



OXFORD LIBRARY OF PSYCHOLOGY

Editor-in-Chief PETER E. NATHAN

The Oxford Handbook of Quantitative Methods

Edited by

Todd D. Little

VOLUME 2: STATISTICAL ANALYSIS

2013

Spatial Analysis

Luc Anselin, Alan T. Murray, and Sergio J. Rey

Abstract

This chapter provides a broad outline of spatial analysis, a collection of methods that share the common characteristic that spatial proximity and relationships matter. We focus on three specific areas: exploratory spatial data analysis, spatial regression and spatial optimization.

Key Words: spatial analysis, geospatial analysis, exploratory spatial data analysis, spatial regression, spatial econometrics, spatial optimization

Introduction

Spatial analysis (now often referred to as *geospatial analysis*) is broadly defined as a “set of methods useful when the data are spatial” (Goodchild and Longley 1999). More specifically, it encompasses a collection of techniques to add value to data contained in a geographic information system (GIS). Such data are *georeferenced*, which means that in addition to value (attribute) the location of the observation is known, typically expressed in a Cartesian coordinate system. As such, spatial analysis forms an important component of the evolving discipline of *Geographic Information Science* (Goodchild 1992). It encompasses many different methodologies that share the common characteristic that spatial proximity and relationships matter. The recent text by de Smith et al. (2007) includes four main components of spatial analysis, in addition to the basic analytic manipulations embedded in most GIS: data exploration and spatial statistics, surface and field analysis, network and location analysis, and geocomputational methods and modeling. In this chapter, we present a slightly different perspective and focus on three broad categories of methodologies: exploratory spatial data analysis, spatial regression analysis and spatial optimization.

An important reason for the growth of geospatial analysis over the past twenty-some years was the realization in the late 1980s that the technology of geographic information systems (and especially desktop systems) provided an excellent opportunity to operationalize and take advantage of the wealth of analytical techniques developed in the quantitative geography literature. In addition, the combination of the computing power in GIS and advanced methods of spatial analysis provided the opportunity to develop integrated systems that contributed not only to practice, but also led to scientific advances and new methods. Early discussions of the integration of spatial analytical methods with GIS can be found in Goodchild (1987), Goodchild et al. (1992), and Anselin and Getis (1992), among others. Goodchild (2010) offers a more recent perspective. An important institutional factor was the establishment in the U.S. of the National Center for Geographic Information and Analysis (Abler 1987), which, through funding from the National Science Foundation provided a major impetus for the development and adoption of spatial analytical methodology. A similar role was played about ten years later by the NSF funded Center for Spatially Integrated Social Science (Goodchild et al. 2000).

Early compilations of methods, applications and software tools for geospatial analysis can be found in Fotheringham and Rogerson (1994), and Fischer and Getis (1997), among others. More recent reviews include Fotheringham and Rogerson (2009), Anselin and Rey (2010), and Fischer and Getis (2010). Extensive technical detail can be found in those references.

In the remainder of this chapter, we provide a concise overview of exploratory spatial data analysis, spatial regression analysis and spatial optimization modeling. We close with some concluding remarks.

Exploratory Spatial Data Analysis

Arguably, one of the first steps in the empirical analysis of any spatially referenced data should be the exploration of the data. This exploration can serve two purposes. First, as part of a model building strategy, application of *Exploratory Spatial Data Analysis* (ESDA) methods to detect any data anomalies, recording errors, or other data related problems can avoid unnecessary work further on the research path where model calibration (see Section 8) and application may have been for naught or misguided were these data problems missed.

While data exploration is good practice in any empirical study, it is particularly critical in the analysis of spatial data as often georeferenced data requires a good deal of preprocessing and data integration, such as georegistration (setting the coordinates for spatial objects), conversion between different data structures (e.g., raster or grid and vector or polygon), and a host of other manipulations often done in the context of a GIS. These manipulations typically involve a chain of sequential steps, each one potentially introducing sources of uncertainty and/or error into the derived data.

Closely related to data checking is the use of ESDA methods for model validation and diagnostics. For example, tests for departures from the assumptions underlying model specification in spatial regression analysis rely on methods that are extensions of ESDA.

These two examples of use cases for ESDA methods reflect their *data validation function*. Perhaps the more commonly encountered role for ESDA methods, however, is the *data insight function* these methods provide. In this sense ESDA can be viewed as a special case of exploratory data analysis (EDA), the branch of computational statistics pioneered by John Tukey (Tukey 1977). EDA is a largely model-free set of statistical and graphical methods designed

to uncover data errors or anomalies, identify patterns, and generate new insights about the processes under study as reflected in the data that would otherwise have remained hidden. As a result of EDA, new hypotheses about the underlying phenomena may be suggested. The generation of hypotheses stands in marked contrast to a model driven analysis. For the latter, one begins from an initial hypothesis that is used to specify models for calibration, estimation and validation. Indeed the genesis for EDA was the view, held by Tukey, that the classic approach to inference was overly dependent on prior hypothesis and thus restrictive in nature.

When viewed as a branch of EDA, ESDA shares many of EDA's goals and philosophy, as well as a reliance on both numerical and graphical methods. However, the development of ESDA methods has been driven by the recognition that the special nature of spatial data required extension of existing EDA methods as well as the development of entirely new methods. Moreover, some of the methods now widely used and recognized as core components of ESDA actually predate the development of both EDA and ESDA, so the nesting of ESDA inside EDA is not as clean as the above might suggest.

In what follows we first highlight the particular characteristics of spatial data that necessitate ESDA methods. From there we focus on a selection of what we see as the central methods in the ESDA toolkit: spatial autocorrelation analysis; spatial clustering; and extensions for space-time data. We recognize that ESDA is a relatively new and rapidly evolving subfield of spatial analysis and we direct the reader interested in further details to overviews in Anselin et al. (2006), Haining et al. (1998), Unwin and Unwin (1998), Anselin (1999), Bailey and Gatrell (1995).

Spatial Data

Before outlining the main techniques in ESDA it is important to distinguish the three types of spatial data commonly considered in spatial analysis, i.e., point patterns, geostatistical data, and lattice data (Cressie 1991).

Point pattern data takes the form of events recorded within some bounded area or region, such as the location of crimes that occurred within a neighborhood, accidents on a street network or retail outlets within a city. Interest centers on determining whether the points display spatial clustering or some other departure (such as dispersion) from that expected in a completely random spatial process. For overviews of point pattern analysis the reader is

directed to Diggle (2003) and the references cited therein.

Geostatistical data are used for phenomena that could conceptually be measured everywhere in space and modeled as a continuous three-dimensional surface, as is commonly encountered in the physical sciences. Examples of such phenomena that are relevant for social sciences would include air temperature and pollution levels in studies of public health. Because it is impossible to in fact measure the phenomena at all locations, often samples are taken at discrete locations and surfaces are constructed for the values of the attribute at other locations using various interpolation methods, such as kriging. Overviews of geostatistical methods can be found in Journel and Huijbregts (1978), Cressie (1991), Chilès and Delfiner (1999) and Rossi et al. (1992), among others.

A final type of spatial data consists of values (attributes) measured for a fixed set of areal units, or so-called *lattice data*. The areal units often are administrative regions such as census tracts/blocks, counties, or states. Here, the focus is on analyzing the variation in the values across the spatial units. Lattice data analysis is distinct from geostatistical data analysis, since in the latter there are an infinite number of locations, while in lattice analysis the number of spatial units is fixed and finite (e.g., the number of counties in a state). The focus on attribute variation across the spatial units in lattice data analysis contrasts with the focus on the relative location of events in point pattern analysis. In what follows we limit our focus to the analysis of lattice data.

Spatial Autocorrelation Analysis

A common characteristic of spatially referenced data is to exhibit similar values in nearby locations. This association between value similarity and spatial similarity is known as spatial autocorrelation, and is a reflection of the so-called *first law of geography*:

"Everything is related to everything else, but near things are more related than distant things" (Tobler 1970).

Spatial autocorrelation can be either positive, reflecting a nonrandom spatial distribution where like values cluster in space, or, less commonly, negative where the autocorrelation reflects nonrandom value dissimilarity in space. In either case the pattern is different from what would be expected if the values were randomly distributed in space.

Spatial autocorrelation can arise in a number of ways. Measurement errors, manipulations such as interpolation (determining values for locations where no observations are available) and problems with the difference between the spatial extent of the process under study and the spatial units at which it is observed are a number of reasons why spatial autocorrelation can be induced in data that was otherwise randomly distributed. With the advent of GIS software, this is an increasingly important concern. Since in this instance the autocorrelation is an artifact of the data construction process, this type of autocorrelation is referred to as *nuisance autocorrelation*. Alternatively, the autocorrelation could reflect the operation of a substantive process, such as in the case of migration or diffusion, interacting agents, or mimicking of nearby behavior by policy making units (such as copy-catting tax rates in adjoining locations), among others. Because the autocorrelation is thus central to an enhanced understanding of the process, it is referred to as *substantive autocorrelation*.

Irrespective of whether the autocorrelation is of a nuisance or substantive nature, spatial autocorrelation has major implications for the statistical analysis of spatially referenced data. Because spatial autocorrelation is a form of statistical dependence, the assumption of random sampling no longer holds. Consequently, carrying out a test for spatial autocorrelation should precede any application of inferential methods to spatially referenced data. Below we outline the main approaches to such testing.

Spatial autocorrelation can also be analyzed from either a global or local perspective. Global autocorrelation is a whole-map property. That is, whether the spatial distribution of attribute values displays clustering or not. Local autocorrelation analysis is relevant when one is interested in detecting departures from the global pattern, or in identifying the specific location of hot (cold) spots that might form individual clusters.

GLOBAL AUTOCORRELATION

The most widely used measure of spatial autocorrelation in an ESDA context is Moran's I (Moran 1948, 1950). For a set of n spatial observations for a variable y , I is given as:

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^n \sum_{j=1}^n z_i w_{ij} z_j}{\sum_{i=1}^n z_i^2} \quad (1)$$

where the z variable represents the deviation from the mean of the original variable, $z_i = y_i - \bar{y}$. The symbol w_{ij} warrants some special attention. It

stands for the so-called spatial weights that define *a priori* which pairs of locations are likely to interact. The weights are non-zero when two locations i and j are "neighbors," usually determined on the basis of a geographical criterion. The most commonly used criteria are sharing a common border or being within a critical distance of each other. The collected weights are typically referred to as a spatial weights "matrix" \mathbf{W} , of the same dimension as the number of observations ($n \times n$) and with zero on the diagonal by convention. The term S_0 is then the sum of all the elements in the weights matrix, or $S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{ij}$. Upon closer examination, Moran's I can be seen to be similar to a correlation coefficient, with a measure of spatial covariance in the numerator (the sum of cross-products of neighboring values), with the numerator similar to a measure of variance. This becomes clear when the statistic is rewritten as:

$$I = \frac{\sum_{i=1}^n \sum_{j=1}^n z_i w_{ij} z_j}{S_0} \quad (2)$$

Inference on I can be based on a number of different approaches. However, irrespective of which approach is adopted, the null hypothesis is that the y values are randomly distributed in space, in which case the expected value of I is:

$$\mathbb{E}[I] = \frac{-1}{n-1} \quad (3)$$

Note that, unlike the familiar correlation coefficient, this mean is not centered on zero, but slightly to the negative (see Cliff and Ord 1981, for a formal derivation). Since the denominator $n - 1$ becomes larger with n , in large samples, the mean will approach zero.

