecol Lett 10:551–563.

Neter J, Kutner MH, Nachtsheim CJ, Wasserman W (1996) Applied linear statistical models. WCB/McGraw-Hill, Boston.

Nychka D (2000) Spatial-process estimates as smoothers. In: Schimek, MG (ed) Smoothing and regression: approaches, computation, and application. John Wiley & Sons, New York.

Olea RA (1984) Sampling design optimization for spatial functions. Math Geol 16:369–392.

Royle JA, Dorazio RM (2006) Hierarchical models of animal abundance and occurrence. J Agric Biol Environ Stat 11:249–263.

Royle JA, Dorazio RM (2008) Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations, and communities. Academic Press, London.

Royle JA, Kery M (2007) A Bayesian state-space formulation of dynamic occupancy models. Ecology 88:1813–1823.

Royle JA, Wikle CK (2005) Efficient statistical mapping of avian count data. Environ Ecol Stat 12:225–243.

Salsburg D (2001) The lady tasting tea: how statistics revolutionized science in the twentieth century. Henry Holt and Company, New York.

Schabenberger O, Gotway CA (2005) Statistical methods for spatial data analysis. Chapman & Hall/CRC, Boca Raton, FL.

Shi T, Cressie NAC (2007) Global statistical analysis of MISR aerosol data: a massive data product from NASA's Terra satellite. Environmetrics 18:665–680.

Shumway R, Stoffer DS (2006) Time series analysis and its applications. Springer, New York.

Stevens DL, Olsen AR (2004) Spatially balanced sampling of natural resources. J Am Stat Assoc 99:262–278.

Stigler SM (1990) The history of statistics: the measurement of uncertainty before 1900. Harvard University Press, Cambridge.

Ver Hoef JM, Peterson E, Theobald D (2006) Spatial statistical models that use flow and stream distance. Environ Ecol Stat 13:449–464.

Waller LA, Gotway CA (2004) Applied spatial statistics for public health data. John Wiley & Sons, Inc. Hoboken, NJ.

Wikle CK, Royle JA (1999) Space-time models and dynamic design of environmental monitoring networks. J Agric Biol Environ Stat 4:489–507.

Wikle CK (2001) A kernel-based spectral approach for spatio-temporal dynamic models. Proceedings of the 1st Spanish workshop on spatio-temporal modelling of environmental processes (METMA), Benicassim, Castellon (Spain), 28–31 October 2001, pp. 167–180.

Wikle CK (2003) Hierarchical Bayesian models for predicting the spread of ecological processes. Ecology 84:1382–1394.

Zhu J, Huang H-C, Wu C-T (2005) Modeling spatial-temporal binary data using Markov random fields. J Agric Biol Environ Stat 10:212–225.

Zhu Z, Stein M (2006) Spatial sampling design for prediction with estimated parameters. J Agric Biol Environ Stat 11:24–44.

Zimmerman DL (2006) Optimal network design for spatial prediction, covariance estimation, and empirical prediction. Environmetrics 17:635–652.