CHAPTER 33

Statistical Analysis of Spatial Crime Data

WIM BERNASCO AND HENK ELFFERS

INTRODUCTION

Social scientists become increasingly aware of the relevance of space and place (Goodchild et al. 2000), and criminology is no exception. As a matter of fact, the geography of crime has been a focal concern of criminologists from the very start of the discipline. In nineteenth century Europe, the "moral statistics" of pioneers, Guerry and Quetelet, empirically demonstrated that crime varied across geographical regions. They not only produced maps that visualized these differences, but also studied statistical relations between crime, poverty, and education.

During the early twentieth century, researchers associated with the University of Chicago studied how crime and other social problems varied across urban communities, again mapping the geographical patterns and using community characteristics to explain these distributions. The work on juvenile delinquency has become a classic example (Shaw and McKay 1942).

Interest in the geography of crime decreased somewhat between 1950 and 1980, to some extend possibly because the ecological fallacy (Robinson 1950) hampered the interpretation of aggregated data. Later, the link between communities and crime was revitalized (Bursik and Grasmick 1993; Sampson et al. 1997). In recent years, a new concern has emerged with micro units of place such as addresses or street segments (Eck and Weisburd 1995; St. Jean 2007).

From the very start, geographical criminology has been an area of research where methodological and statistical innovations were either developed or adopted early. In their times, Guerry and Quetelet were pioneers and innovators, and their work is said to have been the launching pad for much of modern social science (Beirne 1987; Friendly 2007). When hierarchical linear (multilevel) models were developed in the 1980s, criminologists and sociologists who studied the links between community and crime quickly embraced and applied them to model community context effects, and even took a lead by developing a new "econometrics" of crime measurement (Raudenbush and Sampson 1999). When trajectory models were developed to model the criminal development of individuals (Nagin 1999), geographical criminologists soon saw their value for modeling the crime trajectories of geographical entities (Griffiths and Chavez 2004; Weisburd et al. 2004). As another example, spatial econometrics (Anselin 1988) has quickly diffused into the criminology field. In the research on crime

A.R. Piquero and D. Weisburd (eds.), *Handbook of Quantitative Criminology*,
DOI 10.1007/978-0-387-77650-7_33, © Springer Science + Business Media, LLC 2010

location choice, developments in discrete choice modeling have been adopted (Bernasco and Nieuwbeerta 2005). The recent focus on small units of analysis creates new methodological challenges for geographic criminology (Weisburd et al. 2009). In sum, geographic criminology has always been at the cutting edge of major methodological and empirical progress.

The purpose of the present chapter is to provide an up-to-date overview of methods for the statistical modeling of spatial crime data, to review some instructive and innovative applications in the field, and to direct the reader to the relevant literature.

The chapter consists of three sections. The first section introduces and delineates the subject matter. It discusses the relevance of spatial analysis, describes what spatial data are, which spatial units of analysis can be distinguished, and how they are sampled. The section further addresses criminological categories that can be geographically referenced, and delineates spatial modeling from descriptive spatial statistics and from visualization techniques ("crime mapping") that are treated elsewhere in this volume.

We distinguish two types of spatial outcomes that can be modeled: spatial distribution, and movement. The second section deals with the analysis of spatial distributions. We discuss how spatial structure is specified in spatial statistics, address the basic concept of spatial autocorrelation, and review a variety of spatially informed regression models and their uses in criminology. The third section addresses the analysis of movement. We address the length of the journey-to-crime, and discuss spatial interaction models, spatial choice models, and the analysis of mobility triads, again highlighting applications in the field of crime and criminal justice.

This chapter resembles and builds upon a review that appeared nearly a decade ago (Anselin et al. 2000). Compared to that review, the present chapter dedicates less space to theory, to geographic information systems (GIS) and to descriptive spatial analysis methods, and more to the analysis of spatial choice and movement.

What are Spatial Crime Data?