The variance of the test statistic can be derived analytically under an assumption of normality, or an assumption of non-free sampling (randomization). To evaluate the significance of the statistic, the I value is converted into a so-called z-value which is evaluated as a standard normal variate (for technical details, see, e.g., Cliff and Ord 1981).

An alternative to the analytical approach is to rely on random permutations of the observed values across the spatial units to develop the distribution of I under the null of spatial independence. Such *permutation tests* are quite common in spatial analysis whenever analytical results are difficult (or impossible) to obtain. More specifically, in the case of Moran's I , if the observed value is I_{obs} this is compared to the reference distribution which is constructed from M random permutations (or synthetic

maps). For each of these random maps the statistic is recalculated $I_m : m = 1, 2, \dots, M$. A one-tailed, or directional, pseudo significance level for the statistic can be expressed as:

$$\text{Prob}[I \geq I_{obs} | H_0] = \frac{\Phi + 1}{M + 1} \quad (4)$$

where $\Phi = \sum_{m=1}^M \Phi_m$ and:

$$\Phi_m = \begin{cases} 1 & \text{if } I_m \geq I_{obs} \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

This pseudo significance level consists of the ratio of the number of times the simulated statistic equals or exceeds the observed value I_{obs} plus one (for the observed statistic) over the number of random permutations plus one (again, for the observed statistic). For example, if 4 simulated statistics equal or exceed the observed value in 99 random permutations, the pseudo significance level would be $(4 + 1)/(99 + 1) = 0.05$. For a two-tailed, or non-directional, alternative hypothesis the probability in (4) would have to be multiplied by 2 to obtain the correct pseudo significance level.

The Moran's I statistic can be graphically depicted as the slope of a linear regression fit in a so-called Moran scatterplot (Anselin 1996). This graph uses standardized values of the variable of interest on the x-axis and its spatial lag on the y-axis. The spatial lag consists of a weighted average of the values at neighboring locations (see also Section 8 for further discussion of the spatial lag).

An illustration of Moran's I using OpenGeoDa (Anselin et al. 2006) is displayed in Figure 8.1 where the variable of interest is sudden infant death rates for 1979 (SIDR) in 100 North Carolina counties. The spatial weights matrix is based on contiguity between the counties. Depicted are three graphs. At the top of the figure is a choropleth map for the quintiles of the rates, with the darkest shade corresponding to the highest rates. The data are sorted by magnitude and categorized by quintile, with each quintile corresponding to a different shade. To the left at the bottom of the Figure is a Moran scatter plot. The x-axis shows the SIDS rate in 79 and the y-axis shows its spatial lag, \mathbf{W}_SIDR79 . The slope of the linear fit to the scatter plot is 0.1666, as listed at the top of the graph. To the right is a histogram of 999 Moran's I values computed from random permutations of the data. The vertical bar to the right corresponds to the observed value I_{obs} . The pseudo significance level is found to be 0.008, as shown at the top left of the graph. Consequently, the null hypothesis of spatial randomness

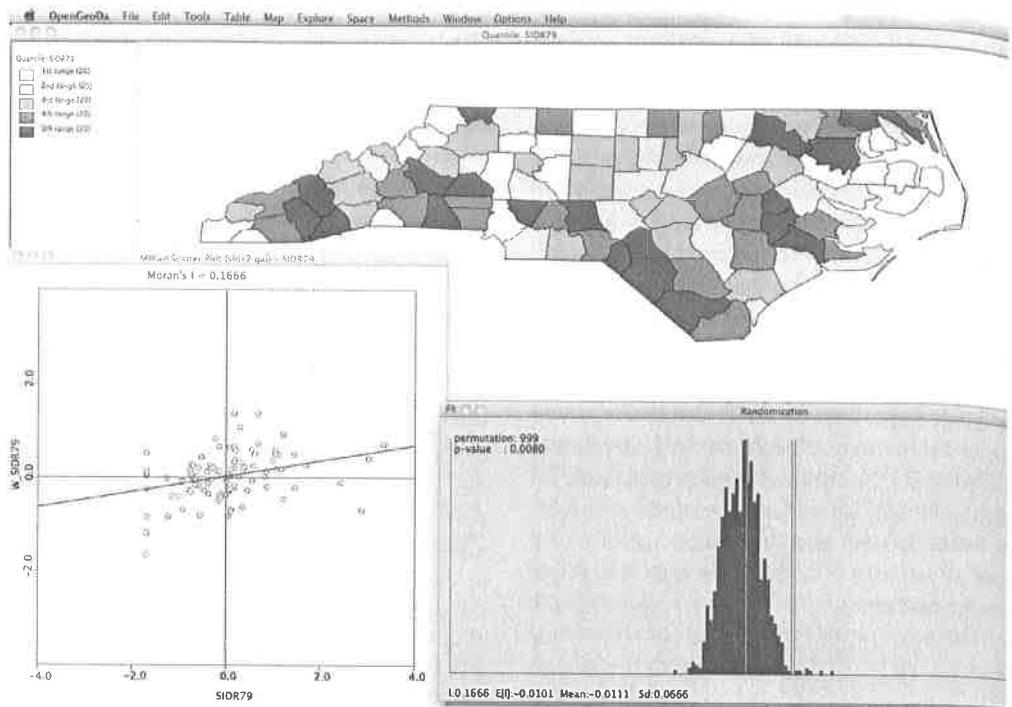


Figure 8.1 Moran's I for Sudden Infant Death Rates in North Carolina Counties (1979)

is rejected. Descriptive statistics for the empirical distribution of the statistics computed from random permutations are listed at the bottom of the graph. These include the analytical mean under the null (-0.0101), the mean of permuted values (-0.0111), and the standard deviation of these values (0.0666).

Moran's I is one of multiple global autocorrelation statistics. For an overview of alternative global autocorrelation statistics see O'Sullivan and Unwin (2003).

LOCAL AUTOCORRELATION

Local autocorrelation statistics are concerned with detecting departures from overall global patterns as well as identifying hot and cold spots, with the latter sometimes referred to as clusters. They do so by developing a measure for each spatial observation that expresses the amount of spatial autocorrelation associated with that observation. Thus, in contrast to the case of global statistics, where there is one value for the entire set of spatial observations, in the local case there are n such measures.

A particular class of local autocorrelation statistics is known as *Local Indicators of Spatial Association* or

LISA, suggested by Anselin (1995). To be considered a LISA a local statistic L_i must satisfy two properties:

1. The global autocorrelation statistic must be a function of the local statistics:

$$\sum_{i=1}^n L_i = \phi \Gamma \quad (6)$$

where Γ is the global autocorrelation statistic and ϕ is a scale factor.

2. It is possible to determine the statistical significance of the pattern of spatial association at individual locations $i = 1, 2, \dots, n$:

$$\text{Prob}[L_i > \delta_i] \leq \alpha_i \quad (7)$$

where δ_i is a critical value and α_i is a significance level.

One example of a LISA is the local counterpart to Moran's I:

$$I_i = z_i \sum_{j=1}^n w_{ij} z_j, \quad (8)$$

where the z variable is the same deviation from the mean as in equation 1 and the w_{ij} are the spatial weights.

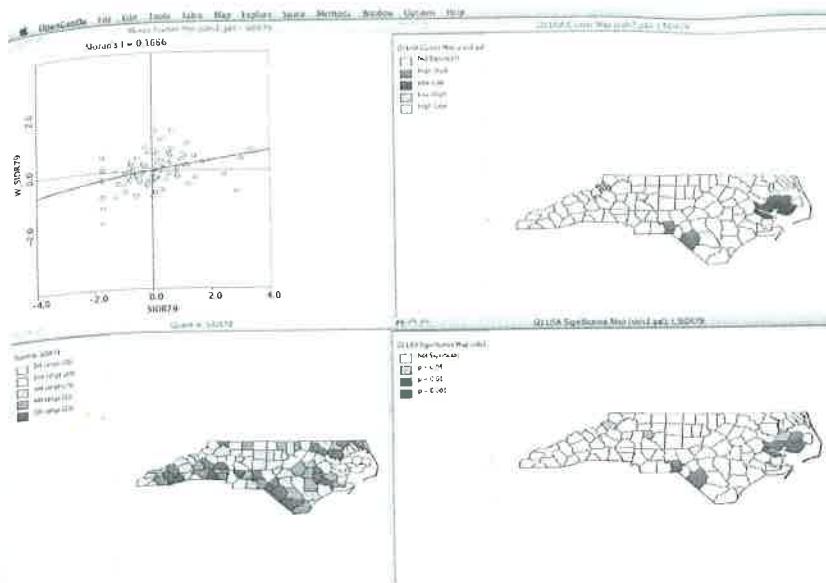


Figure 8.2 Local Statistics for Sudden Infant Death Rates in North Carolina Counties (1979)

Local statistics have seen wide application in recent years, in part driven by their implementation in commercial GIS software and their ability to provide location specific information about spatial autocorrelation which can be then visualized on a map. An illustration using the same North Carolina data as in Figure 8.1 is shown in Figure 8.2. The left part of the figure consists of the Moran scatter plot and the quintile map shown previously. The right part of the figure contains two maps. The bottom map highlights the counties for which the local Moran statistic is significant, with the shading corresponding to the three pseudo p-values of 0.05, 0.01 and 0.001. A darker shading indicates a lower p-value (greater significance). Significance as such does not provide any information as to what type of local spatial association is present (positive or negative). This is obtained by identifying with which quadrant of the Moran scatter plot the significant observations correspond. This allows for the classifications of locations as clusters (high surrounded by high, or low surrounded by low) and spatial outliers (high surrounded by low, or low surrounded by high). A visual representation of the combined classification and significance is illustrated in the top map on the right, a so-called cluster map. Here, the same significant locations as in the bottom map are classified by type, illustrating the presence of high and low clusters as well as some spatial outliers.

Figure 8.2 also illustrates the interactive functionality in OpenGeoDa. The cold-spot county in the northwestern part of the state has been selected on the cluster map, reflected by a cross-hatching in the figure. In turn, the observations associated with this county are then highlighted in the three other views. On the maps, the matching county is also cross-hatched and on the Moran scatter plot the point corresponding to this county is highlighted. This *linking* is one of multiple forms of dynamic graphics that provide powerful mechanisms for the user to explore different dimensions of their data.

While local statistics enjoy much popularity, there are some complications in their interpretation and use. Analytical results for the sampling distribution of the statistics are generally unavailable and, as a result, inference is often based on a conditional randomization of the values surrounding each observation. Given that there are now n tests being carried out the issue of multiple comparisons becomes relevant and several adjustments to the critical values and marginal significance have been suggested in the literature. The n tests will also be dependent since neighboring LISAs will utilize a common subset of observations in their construction. Further discussions of local statistics can be found in de Castro and Singer (2006), Anselin (1995) and Ord and Getis (1995).

Spatial Clustering

As mentioned above, one of the applications of LISAs is to identify clusters within a map pattern. In this regard, the LISA can be seen as a special case of the more general problem of spatial clustering, which is broadly concerned with grouping spatial observations together in such a way that the internal group variance is minimized while the inter-group variance is maximized. Spatial clustering is a large literature with overviews available in Haining (2003, p. 251–265), Lawson and Denison (2002) and Murray and Estivill-Castro (1998).

The variety of spatial clustering methods can be organized into three classes: [1] clustering which is exhaustive and mutually exclusive; [2] clustering which is not exhaustive but is mutually exclusive; and [3] there is a focal point around which one is interested in determining if there is a cluster.

In the first two classes, the number of clusters may or may not be specified *a priori*. If not, then a common strategy is to run a clustering algorithm for different numbers of clusters and select the solution that performs best. Examples of the first class arise in studies of urban neighborhoods and geodemographics (Harris et al. 2005) where there is a need to define a complete partitioning of primitive spatial units such as census blocks or tracts into homogeneous and spatially contiguous neighborhoods or market segments. The second class of clustering methods is widely applied in the analysis of disease and crime rate patterns where the focus is on identifying areas where the rates are elevated. In these cases only a subset of the spatial observations are assigned into clusters. The third class of clustering problems concerns a focal point of interest, such as a noxious facility, so the core of a potential cluster is specified *a priori* which is not the case for the first two sets of clustering methods. The statistical methods to assess whether there is a cluster around the focal point are, however, similar to those applied in the first two cases with unknown cluster locations and generally compare the observed number of events falling within the potential cluster boundary (typically a circle) to what the expected number of such events should be assuming the data generating process is characterized by complete spatial randomness. It is important to adjust the expected counts for spatial variation in the underlying population at risk – so for example in the study of disease cases, spurious clusters, due to population concentrations and uniform risks, are not detected and instead only clusters displaying truly elevated risks are identified.

EXTENSIONS TO SPACE-TIME

With the growing use of geospatial technologies such as global positioning systems there is an increasing amount of data that is not only spatial but also includes a time dimension. An active area of research within ESDA is the development of methods that incorporate this time dimension in a number of different ways. One branch of this research is developing new statistical measures that are designed to characterize the overall spatial dynamics, that is the role of spatial clustering in the evolution of a value distribution over time (Rey 2001, Rey and Anselin 2007). For example, in the literature on regional income convergence and divergence (Rey and Le Gallo 2009) interest has centered on the identification of so called poverty traps or growth clubs consisting of geographically clustered regional economies that display distinct collective income growth patterns over time.

Coupled with these new space-time statistics are a collection of interactive and dynamic graphics that extend the brushing and linking capabilities seen above in the case of OpenGeoDa, to include the temporal dimension. Representative examples of software packages implementing these views are STARS (Rey and Janikas 2006) and CommonGIS (Andrienko and Andrienko 2006).

We illustrate some of these concepts in Figures 8.3 and 8.4, using the STARS software. The variable of interest is the evolution of regional income for states in Mexico over the period 1940–2000. Just like OpenGeoDa, STARS implements the concepts of linking and brushing, connecting all the different views of the data. Figure 8.3 illustrates the situation at the beginning of the period. It contains four graphs. At the top left is a box plot that shows the relative distribution of per capita regional incomes. The three points above the upper bar are outliers, they are more than 1.5 times the interquartile range above the 75-percentile, using the standard approach in EDA. To the right on top is a quintile map of the same variable for the Mexican states. To the left on the bottom is a Moran scatter plot, with the spatial lag of the regional income on the vertical axis and the income on the horizontal axis. This Moran scatter plot pertains to the data for 1940. In the plot on the bottom right of the figure a time series plot is drawn that shows the evolution of the Moran's *I* (the slope in the scatter plot on the left) at the different points in time. The vertical axis represents the magnitude of Moran's *I* with the time periods on the horizontal axis. In the dynamic version of the graph, the visualization moves through

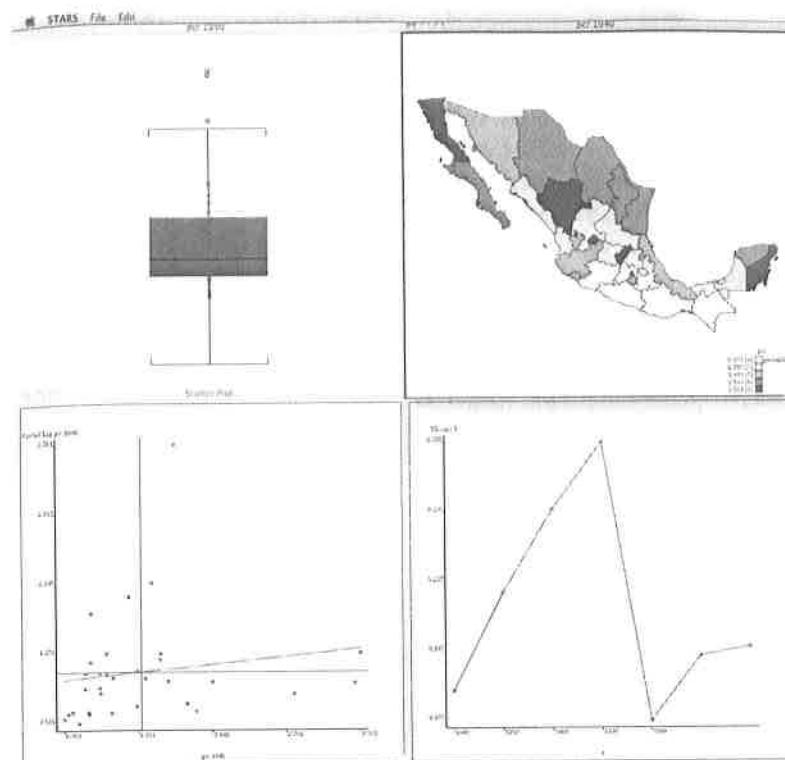


Figure 8.3 Exploratory space-time analysis of Mexican state incomes 1940–2000 using STARS

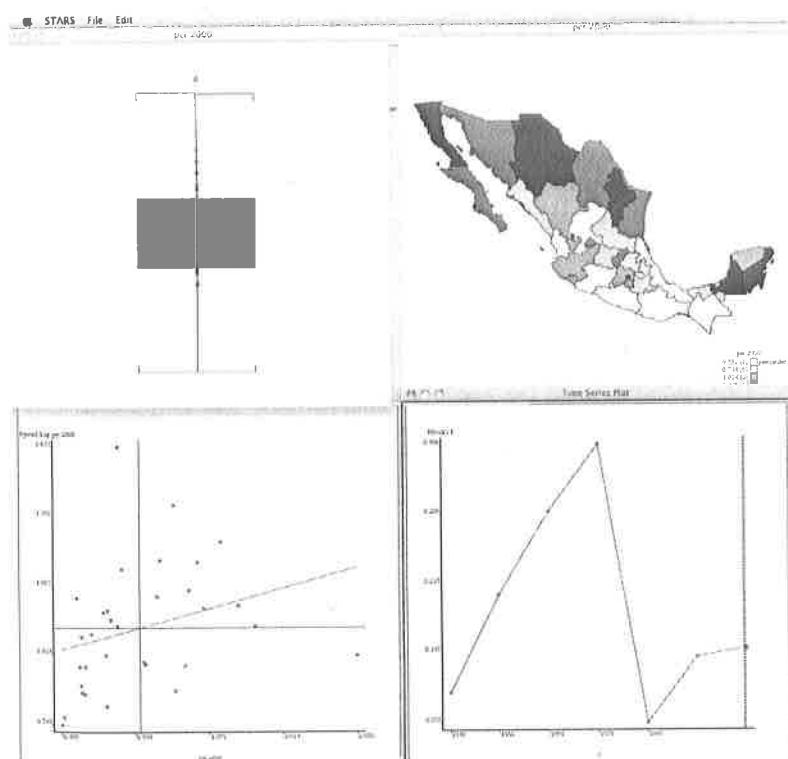


Figure 8.4 Time-sliding as a form of linking

each year in turn, which changes the box plot, the quintile map and the Moran scatter plot. The corresponding Moran's I is highlighted as a point on the time series plot. Figure 8.4 shows the situation at the end of the period, in 2000. The time series plot (bottom right) is unchanged, but it now shows a vertical bar at the year 2000. The other three graphs are different, illustrating different outliers (only one state is an outlier in 2000), a different spatial distribution (in the map) and a different Moran's I (the slope of the Moran scatter plot). The software allows the user to move the vertical bar, a so-called time slider, interactively over the different time periods. This time sliding allows for the exploration of the dynamics of spatial autocorrelation, since the graphs are updated for each time period.

Extending ESDA to include a temporal dimension opens up the scope of analysis to a wide array of data types. In the case of STARS above, the spatial units are fixed over time while their attribute values change over time. In other cases of space-time data, the location of events may be changing. For example, Figure 8.5 illustrates the origin-destination pattern of convicted sex offenders in relation to the location of schools and the associated spatial restriction zones (Murray et al. 2012). Each arrow shows the residence of the offender at the beginning and at the end of the time period under consideration. The schools are the blue areas surrounded by their spatial restriction zone. The complex patterns generated by such movement data necessitate the development of both new interactive visualization devices as well as exploratory statistics in order to identify interesting structures within this multidimensional data. This remains a very active area of ongoing research.

Spatial Regression Analysis

Spatial regression deals with the specification, estimation, diagnostic checking and prediction of regression models that incorporate spatial effects (e.g., Anselin 2006). In the social sciences literature, it is also often referred to as *spatial econometrics*, a term coined in the early 1970s by the Belgian economist Jean Paelinck (Paelinck and Klaassen 1979). In a regression context, two broad classes of spatial effects may be distinguished, referred to as spatial dependence and spatial heterogeneity (Anselin 1988). In this section, we will focus on how these spatial effects affect regression analysis, with a particular emphasis on the linear regression model, which is the most frequently used in practice.

Early interest in the statistical implications of estimating spatial regression models dates back to the pioneering results of the statisticians Whittle (1954), Besag (1974), Ord (1975) and Ripley (1981). By the late 1980s and early 1990s, several compilations had appeared that included technical reviews of a range of models, estimation methods and diagnostic tests, including Anselin (1988), Griffith (1988) and Haining (1990). Most importantly, with the publication of the text by Cressie (1991), a near-comprehensive technical treatment of the statistical foundations for the analysis of spatial data was provided.

Spatial regression analysis is a core aspect of the *spatial* methodological toolbox and several recent texts covering the state of the art have appeared, including Haining (2003), Waller and Gotway (2004), Banerjee et al. (2004), Fortin and Dale (2005), Schabenberger and Gotway (2005), Arbia (2006), Pfeiffer et al. (2008), Lawson (2009), and LeSage and Pace (2009). Recent surveys of more

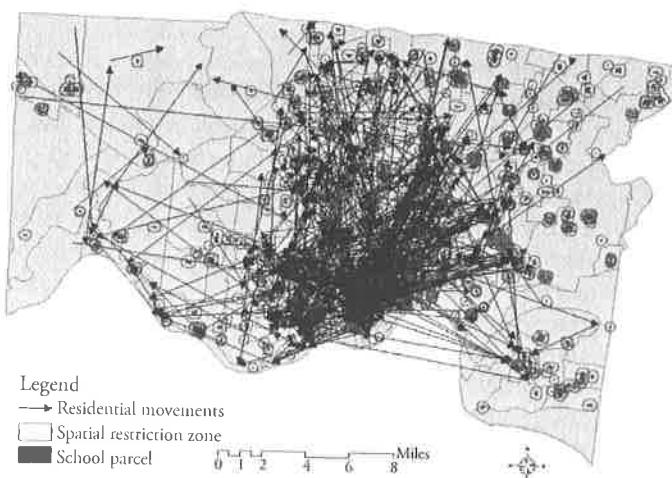


Figure 8.5 Exploratory visualization of residential movements of sex offenders

advanced methodological issues in spatial regression analysis (and spatial econometrics) can be found in Anselin and Bera (1998) and Anselin (2001, 2006, 2010).