All methods discussed in this chapter apply to spatial crime data. Crime data are simply data that bear a direct relation to crime. Often the data apply to people in their roles of offenders, accomplices, fences, victims, bystanders, police officers or judges. They can also be crime targets, such as houses (for burglary), empty walls (for graffiti), cars (for theft), or airplanes (for hijacking). Most often, however, the data are the criminal events themselves: the burglaries, rapes, arsons, robberies, assaults, and murders.

What makes crime data *spatial* crime data is that the units of analysis are geographically referenced. This means that they have attributes (e.g., a pair of geographical coordinates) that can be used to establish where they are situated relative to the other units in the sample. In modeling spatial distributions, a *weight matrix* (see section "Specification of Spatial Structure: The Spatial Weight Matrix and Chap. 6 by Tita and Radil") specifies the spatial relations between all pairs of observations.

Thus, like in network data and in hierarchically structured data, in spatial data the observational units are interrelated. In spatial data, this relation is geographic in nature. For example, two units are adjacent or non-adjacent, they are nearby or distant, they are nearest neighbors or not.

Many textbooks distinguish spatial data by the spatial characteristics of the units of analysis, e.g., whether the data refer to points, to cells of a grid or to areas (depending on the contexts also referred to as zones, lattices, or polygons). For the purpose of the present review, however, it is more useful to make another distinction, namely between stationary (time invariant) spatial distributions on the one hand, and movement between origins and destinations on the other hand. The first type of data may be referred to as spatial distribution data, the second as movement data. Here are some examples of spatial distribution data on crime and criminal justice issues:

– Geographical coordinates of the home addresses of convicted juvenile offenders in Chicago (Shaw and McKay 1942)
– Numbers of homicides per county in the USA (Baller et al. 2001)
– Percentage of residents reporting to be victims of violent assault in their own neighborhood, for each neighborhood cluster in Chicago (Sampson et al. 1997)
– Geographical coordinates and dates of police reported burglary incidents in Liverpool, England (Bowers and Johnson 2005)
– Numbers of police recorded crimes per street segment in Seattle over a period of 14 years (Weisburd et al. 2004)

Spatial mobility data involve movement between two or more locations. Here are some examples of movement data on crime and criminal justice issues:

– The distance between the home and the place of the offence of serial rapists (Warren et al. 1998).
– Robbery incidents in Chicago, georeferenced according to the census tract of residence and the census tract of the robbery incident (Bernasco and Block 2009).
– Homicides in Washington, DC, georeferenced according to the geographical coordinates of the offender's home, the victim's home and the location of the homicide (Groff and McEwen 2007).
– Numbers of crime trips (linking offender's home to crime site) between neighborhoods in The Hague, the Netherlands (Elffers et al. 2008).

What is Spatial Modeling?

Although spatial models require spatial data, spatial data need not necessarily be analyzed with spatial models. As a matter of fact, most spatial crime data have been analyzed without spatial models. For example, with a few exceptions (e.g., Heitgerd and Bursik 1987; Morenoff et al. 2001) spatial models have not been used in the century-old ecological tradition that studies how neighborhood crime rates are influenced by neighborhood conditions, while neighborhoods are clearly spatial entities. Neither have they been used in cross-national comparisons of crime phenomena, although like neighborhoods, countries are spatial entities. The present chapter will obviously focus on the methods of analysis that actually utilize the spatial nature of the data.

We distinguish between two types of spatial analysis methods. The first is often referred to as exploratory spatial data analysis (acronym ESDA) and is concerned with the description and exploration of spatial data. Typically, the results of these analytical methods are visualized with the use of geographic information systems (GIS). Geographical information systems are software tools for digital cartography that help to process, organize, analyze, and visualize geographically referenced information. Applied to crime and justice topics, this is commonly referred to as "crime mapping." Textbooks that discuss ESDA methods are Bailey and Gatrell (1995) and Haining (2003). Visualization and crime mapping issues are comprehensively dealt with in Chainey and Ratcliffe (2005), and more concisely in the chapter by Ratcliffe (Chap. 2)