In the remainder of this section, we will primarily focus on the various specifications through which spatial effects can be introduced into a regression model. We start with a discussion of the overall problem and then move to spatial dependence and spatial heterogeneity in linear regression models. We close with a brief review of three other spatial models, i.e., specifications of spatial panel data models, spatial latent variable models and Bayesian hierarchical spatial models. Where appropriate, we will address the three other aspects of spatial regression (estimation, specification tests and prediction) with selected references. Due to the highly technical nature of these topics, we will limit ourselves to a review of the main principles and focus primarily on model specification. We refer the interested reader to the review articles cited above and the materials cited in them for technical details and a more in-depth treatment.

Spatial Effects in Regression Specification

In the context of regression analysis, spatial dependence is viewed as a special case of cross-sectional dependence, in the sense that the *structure* of the correlation or covariance between random variables at different locations is derived from a specific ordering, determined by the relative position (distance, spatial arrangement) of the observations in geographic space (or, in general, in network space). While similar to correlation in the time domain, the distinct nature of spatial dependence requires a specialized set of techniques. Importantly, these are *not* a straightforward extension of time series methods to two dimensions.

Spatial heterogeneity is a special case of observed or unobserved heterogeneity, a familiar problem in standard econometrics. In contrast to spatial dependence, tackling spatial heterogeneity does not always require a separate set of methods. The only spatial aspect of the heterogeneity is the additional information that may be provided by spatial structure. For example, the information on spatial structure may inform models for heteroscedasticity, spatially varying coefficients, random coefficients and spatial structural change.

Spatial heterogeneity becomes particularly challenging since it is often difficult to separate from spatial dependence. This difficulty is known in the literature as the inverse problem. It is also related to

the impossible distinction between true and apparent contagion. The essence of the problem is that cross-sectional data, while allowing the identification of clusters and patterns, do not provide sufficient information to identify the *processes* that led to the patterns. As a result, it is impossible to distinguish between the case where the cluster is due to structural change (apparent contagion) or follows from a true contagious process. This problem is specific to a pure cross-sectional setting and can be remedied by resorting to observations across space and over time. In a regression context, models for such pooled cross-section and time series data are referred to as spatial panel models.

Spatial Dependence in the Linear Regression Model

The point of departure in our discussion of model specification is the standard linear regression model. To fix notation, we consider, for each observation (location) i , the following linear relationship between a dependent variable y and k explanatory variables x_h :

$$y_i = \sum_{h=1}^k x_{hi} \beta_h + \epsilon_i, \quad (9)$$

where the β_h are the associated regression coefficients and ϵ_i is a random error term.

Spatial dependence is introduced into the regression specification in two fundamentally different ways. In the first, the dependence is conceptualized as following from an interaction process between the observational units. Examples of such interaction processes are externalities, copy catting, peer-effects, etc. This interaction corresponds to the notion of *substantive* spatial dependence introduced in Section 8. In essence, the dependent variable at one location (observation) is specified as a function of its value at neighboring locations. For the sake of simplicity, we consider the linear case only.

The effect of the *neighbors* is encapsulated in a so-called spatially lagged dependent variable, which we will designate as Wy . Technically, an observation on the spatially lagged dependent variable, or spatial lag, is obtained as a weighted average of neighboring values, with the weights specified in a spatial weights matrix, similar to what is necessary for a spatial autocorrelation coefficient (see Section 8). For location i , this becomes:

$$Wy_i = \sum_{j=1}^n w_{ij} y_j \quad (10)$$

with the w_{ij} as the spatial weights. Since the spatial weights are typically row standardized (such that $\sum_j w_{ij} = 1$), the weighted sum in Equation 10 boils down to an averaging of the value of y for the neighboring locations. The operation of creating a new variable as a weighted average of neighboring values is called a *spatial lag operation*. It can be performed on the dependent variable, the explanatory variables, or the error term (for details, see Anselin 2003).

A regression specification that includes a spatially lagged dependent variable is referred to as a mixed regressive, spatial autoregressive model, or *spatial lag* model in short. Conceptually, it is the expression of the equilibrium outcome of a process of social and spatial interaction, although in practice it is often used as a mechanism to filter the dependent variable of the effect of spatial correlation. Formally, the spatial lag model is then:

$$y_i = \rho \sum_{j=1}^n w_{ij} y_j + \sum_{h=1}^k x_{hi} \beta_h + \epsilon_i, \quad (11)$$

where ρ is the spatial autoregressive coefficient, i.e., the regression coefficient associated with the spatially lagged dependent variable $\sum_{j=1}^n w_{ij} y_j$. The rest of Equation 11 is the familiar linear regression specification.

The inclusion of the spatial lag is similar to an autoregressive term in a time series context, hence it is called a spatial autoregressive model, although there is a fundamental difference. Unlike time dependence, dependence in space is multidirectional, implying feedback effects and simultaneity. More precisely, if i and j are neighboring locations, then y_j enters on the right hand side in the equation for y_i , but y_i also enters on the right hand side in the equation for y_j (the neighbor relation is symmetric). The endogeneity implied by this feedback must be accounted for in the estimation process and is qualitatively different from the one-directional dependence in time series.

An important aspect of the spatial lag model is the concept of a spatial multiplier (for details, see Anselin 2003). The multiplier follows from the solution of the model in which all the dependent variables are removed from the right hand side of the equation, the so-called *reduced form*. Using matrix notation, the spatial lag model becomes:

$$y = \rho W y + X \beta + \epsilon, \quad (12)$$

where y is now a n by 1 vector of observations on the dependent variables, W is a n by n spatial weights matrix, X is a n by k matrix of observations on the

explanatory variables, ϵ is a n by 1 vector of error terms, and the coefficients are as before. Solving Equation 12 by means of a matrix inverse operation yields:

$$y = (I - \rho W)^{-1} X \beta + (I - \rho W)^{-1} \epsilon, \quad (13)$$

with I as an n by n identity matrix. Equation 13 reveals the spatial multiplier, in the sense that the value of y at any location i is not only determined by the values of x at i , but also of x at all other locations in the system. A simple expansion of the inverse matrix term (for $|\rho| < 1$ and with a row-standardized W), and using the expected value (since the errors all have mean zero) further reveals the structure of the multiplier:

$$E[y|X] = X\beta + \rho W X \beta + \rho^2 W^2 X \beta + \dots \quad (14)$$

The powers of ρ matching the powers of the weights matrix (higher orders of neighbors) ensure that a distance decay effect is present.

In the second class of spatial regression model, the spatial dependence does not enter into the substantive part of the regression specification, but affects the covariance structure of the random error terms. The typical motivation for a *spatial error* specification is that there is a mismatch between the scale of observation and the scale at which the phenomenon under study manifests itself. The mismatch implies that neighboring locations share unobserved effects (sometimes called common factors) which results in non-zero off-diagonal elements in the error variance-covariance matrix. The non-spherical error variance-covariance is the expression of *nuisance* spatial autocorrelation (see Section 8).

Spatial error autocorrelation is thus a special case of a non-spherical error covariance matrix, i.e., where $E[\epsilon_i \epsilon_j] \neq 0$, for $i \neq j$, or, in matrix notation, with $E[\epsilon \epsilon'] = \Sigma$. The value and pattern of the non-zero covariances are the outcome of a spatial ordering. In a cross-section, it is impossible to extract this ordering from the data directly, since there are potentially $[n \times (n - 1)]/2$ covariance parameters and only n observations to estimate them from. Hence, it is necessary to impose structure.

The spatial covariance structure can be obtained in a number of ways, yielding a wide array of specifications. One of the earliest suggestions was to express the covariance terms as a function of the distance between the observations in question. In this so-called *direct representation*, the covariance between error terms is a function of distance, $E[\epsilon_i \epsilon_j] = \sigma^2 f(d_{ij}, \phi)$, with f as a *proper* function, such that the resulting variance-covariance matrix

is positive definite. This approach is most common in the so-called geostatistical literature and provides a way to improve on the precision of spatial prediction, so-called kriging (see Schabenberger and Gotway 2005, for an extensive treatment).

Arguably the most commonly used approach to specify spatial error dependence is to select a spatial stochastic process model for the error term. Typical choices include a spatial autoregressive form (similar to the spatial lag model, but expressed for the error term) and a spatial moving average form. A special case of the spatial autoregressive specification is the conditional autoregressive or CAR model, a common choice as prior in a Bayesian spatial hierarchical model (for reviews, see, e.g. Banerjee et al. 2004).

Specification of Spatial Heterogeneity

A good starting point to discuss the complexities introduced by spatial heterogeneity is to contrast complete homogeneity and extreme structural instability. Under homogeneity, the standard linear regression specification is fixed across observations, the unknown parameter values are constant and the error terms are independent and identically distributed (*i.i.d.*). Formally, for each observation i , this is expressed as in equation (9) above. In contrast, in the case of extreme heterogeneity, for each observation, there is potentially a different functional form, encompassing the situation of different parameter values and/or different explanatory variables, and the error is not independent and not identically distributed (*n.i.n.i.d.*). Formally, extreme heterogeneity can be expressed as:

$$y_i = f_i \left(\sum_{b=1}^{K_i} x_{ib} \beta_{ib} \right) + \epsilon_i, \quad (15)$$

where a different functional form f_i pertains to each observations, with a different set of regression coefficients β_{ib} and an error term with a different distribution for each i . This expression suffers from the *incidental parameter* problem, in that there is a different set of parameters for each observation. It is therefore not operational. The number of coefficients to be estimated increases with the number of observations, such that the sample never provides sufficient information to obtain reliable estimates. We solve this problem by imposing structure of a *spatial* form, hence the term spatial heterogeneity.

There are two main approaches to imposing structure. In one, the instability is categorized into a small number of subsets within which the specification is stable. This approach is referred to as

discrete spatial heterogeneity. The best-known form of this approach is the inclusion of so-called spatial fixed effects in the regression specification. Spatial fixed effects are indicator variables that correspond to a spatial subset of observations, such as all counties within the same state, or blocks within the same census tract. The estimated coefficients of the spatial fixed effects indicate the extent to which individual subregions deviate from the common mean. The notion of a fixed effect applied to the regression intercept can be readily generalized to all the regression coefficients, in a so-called *spatial regimes* model (Anselin 1990). The regression is specified such that different subregions have different coefficient sets, which allows one to test the null hypothesis of regional homogeneity.

The second approach specifies the structural instability in the form of a smooth or continuous variation of the model coefficients, as a special case of varying coefficients. We refer to this approach as *continuous spatial heterogeneity*. An early example of this in the quantitative geography literature is the so-called *spatial expansion* method of Casetti (1972, 1997). In its initial form, the expansion method consisted of fitting a spatial trend surface (a polynomial regression in the coordinates of the observations) to each of the regression coefficients, allowing one to map the *spatial drift* of the estimates. Later, the trend surface was generalized to include any type of expansion variable. In the statistical literature, spatially varying coefficients are viewed as a special case of models in which the coefficients are allowed to vary as smooth functions of other variables (for a general discussion, see Hastie and Tibshirani 1993, among others).

A local form of this principle is reflected in the *geographically weighted regression* or GWR (Fotheringham et al. 1998, 2002). In a GWR, each model coefficient is estimated at each location as a locally weighted (kernel) estimate, using geographically nearby observations as the support. The collection of local coefficients also provides the opportunity to map the spatial variability of each coefficient. The principle behind GWR has been applied to many other estimation contexts besides the linear regression model and continues to be a subject of active research. For example, a recent comparison of the performance of GWR and spatially varying coefficient models in an empirical context is given in Waller et al. (2007).

In spatially varying coefficient models such as the GWR, the continuous heterogeneity is expressed as a function. An alternative that has wide application

in Bayesian hierarchical spatial modeling is to specify the variability in the form of a (prior) distribution for the model parameters. This distribution can itself encompass a spatial stochastic process structure, which allows for the spatial variability to be expressed in terms of a small number of parameters. Examples of models with random spatial parameter variation are given in Gelfand et al. (2003) and Assunçao (2003).

Other Spatial Models

The range of possible spatial regression specifications is much greater than the linear spatial lag and spatial error models considered so far. For example, each of these models can be taken in combination with spatial heterogeneity, such as spatial regimes or heteroskedasticity. More fundamentally, there are three broad dimensions in which the spatial specifications can be extended. One consists of introducing spatial dependence and heterogeneity into models that combine observations across space and over time, in so-called spatial panel data models (for recent overviews, see, e.g., Elhorst 2003, Anselin et al. 2008, Lee and Yu 2010). A second considers spatial effects in models with latent variables, such as discrete choice models, which includes as a special case the spatial probit and tobit models (see Fleming 2004). A third direction pertains to the very large volume of Bayesian spatial hierarchical models, which have seen wide application in modern statistics (e.g., Banerjee et al. 2004). In addition to these three broad areas, there is an increasing interest in semi-parametric spatial models (e.g., Pinkse et al. 2002, Basile and Gress 2005, Pinkse and Slade 2010). However, a detailed discussion of semi-parametric models is beyond the current scope.

We next give a brief overview of each class of models (due to space constraints, we direct the interested reader to the references cited for additional examples and technical details). It should be noted that each of these areas are still undergoing very active research and the state of the art is rapidly moving forward.

Spatial Panel Regression Models

Spatial dependence of the error or lag form can be introduced into a standard panel model specification in a straightforward manner. The point of departure is the model:

$$y_{it} = \mathbf{x}'_{it}\beta + \epsilon_{it}, \quad (16)$$

where i is an index for the cross-sectional dimension, with $i = 1, \dots, n$, and t is an index for the

time dimension, with $t = 1, \dots, T$. Using customary notation, y_{it} is an observation on the dependent variable at i and t , \mathbf{x}_{it} a $k \times 1$ vector of observations on the (exogenous) explanatory variables, β a matching $k \times 1$ vector of regression coefficients, and ϵ_{it} an error term. The setting considered here is where the cross-sectional dimension dominates, with $n \gg T$. Also, even though the basic design is referred to as "space" and "time," the second dimension could equally pertain to different cross-sections, such as in a study of industrial sectors or household types. In stacked matrix form, the simple pooled regression then becomes:

$$\mathbf{y} = \mathbf{X}\beta + \epsilon, \quad (17)$$

with \mathbf{y} as a $nT \times 1$ vector, \mathbf{X} as a $nT \times k$ matrix and ϵ as a $nT \times 1$ vector. Note that in order to incorporate spatial effects, the stacking is for a complete cross-section at a time, and not for each individual cross-section over time.

The key to incorporating spatial dependence into this specification is the use of a spatial weights matrix for the panel dimension, by creating a block diagonal $nT \times nT$ matrix with the n -dimensional cross-sectional weights as the diagonal elements, or:

$$\mathbf{W}_{nT} = \mathbf{I}_T \otimes \mathbf{W}_n, \quad (18)$$

where \mathbf{I} is an identity matrix and the subscripts refer to the matrix dimension, with \otimes as the Kronecker product.

A spatial lag model can then be expressed as:

$$\mathbf{y} = \rho(\mathbf{I}_T \otimes \mathbf{W}_n)\mathbf{y} + \mathbf{X}\beta + \epsilon, \quad (19)$$

where ρ is the spatial autoregressive parameter (constant over the time dimension), and the other notation is as before. Similarly, a model with spatial SAR error dependence results in an $nT \times nT$ non-spherical error variance-covariance matrix of the form:

$$\Sigma_{nT} = \sigma_u^2 [\mathbf{I}_T \otimes [(\mathbf{I}_n - \lambda \mathbf{W}_n)'(\mathbf{I}_n - \lambda \mathbf{W}_n)]^{-1}], \quad (20)$$

where σ_u^2 is a common variance term, and the spatial autoregressive coefficient λ is assumed to be constant over the time dimension. More complex model specifications are reviewed in Anselin et al. (2008).

One class of space-time models has received considerable attention in spatial analysis, particularly in applied economics. In so-called error component models the spatial (and time) dependence is introduced into a classic two-way error component specification (e.g., Baltagi 2001, p. 31). In this model, each error term is decomposed into three terms:

$$\epsilon_{it} = \alpha_i + \phi_t + u_{it}, \quad (21)$$

where α_i is a cross-sectional unobserved random effect, ϕ_i is a random time effect, and u_{it} is an idiosyncratic component. Spatial dependence can be introduced in the form of a spatial autoregressive process for the idiosyncratic component, but a number of other specifications have been suggested as well (for reviews, see Anselin 1988, Baltagi et al. 2003, 2006, Kapoor et al. 2007). The error component specification is also one of the main ways in which spatial dependence is incorporated in a Bayesian approach (see Section 8).

SPATIAL LATENT VARIABLES

Spatial effects in latent variables models are particularly relevant in applied micro-econometrics, where the observed dependent variable often only takes on discrete values (e.g., a binary dependent variable that takes on values of 0 and 1). Incorporating spatial dependence into these specifications is not straightforward. One common approach (especially in the statistics literature) is to take a Bayesian perspective (see Section 8). Another approach is to specify the dependence for an unobserved latent variable, say y_i^* , which is defined as a linear function of an "index function" and a random error term:

$$y_i^* = \mathbf{x}'_i \beta + \epsilon_i, \quad (22)$$

with $\mathbf{x}'_i \beta$ as the index function, where \mathbf{x}_i is a $k \times 1$ vector of observations on the explanatory variables, and β is a matching vector of coefficients. The observed counterpart of y_i^* , the discrete dependent variable y_i , equals one for $y_i^* > 0$ and zero otherwise. Interest therefore centers on the probability of observing an event, i.e., $P[y_i^* > 0] = P[\mathbf{x}'_i \beta + \epsilon_i > 0]$. By specifying a distribution for the random error term, estimates for β can be obtained. In order to incorporate spatial dependence, the multivariate Gaussian distribution is particularly attractive, which leads to the so-called *spatial probit* model.

The key to modeling spatial dependence in this context is not to express the dependence for the observed variable y_i , but instead to use a spatial model for the latent variable in expression (22). For example, a spatial lag model would be:

$$y_i^* = \rho \sum_j w_{ij} y_j^* + \mathbf{x}'_i \beta + \epsilon_i, \quad (23)$$

or, in matrix notation, using the familiar reduced form:

$$\mathbf{y}^* = (\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{X} \beta + (\mathbf{I} - \rho \mathbf{W})^{-1} \boldsymbol{\epsilon}. \quad (24)$$

In this *simultaneous* model, the latent variables are jointly determined, both by the values for x at their

own location and by the values for x at all other locations, subject to a distance decay effect. One immediate consequence of the simultaneity is that the usual marginal condition $P[y_i^* > 0]$ now pertains to the marginal probability of a (complex) *multivariate* normal distribution with a variance-covariance matrix that includes a spatial parameter. A similar complication occurs for the spatial error specification for a latent variable. As a result, there is no longer an analytical solution to the estimation problem and one typically has to resort to simulation estimators. Overviews of some of the technical issues can be found in Pinkse and Slade (1998), LeSage (2000), Kelejian and Prucha (2001), Fleming (2004), and Beron and Vijverberg (2004), among others.

BAYESIAN HIERARCHICAL SPATIAL MODELS

The application of the Bayesian perspective to spatial regression modeling has seen tremendous growth since the advent of readily available simulation estimators, especially Markov Chain Monte Carlo (MCMC) methods, such as the Gibbs sampler and the Metropolis-Hastings algorithm (see, e.g., Geman and Geman 1984, Gilks et al. 1996). Early applications were severely restricted in terms of the choice of conjugate prior distributions to ensure a proper posterior distribution and an analytical solution to the estimation problem. However, with simulation estimators, increasingly complex specifications can now be analyzed, including complex space-time correlation structures (for a recent review, see Banerjee et al. 2004).

We can make a distinction between Bayesian approaches to the spatial lag and spatial error specifications (as exemplified in the work of LeSage, summarized in LeSage 1997, LeSage and Pace 2009) and models where the spatial effects are introduced in a hierarchical fashion. We focus on the latter.

A major area of application of hierarchical spatial models is the analysis of rates or events as an approach to estimate the underlying risk surface, such as in epidemiology, public health and criminology (Best et al. 1999). The basic principle behind the hierarchical model can be illustrated with a specification for counts of events as a realization of a Poisson distribution with a heterogeneous mean, $y_i \sim Poi(\mu_i)$, where *Poi* stands for the Poisson distribution and μ_i is a location-specific mean. Standard practice in epidemiology is to express the mean as the product of the "expected count" and the relative risk. The expected count is typically based on a form of standardization and the main interest

focuses on the relative risk, i.e., to what extent does the risk at individual locations differ from what should be expected. The model specification thus becomes $y_i|\eta_i \sim Poi(E_i\eta_i)$, with E_i as the expected count and η_i as the relative risk, the parameter of interest.

The hierarchical aspect of this model appears as a random coefficient specification in the form of a distribution for the parameter η_i . Typically, a Gamma distribution is taken as the distribution, itself a function of two parameters, for which prior distributions need to be specified. This is referred to as the Poisson-Gamma model. While it is possible to include spatial effects into this model, this is typically quite complex (Wolpert and Ickstadt 1998).

A simpler approach, which allows for ready inclusion of spatial dependence and spatial heterogeneity, is to take the distribution for the log of the relative risk, $\phi_i = \log \eta_i$, to follow a Gaussian distribution. This approach is referred to as the Poisson-Lognormal model. The conditional distribution of the observed counts, conditional upon the log-relative risk is then $y_i \sim Poi(E_i e^{\phi_i})$. The hierarchical aspect comes in the form of a linear regression specification for the relative risk parameter:

$$\phi_i = \mathbf{x}'_i \beta + \theta_i + \psi_i, \quad (25)$$

where \mathbf{x}'_i is a vector of variables that explain the heterogeneity of the risk parameter, with an associated coefficient vector β , and θ_i and ψ_i are random components that follow a Gaussian distribution. The first of the random components, θ_i reflects the heterogeneity across space and often takes on a Gaussian prior with a given variance (or, preferably, the inverse of the variance, or precision). The second component incorporates spatial autocorrelation, either in the form of a direct representation (a multivariate Gaussian distribution with a variance-covariance that is a function of inverse distance) or of a CAR process. Estimation of the parameters of the model is carried out by means of MCMC. Both models have seen wide application (see Banerjee et al. 2004, Lawson 2009, for a more extensive discussion and many illustrations).

Bayesian hierarchical spatial modeling continues to be an area of active research.

Spatial Optimization Modeling

There are many ways in which optimization is important in spatial analysis. Murray (2007) discusses a range of contexts, spanning GIS database management, the arrangement and representation

of geographic space as well as spatial optimization in planning and decision making. Of course, much of the statistical methods detailed in this chapter rely on optimization to derive best or efficient model parameter estimates. In the remainder of this section we focus on optimization, and spatial optimization modeling more specifically.

Optimization

Optimization is a term that is widely used. Even in basic calculus one encounters functions that can be characterized in terms of local and global optima, with the challenge of identifying such *critical points* and interpreting their significance (see Miller 2000). Of course, local/global optima represent instances of the function that correspond to extreme values, with an economic interpretation of being the most efficient or most profitable when the function describes return on investment, as an example. Optimization is necessarily an area of applied mathematics, but the fields of operations research and management sciences have come to be synonymous with optimization because of their emphasis on identifying, developing and applying optimization based approaches. Therefore, the mathematical view of optimization is that it seeks to obtain values to variables (unknowns) that either maximize or minimize a function subject to constraining conditions (also functions).

A generic optimization problem can be stated as follows:

$$\text{Maximize} \quad f(\mathbf{x}) \quad (26)$$

$$\text{Subject to } g_i(\mathbf{x}) \leq \beta_i \forall i \quad (27)$$

$$\mathbf{x} \geq 0, \quad (28)$$

where $f(\mathbf{x})$ is a function, \mathbf{x} is an $n \times 1$ vector of decision variables, $g_i(\mathbf{x})$ is function i (m in total), and β_i is a coefficient specifying the bound or limit on the value of function i . The objective, (26), represents a function to be optimized. Note that it is also possible to minimize the objective by simply multiplying the associated function by -1 . The model constraints are specified in (27), where there are m functions, each bounded by an inequality condition. Finally, decision variable stipulations are given in (28). The idea then is that we need to make decisions on what the best values of the unknown variables should be in order to optimize the objective, but must not violate any of the imposed constraining conditions.

If there are no constraints (27), then from calculus it is theoretically possible to solve this optimization problem by taking the derivative and setting it equal to zero. Unfortunately, in practice it may not be possible to do this. One reason is that the derivative may not exist. Alternatively, it may not be possible to isolate individual variables. Finally, the derivative(s) may simply be too complex/difficult to solve. If there are constraints (27), optimally solving associated problems is all the more challenging. For these reasons, non-linear optimization problems, in general, remain difficult to solve, though many techniques exist for their solution, some available in commercial software.

Many important optimization models involve linear functions in the objective and the constraints. The generic optimization problem can therefore be restated as follows for linear functions:

$$\text{Maximize } \mathbf{c}\mathbf{x} \quad (29)$$

$$\text{Subject to } \mathbf{Ax} \leq \mathbf{b} \quad (30)$$

$$\mathbf{x} \geq \mathbf{0} \quad (31)$$

where \mathbf{c} is a $1 \times n$ vector of benefits, \mathbf{x} is an $n \times 1$ vector of decision variables, \mathbf{A} is a $m \times n$ matrix of constraint coefficients, and \mathbf{b} is a $m \times 1$ vector for right hand side limits. An optimization problem with linear functions is known as a *linear program*, and may be solved using linear programming based methods. There are many commercial software packages for solving linear programs. Further, it is possible to structure linear approximations for many non-linear problems.

Spatial Optimization

With the basics of optimization in hand, we can proceed to discussing its significance in spatial analysis. In general terms, *spatial optimization modeling* extends or applies optimization to a geographic context, focusing on situations where objectives and constraints are inherently defined by space and spatial relationships.

There are three different types of spatial optimization models that have been applied: location models; land use planning models; and, network design and protection models. Much of the current work in these areas makes extensive use of GIS as it enables access to spatial information, provides the capacity to understand, extract and structure spatial relationships, and facilitates the development and solution of model abstractions that reflect the increased reality and complexity faced by planners and decision makers. In the remainder of this section we detail

an example of a location model and a land use planning model, focusing on how space is structured mathematically.

LOCATION MODELS

Good overviews of location models can be found in Mirchandani and Francis (1990), Drezner and Hamacher (2002), and Church and Murray (2009). Weber (1909) was an early example of applying spatial optimization, where the interest was finding a location to site a factory in order to minimize transportation costs associated with both raw material inputs and delivery of finished goods to a market. Much location modeling work has followed associated with siting all sorts of public and private sector facilities and services, including libraries, fire stations, telecommunications infrastructure, distribution centers, emergency warning sirens, retail outlets, schools, oil/gas transmission corridors (see Church and Murray 2009). The range of applications and the variety of models are considerable. Given space limitations, it is not possible to review and/or discuss them all here. However, we will detail one model that reflects the significance of geography and spatial relationships in location models, and the inherent challenges of mathematically structuring this.

An important spatial optimization model involves siting a minimal number of facilities in order to ensure that all those areas utilizing the facility are within a maximum service or distance standard. Edmonds (1962) was among the first to discuss this optimization model, the *set covering* problem, but in the context of geographic space Toregas et al. (1971) formulated the problem to locate fire stations such that every area could be responded to within some maximum time (e.g., 10 minutes). Ando et al. (1998) used the same model to examine species protection through biological reserve siting. While the basic problem remains important and widely applied in practice, the emergence of GIS and more detailed spatial information has led the way to enhanced conceptualization of this problem, involving greater mathematical specification and structure. Murray (2005) formulated and solved a generalization of this problem to site emergency warning sirens in order to cover neighborhood areas (represented as polygons). Consider the following notation:

i = index of areas to be covered;

j = index of potential facility locations;

l = index of coverage levels ($1, 2, 3, \dots, L$);

δ_l = minimum acceptable coverage percentage at level l ;

Ω_{il} = set of potential facilities j partially covering area i at least δ_l ;

α_l = minimum number of facilities needed for complete coverage at level l ;

$$x_j = \begin{cases} 1 & \text{if a facility is sited at potential location } j \\ 0 & \text{otherwise} \end{cases}$$

$$y_{il} = \begin{cases} 1 & \text{if area } i \text{ is covered at level } l \\ 0 & \text{otherwise} \end{cases}$$

The basic coverage model suggested by Murray (2005), and detailed in Church and Murray (2009) as well as in Murray et al. (2010), is as follows:

$$\text{Minimize} \quad \sum_j x_j \quad (32)$$

Subject to

$$\sum_{j \in \Omega_{il}} x_j \geq \alpha_l y_{il} \quad \forall i, l \quad (33)$$

$$\sum_l y_{il} = 1 \quad \forall i \quad (34)$$

$$x_j = \{0, 1\} \quad \forall j \quad (35)$$

$$y_{il} = \{0, 1\} \quad \forall i, l$$

The objective, (32), is to minimize the number of facilities sited. Constraints (33) relate facility siting to coverage of demand area i , where coverage is not provided until at least α_l facilities are sited that serve the area. Constraints (34) require that coverage must be provided at some level k . Constraints (35) stipulate restrictions on decision variables.

There are two inherent spatial components to this particular model. First, there are the obvious geographic siting decisions, x_j , that indicate where facilities should be placed. Second, there is the tracking of coverage provided to areas by sited facilities, y_{il} . Of course these two components are intimately connected as the siting decisions dictate what will be covered. Beyond this spatial connection, what is significant about this coverage model is that it allows geographic objects in the form of points, lines and areas to be modeled. Thus, the work of Toregas et al. (1971), as an example, is a special case of this approach where the number of coverage levels is one ($L = 1$). In order to address line and area features (e.g., roads and neighborhoods, respectively), however, the model must recognize that complete

coverage of an area i may not be possible with only one facility, but rather will require multiple facilities in some cases. The index l therefore represents cases of multiple coverage when l is greater than one. As an example, if $l = 2$ the set Ω_{i2} accounts for those facilities capable of covering area i at least $\delta_2 = 60\%$, but less than 100%. If $\delta_2 = 60\%$, then this set contains all those potential facilities j that cover at least 60% of area i (but less than 100%). As a result, α_2 equal to two in constraint (33) would make sense, as this means that one would need at least two facilities covering area i at least 60% for it to actually be covered. The model therefore incorporates geographic siting decisions as well as tracks which areas are covered by what located facilities, where coverage of an area could be by a single facility, two facilities, three facilities, etc., up to L facilities. The ability both to spatially reference as well derive spatial relationships is critical for applying this coverage model. GIS facilitates both, enabling location models to be structured, as discussed in Murray (2010).

LAND USE PLANNING MODELS

A fundamental land use planning problem involves selecting land to acquire for some purpose. There are many contexts where land acquisition is necessary, such as waste disposal siting, building residential subdivisions, parks and recreation area designation, natural resource management, and establishing nature reserves to protect endangered species. In land acquisition there are many criteria that are important to consider. Among the most critical are often cost and suitability, but shape, area and proximity, among others, can be influential factors in what land should be acquired. The idea is to select land for use, development or preservation, taking into account budgetary and quality issues in the process.

There has been a considerable amount of research devoted to the development and application of a broad range of land acquisition models (see Thompson et al. 1973, Wright et al. 1983, Snyder et al. 1999, Williams 2002, Fischer and Church 2003, Zhang and Wright 2004, Shirabe 2005, Downs et al. 2008). While it is common in the literature to see references to a so called "land acquisition model," there is no one such model. Rather there are a variety of land acquisition models, just as there are numerous location models and many spatial optimization models. Wright et al. (1983) relied on perimeter to define shape in their developed mathematical model. The rationale behind this is that minimizing perimeter necessarily encourages

compactness (a circle is the most compact spatial object) and promotes contiguity, or an interconnectivity between acquired land. Others have relied on compactness and perimeter minimization in developed land use planning models, such as Nalle et al. (2002), Fischer and Church (2003), Onal and Briers (2003).

Given that a primary approach for addressing shape is through the use of perimeter, we now formulate a land use planning model based on the work of Wright et al. (1983) that was detailed in Wu and Murray (2007) as well as Church and Murray (2009). Consider the following notation:

i = index of land parcels;

b_i = benefit of acquiring land parcel i ;

c_i = cost of acquiring land parcel i ;

Φ_i = set of parcels that are adjacent to parcel i ;

p_{ij} = edge length between adjacent parcels i and j ;

Ψ = set of parcels on region boundary;

ω = importance weight for perimeter minimization;

μ = acquisition budget;

$$x_i = \begin{cases} 1 & \text{if parcel } i \text{ is acquired for a particular} \\ & \text{land use} \\ 0 & \text{otherwise} \end{cases}$$

$$e_{ij}^+ = \begin{cases} 1 & \text{if } x_i = 1 \text{ and } x_j = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$e_{ij}^- = \begin{cases} 1 & \text{if } x_i = 0 \text{ and } x_j = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} \text{Maximize} & (1 - \omega) \sum_i b_i x_i \\ & -\omega \left[\sum_i \sum_{j \in \Phi_i} p_{ij} (e_{ij}^+ + e_{ij}^-) + \sum_{i \in \Psi} p_{ii} x_i \right] \end{aligned} \quad (36)$$

Subject to

$$\sum_i c_i x_i \leq \mu \quad (37)$$

$$x_i - x_j - e_{ij}^+ + e_{ij}^- = 0 \quad \forall i, j \in \Phi_i \quad (38)$$

$$x_i = \{0, 1\} \quad \forall i \quad (39)$$

$$e_{ij}^+, e_{ij}^- = \{0, 1\} \quad \forall i, j \in \Phi_i \quad (40)$$

The objective, (36), optimizes the weighted combination of benefit and shape. Specifically, the first component represents the maximization of total

benefit and the second component minimizes total perimeter in an attempt to encourage compactness and contiguity associated with acquired land. Constraint (38) limits total acquired land by the project budget. Constraints (39) track perimeter resulting from land configuration by accounting for instances where one of two neighboring parcels is selected. Constraints (40) impose integer restrictions on decision variables.

The land use planning model explicitly tracks external edges to account for perimeter. If two neighboring parcels, i and j , are both selected, then $x_i = x_j = 1$. This forces $e_{ij}^+ = e_{ij}^- = 0$ in constraint (39) given the objective component of minimizing external edge in (36). This is what should happen in this case because both parcels are selected and there is no external edge that results between these two parcels. A similar situation occurs when neither are selected. When an external edge is produced (only one of the two neighbors is selected), then the edge must be accounted for. This happens in the model in constraints (39) combined with the minimization objective for total resulting external perimeter. For region boundary parcels, perimeter edge is produced and is accounted for through $\sum_i p_{ii} x_i$.

As with the location model, one facet of the spatial nature of the land use model is that parcels with a specific geographic reference are to be selected. Beyond this, the spatial relationship between neighboring parcels is essential. In this case, the spatial relationship has to do with adjacent parcels, and tracking the resulting perimeter that results from parcel selection. Again, GIS facilitates both aspects of structuring space mathematically.

Conclusion

Spatial analysis techniques have become increasingly accessible to non-specialists, because they have become incorporated into many user friendly software implementations. These include traditional desktop GIS, but also statistical packages and special-purpose software tools. The presence of spatial analytical techniques ranges from software for handheld devices to high performance computing and includes both commercial as well as open source solutions. One particularly exciting recent development is the move towards a cyberinfrastructure for spatial analysis, or cyberGIS (Wang and Armstrong 2009, Wang 2010) which should provide the basis for extensive collaboration across disciplines, using state of the art tools that take into account location, spatial proximity and relationships.

In this chapter, we have attempted to provide a sense of the way in which explicitly accounting for space extends a number of methods. This ranged from data exploration and visualization to regression and optimization. The examples provided are only a small subset of the range of techniques encompassed under the term *spatial analysis*. The references included (as well as references contained in them) should provide an initial guide to the wide range of application of these techniques and their methodological foundations.

Author Note

1. Luc Anselin is Regents' Professor and Walter Isard Chair in the School of Geographical Sciences and Urban Planning and Director of the GeoDa Center for Geospatial Analysis and Computation, Arizona State University, luc.anselin@asu.edu
2. Alan T. Murray is Professor in the School of Geographical Sciences and Urban Planning and the GeoDa Center for Geospatial Analysis and Computation, Arizona State University, atmurray@asu.edu
3. Sergio J. Rey is Professor in the School of Geographical Sciences and Urban Planning and the GeoDa Center for Geospatial Analysis and Computation, Arizona State University, srey@asu.edu

References

- Abler, R. (1987). The National Science Foundation National Center for Geographic Information and Analysis. *International Journal of Geographical Information Systems*, 1:303–326.
- Ando, A., Camm, J., Polasky, S., and Solow, A. (1998). Species distributions, land values, and efficient conservation. *Science*, 279:2126–2168.
- Andrienko, N. and Andrienko, G. (2006). *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. Springer Verlag, Heidelberg.
- Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers, Dordrecht.
- Anselin, L. (1990). Spatial dependence and spatial structural instability in applied regression analysis. *Journal of Regional Science*, 30:185–207.
- Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis*, 27:93–115.
- Anselin, L. (1996). The Moran scatterplot as an ESDA tool to assess local instability in spatial association. In Fischer, M., Scholten, H., and Unwin, D., editors, *Spatial Analytical Perspectives on GIS in Environmental and Socio-Economic Sciences*, pages 111–125. Taylor and Francis, London.
- Anselin, L. (1999). Interactive techniques and exploratory spatial data analysis. In Longley, P. A., Goodchild, M. F., Maguire, D. J., and Rhind, D. W., editors, *Geographical Information Systems: Principles, Techniques, Management and Applications*, pages 251–264. John Wiley, New York, NY.
- Anselin, L. (2001). Spatial econometrics. In Baltagi, B., editor, *A Companion to Theoretical Econometrics*, pages 310–330. Blackwell, Oxford.
- Anselin, L. (2003). Spatial externalities, spatial multipliers and spatial econometrics. *International Regional Science Review*, 26:153–166.
- Anselin, L. (2006). Spatial econometrics. In Mills, T. and Patterson, K., editors, *Palgrave Handbook of Econometrics: Volume I, Econometric Theory*, pages 901–969. Palgrave Macmillan, Basingstoke.
- Anselin, L. (2010). Thirty years of spatial econometrics. *Papers in Regional Science*, 89:2–25.
- Anselin, L. and Bera, A. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. In Ullah, A. and Giles, D. E., editors, *Handbook of Applied Economic Statistics*, pages 237–289. Marcel Dekker, New York, NY.
- Anselin, L. and Getis, A. (1992). Spatial statistical analysis and geographic information systems. *The Annals of Regional Science*, 26:19–33.
- Anselin, L., Le Gallo, J., and Jayet, H. (2008). Spatial panel econometrics. In Matyas, L. and Sevestre, P., editors, *The Econometrics of Panel Data, Fundamentals and Recent Developments in Theory and Practice (3rd Edition)*, pages 627–662. Springer-Verlag, Berlin.
- Anselin, L. and Rey, S. J. (2010). *Perspectives on Spatial Analysis*. Springer-Verlag, Heidelberg.
- Anselin, L., Syabri, I., and Kho, Y. (2006). GeoDa, an introduction to spatial data analysis. *Geographical Analysis*, 38:5–22.
- Arbia, G. (2006). *Spatial Econometrics: Statistical Foundations and Applications to Regional Convergence*. Springer-Verlag, Berlin.
- Assunção, R. M. (2003). Space varying coefficient models for small area data. *Environmetrics*, 14:453–473.
- Bailey, T. C. and Gatrell, A. C. (1995). *Interactive Spatial Data Analysis*. John Wiley and Sons, New York, NY.
- Baltagi, B. H. (2001). *Econometric Analysis of Panel Data (Second Edition)*. John Wiley & Sons, Chichester.
- Baltagi, B. H., Egger, P., and Pfaffermayr, M. (2006). A generalized spatial panel data model with random effects. Working paper, Syracuse University, Syracuse, NY.
- Baltagi, B. H., Song, S. H., and Koh, W. (2003). Testing panel data regression models with spatial error correlation. *Journal of Econometrics*, 117:123–150.
- Banerjee, S., Carlin, B. P., and Gelfand, A. E. (2004). *Hierarchical Modeling and Analysis for Spatial Data*. Chapman & Hall/CRC, Boca Raton.
- Basile, R. and Gress, B. (2005). Semi-parametric spatial autocovariance models of regional growth in europe. *Région et Développement*, 21:93–118.
- Beron, K. J. and Vijverberg, W. P. (2004). Probit in a spatial context: a Monte Carlo analysis. In Anselin, L., Florax, R. J., and Rey, S. J., editors, *Advances in Spatial Econometrics*, pages 169–195. Springer-Verlag, Heidelberg.
- Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society B*, 36:192–225.
- Best, N. G., Arnold, R. A., Thomas, A., Waller, L. A., and Conlon, E. M. (1999). Bayesian models for spatially correlated disease and exposure data. In Bernardo, J., Berger, J., Dawid, A., and Smith, F., editors, *Bayesian Statistics 6*, pages 131–156. Oxford University Press, New York, NY.
- Casetti, E. (1972). Generating models by the expansion method: Applications to geographical research. *Geographical Analysis*, 4:81–91.

- Casetti, E. (1997). The expansion method, mathematical modeling, and spatial econometrics. *International Regional Science Review*, 20:9–33.
- Chilès, J.-P. and Delfiner, P. (1999). *Geostatistics, Modeling Spatial Uncertainty*. John Wiley & Sons, New York, NY.
- Church, R. L. and Murray, A. T. (2009). *Business Site Selection, Location Analysis and GIS*. John Wiley, New York, NY.
- Cliff, A. and Ord, J. K. (1981). *Spatial Processes: Models and Applications*. Pion, London.
- Cressie, N. (1991). *Statistics for Spatial Data*. Wiley, New York, NY.
- de Castro, M. C. and Singer, B. H. (2006). Controlling the false discovery rate: A new application to account for multiple and dependent tests in local statistics of spatial association. *Geographical Analysis*, 38:180–208.
- de Smith, M. J., Goodchild, M. F., and Longley, P. A. (2007). *Geospatial Analysis. A Comprehensive Guide to Principles, Techniques and Software Tools (2nd Edition)*. Matador, Leicester.
- Diggle, P. J. (2003). *Statistical Analysis of Point Patterns*. Arnold, New York, NY.
- Downs, J., Gates, R., and Murray, A. (2008). Estimating carrying capacity for sandhill cranes using habitat suitability and spatial optimization models. *Ecological Modelling*, 214:284–292.
- Dreicer, Z. and Hamacher, H. (2002). *Facility Location: Applications and Theory*. Springer, Berlin.
- Edmonds, J. (1962). Covers and packings in a family of sets. *Bulletin of the American Mathematical Society*, 68:494–499.
- Elhorst, J. P. (2003). Specification and estimation of spatial panel data models. *International Regional Science Review*, 26:244–268.
- Fischer, D. and Church, R. L. (2003). Clustering and compactness in reserve site selection: an extension of the bioniversity management and selection model. *Forest Science*, 49:555–565.
- Fischer, M. M. and Getis, A. (1997). *Recent Developments in Spatial Analysis*. Springer-Verlag, Berlin.
- Fischer, M. M. and Getis, A. (2010). *Handbook of Applied Spatial Analysis. Software Tools, Methods and Applications*. Springer-Verlag, Heidelberg.
- Fleming, M. (2004). Techniques for estimating spatially dependent discrete choice models. In Anselin, L., Florax, R. J., and Rey, S. J., editors, *Advances in Spatial Econometrics*, pages 145–168. Springer-Verlag, Heidelberg.
- Fortin, M.-J. and Dale, M. (2005). *Spatial Analysis: A Guide for Ecologists*. Cambridge University Press, Cambridge.
- Fotheringham, A. S., Brundson, C., and Charlton, M. (1998). Geographically weighted regression: A natural evolution of the expansion method for spatial data analysis. *Environment and Planning A*, 30:1905–1927.
- Fotheringham, A. S., Brundson, C., and Charlton, M. (2002). *Geographically Weighted Regression*. John Wiley, Chichester.
- Fotheringham, A. S. and Rogerson, P. (1994). *Spatial Analysis and GIS*. Taylor and Francis, London.
- Fotheringham, A. S. and Rogerson, P. (2009). *The SAGE Handbook of Spatial Analysis*. Sage Publications, London.
- Gelfand, A. E., Kim, H.-J., Sirmans, C., and Banerjee, S. (2003). Spatial modeling with spatially varying coefficient processes. *Journal of the American Statistical Association*, 98:387–396.
- Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741.
- Gilks, W., Richardson, S., and Spiegelhalter, D. (1996). *Markov Chain Monte Carlo in Practice*. Chapman and Hall, London.
- Goodchild, M. F. (1987). A spatial analytical perspective on geographical information systems. *International Journal of Geographical Information Systems*, 1:31–45.
- Goodchild, M. F. (1992). Geographical information science. *International Journal of Geographical Information Systems*, 6:31–45.
- Goodchild, M. F. (2010). Whose hand on the tiller? revisiting “spatial statistical analysis and GIS”. In Anselin, L. and Rey, S. J., editors, *Perspectives on Spatial Analysis*, pages 49–59. Springer-Verlag, Heidelberg.
- Goodchild, M. F., Anselin, L., Appelbaum, R., and Harthorn, B. (2000). Toward spatially integrated social science. *International Regional Science Review*, 23:139–159.
- Goodchild, M. F., Haining, R. P., Wise, S., and others (1992). Integrating GIS and spatial analysis – problems and possibilities. *International Journal of Geographical Information Systems*, 1:407–423.
- Goodchild, M. F. and Longley, P. (1999). The future of GIS and spatial analysis. In Longley, P., Goodchild, M. F., Maguire, D. J., and Rhind, D. W., editors, *Geographical Information Systems: Principles, Techniques, Applications and Management*, pages 567–580. Wiley, New York, NY.
- Griffith, D. A. (1988). *Advanced Spatial Statistics*. Kluwer, Dordrecht.
- Haining, R. (1990). *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge University Press, Cambridge.
- Haining, R. (2003). *Spatial Data Analysis: Theory and Practice*. Cambridge University Press, Cambridge.
- Haining, R., Wise, S., and Ma, J. (1998). Exploratory spatial data analysis. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47:457–469.
- Harris, R., Sleight, P., and Webber, R. (2005). *Geodemographics, GIS and Neighborhood Targeting*. John Wiley & Sons Inc, New York, NY.
- Hastie, T. and Tibshirani, R. (1993). Varying-coefficient models. *Journal of the Royal Statistical Society B*, 55:757–796.
- Journel, A. and Huijbregts, C. (1978). *Mining Geostatistics*. Academic Press, London.
- Kapoor, M., Kelejian, H. H., and Prucha, I. (2007). Panel data models with spatially correlated error components. *Journal of Econometrics*, 140:97–130.
- Kelejian, H. H. and Prucha, I. (2001). On the asymptotic distribution of the Moran I test statistic with applications. *Journal of Econometrics*, 104:219–257.
- Lawson, A. and Denison, D. (2002). *Spatial Cluster Modelling*. CRC Press, Boca Raton, FL.
- Lawson, A. B. (2009). *Bayesian Disease Mapping, Hierarchical Modeling in Spatial Epidemiology*. CRC Press, Boca Raton, FL.
- Lee, L.-F. and Yu, J. (2010). Some recent developments in spatial panel data models. *Regional Science and Urban Economics*, 40:255–271.
- LeSage, J. P. (1997). Bayesian estimation of spatial autoregressive models. *International Regional Science Review*, 20:113–129.
- LeSage, J. P. (2000). Bayesian estimation of limited dependent variable spatial autoregressive models. *Geographical Analysis*, 32:19–35.

- LeSage, J. P. and Pace, R. K. (2009). *Introduction to Spatial Econometrics*. CRC Press, Boca Raton, FL.
- Miller, R. E. (2000). *Optimization*. Wiley, New York, NY.
- Mirchandani, P. and Francis, R. (1990). *Discrete Location Theory*. Wiley, New York, NY.
- Moran, P. A. P. (1948). The interpretation of statistical maps. *Biometrika*, 35:255–260.
- Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 37:17–23.
- Murray, A. T. (2005). Geography in coverage modeling: exploiting spatial structure to address complementary partial service of areas. *Annals of the Association of American Geographers*, 95:761–772.
- Murray, A. T. (2007). Optimization. In Kemp, K. K., editor, *Encyclopedia of Geographic Information Science*, pages 333–334. Sage, Thousand Oaks, CA.
- Murray, A. T. (2010). Advances in location modeling: GIS linkages and contributions. *Journal of Geographical Systems*, 12:335–354.
- Murray, A. T. and Estivill-Castro, V. (1998). Cluster discovery techniques for exploratory spatial data analysis. *International Journal of Geographical Information Science*, 12:431–443.
- Murray, A. T., Grubesic, T. H., Mack, E. A., Wei, R., Rey, S. J., Anselin, L., and Griffin, M. (2012). Convicted sex offender residential movements. In Leitner, M., editor, *Crime Modeling and Mapping Using Geospatial Technologies*. New York: Springer.
- Murray, A. T., Tong, D., and Kim, K. (2010). Enhancing classic coverage location models. *International Regional Science Review*, 33:115–133.
- Nalle, D. J., Arthur, J. L., and Sessions, J. (2002). Designing compact and contiguous reserve networks with a hybrid heuristic algorithm. *Forest Science*, 48:59–68.
- Onal, H. and Briers, R. A. (2003). Selection of a minimum-boundary reserve network using integer programming. *Proceedings of the Royal Society London B*, 270:1487–1491.
- Ord, J. K. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, 70:120–126.
- Ord, J. K. and Getis, A. (1995). Local spatial autocorrelation statistics: distributional issues and an application. *Geographical Analysis*, 27:286–306.
- O'Sullivan, D. and Unwin, D. (2003). *Geographic Information Analysis*. John Wiley & Sons Inc, Chichester.
- Paelinck, J. and Klaassen, L. (1979). *Spatial Econometrics*. Saxon House, Farnborough.
- Pfeiffer, D. U., Robinson, T. P., Stevenson, M., Stevens, K. B., Rogers, D. J., and Clemens, A. C. A. (2008). *Spatial Analysis in Epidemiology*. Oxford University Press, Oxford.
- Pinkse, J. and Slade, M. E. (1998). Contracting in space: An application of spatial statistics to discrete-choice models. *Journal of Econometrics*, 85:125–154.
- Pinkse, J. and Slade, M. E. (2010). The future of spatial econometrics. *Journal of Regional Science*, 50:103–117.
- Pinkse, J., Slade, M. E., and Brett, C. (2002). Spatial price competition: A semiparametric approach. *Econometrica*, 70:1111–1153.
- Rey, S. J. (2001). Spatial empirics for economic growth and convergence. *Geographical Analysis*, 33:195–214.
- Rey, S. J. and Anselin, L. (2007). PySAL: A Python library of spatial analytical methods. *The Review of Regional Studies*, 37:5–27.
- Rey, S. J. and Janikas, M. V. (2006). STARS: Space-Time Analysis of Regional Systems. *Geographical Analysis*, 38:67–86.
- Rey, S. J. and Le Gallo, J. (2009). Spatial analysis of economic convergence. In Patterson, K. and Mills, T. C., editors, *Palgrave Handbook of Econometrics*, pages 1251–1290. Palgrave, Basingstoke.
- Ripley, B. D. (1981). *Spatial Statistics*. Wiley, New York, NY.
- Rossi, R., Mulla, D., Journel, A., and Franz, E. (1992). Geostatistical tools for modeling and interpreting ecological spatial dependence. *Ecological Monographs*, 62:277–314.
- Schabenberger, O. and Gotway, C. A. (2005). *Statistical Methods for Spatial Data Analysis*. Chapman & Hall/CRC, Boca Raton, FL.
- Shirabe, T. (2005). A model of contiguity for spatial unit allocation. *Geographical Analysis*, 37:2–16.
- Snyder, S., Tyrrell, L., and Haight, R. (1999). An optimization approach to selecting research natural areas in National Forests. *Forest Science*, 45:458–469.
- Thompson, E., Halterman, B., Lyon, T., and Miller, R. (1973). Integrating timber and wildlife management planning. *Forestry Chronicle*, 47:247–250.
- Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46:234–240.
- Toregas, C., Swain, R., ReVelle, C., and Bergman, L. (1971). The location of emergency service facilities. *Operations Research*, 19:1363–1373.
- Tukey, J. (1977). *Exploratory Data Analysis*. Addison-Wesley, New York, NY.
- Unwin, A. and Unwin, D. (1998). Exploratory spatial data analysis with local statistics. *The Statistician*, 47:415–421.
- Waller, L., Zhu, L., Gotway, C., Gorman, D., and Gruenwald, P. (2007). Quantifying geographic variations in associations between alcohol distribution and violence: A comparison of geographically weighted regression and spatially varying coefficient models. *Stochastic Environmental Research and Risk Assessment*, 21:573–588.
- Waller, L. A. and Gotway, C. A. (2004). *Applied Spatial Statistics for Public Health Data*. John Wiley, Hoboken, NJ.
- Wang, S. (2010). A cyberGIS framework for the synthesis of cyberinfrastructure, GIS, and spatial analysis. *Annals of the Association of American Geographers*, 100:535–557.
- Wang, S. and Armstrong, M. P. (2009). A theoretical approach to the use of cyberinfrastructure in geographical analysis. *International Journal of Geographical Information Science*, 23:169–193.
- Weber, A. (1909). *Ueber den Standort der Industrien* (English translation C.J. Friedrich, 1929, *Theory of the Location of Industries*. Chicago: University of Chicago Press, Tubingen).
- Whittle, P. (1954). On stationary processes in the plane. *Biometrika*, 41:434–449.
- Williams, J. (2002). A zero-one programming model for contiguous land acquisition. *Geographical Analysis*, 34:330–349.
- Wolpert, R. L. and Ickstadt, K. (1998). Poisson/Gamma random field models for spatial statistics. *Biometrika*, 85:251–267.
- Wright, J., ReVelle, C., and Cohen, J. (1983). A multiobjective integer programming model for the land acquisition problem. *Regional Science and Urban Economics*, 13:31–53.
- Wu, X. and Murray, A. T. (2007). Spatial contiguity optimization in land acquisition. *Journal of Land Use Science*, 2:243–256.
- Zhang, Y. and Wright, J. (2004). Global optimization of combined region aggregation and leveling model. *Journal of Computing in Civil Engineering*, 18:154–161.