Exploratory Spatial Data Analysis Techniques for Examining Urban Crime

Article *in* British Journal of Criminology · March 2001

DOI: 10.1093/bjc/41.2.309 · Source: OAI

Alan T. Murray
University of California, Santa Barbara

Ingrid Mcguffog
State University of New York College at Brockport

BRIT. J. CRIMINOL. (2001) **41**, 309–329

EXPLORATORY SPATIAL DATA ANALYSIS TECHNIQUES FOR EXAMINING URBAN CRIME

Alan T. Murray, Ingrid McGuffog, John S. Western and Patrick Mullins*

Regional planners, policy makers and policing agencies all recognize the importance of better understanding the dynamics of crime. Theoretical and application-oriented approaches which provide insights into why and where crimes take place are much sought after. Geographic information systems and spatial analysis techniques, in particular, are proving to be essential for studying criminal activity. However, the capabilities of these quantitative methods continue to evolve. This paper explores the use of geographic information systems and spatial analysis approaches for examining crime occurrence in Brisbane, Australia. The analysis highlights novel capabilities for the analysis of crime in urban regions.

The occurrence of criminal activity in the form of thefts, assaults, homicides, etc. is something that takes place every day in almost all reaches of our world. There is a great deal of debate on the causes of crime (Fattah 1997). Further, policing agencies are not particularly effective at foreseeing where and when specific future crimes will take place. This is not unexpected, as criminal behaviour is not well understood. Regardless of the reasons for its occurrence, criminal behaviour puts a strain on the communities, towns and cities in which we live. There are significant monetary costs associated with policing crime and prosecuting those who commit crimes and, in non-monetary terms, there are social costs associated with crime. These costs are reflected in changing perceptions of quality of life, mental health and physical security in our daily activities.

As computing technology has advanced, there has been continued interest in the use and development of techniques for helping to explain the occurrence of criminal activity. Perhaps the most influential computer-based tools thus far have been geographical information systems (GIS) and crime mapping software. These two related technologies have facilitated the exploration of the spatial distribution of crime (Ratcliffe and McCullagh 1998a). However, GIS and crime mapping have only scratched the surface of their potential application to the investigation of criminal activity. As noted in the fields of urban geography and sociology (see Smith 1995; Wright *et al.* 1997; Openshaw 1998), there is a constant search for more sophisticated GIS-based tools for studying human behaviour in physical space. It is the ability to combine spatial information with other data that makes GIS so valuable for urban researchers. Nevertheless, the sheer magnitude of information also creates the need for more effective

* Alan T. Murray, Department of Geography, The Ohio State University, Columbus, Ohio; Ingrid McGuffog, John S. Western and Patrick Mullins, University of Queensland Brisbane, Queensland, Australia. This research was supported in part by a grant from the Australian Research Council while the first author was a research fellow at the University of Queensland in the Australian Housing and Urban Research Institute.

approaches within GIS for integrating information and identifying patterns and relationships.

The use of GIS and crime mapping software for studying criminal activity has generally been in the area of environmental criminology, which is concerned with the physical characteristics of areas and their impact on preventing or encouraging crime (see Harries 1980; Brantingham and Brantingham 1981). The relationship between the physical layout of an area, proximity to various services, and land use mixes are seen as important factors which are likely to influence criminal behaviour (Greenburg and Rohe 1984). Environmental criminologists have proposed that certain crimes are more likely to be perpetrated because of issues of access, exposure and opportunity (Brantingham and Brantingham 1981; Byrne and Sampson 1986). That is, significantly high numbers of property crimes or robberies, as an example, may be a result of the design of a suburb, where there are residential and industrial land use mixes, building entrances not visible from main roads, etc. In general, a major limitation of the application of GIS and crime mapping software has been its descriptive use for depicting crime activity rather than any focus on data integration issues or the development of analytically sophisticated tools.

There is significant interest in the occurrence of crime and its relationship to urban change in the south-east Queensland region of Australia, not unlike most metropolitan regions. This region contains approximately 2.1 million residents covering 22,320 square kilometres, with the city of Brisbane being the major urban centre representing over half of the total regional population. South-east Queensland continues to experience significant and sustained population growth, a trend of the past few decades (Stimson and Taylor 1999). Thus, there has been considerable interest in examining the impact of growth and development in this region (Murray *et al.* 1998; Stimson *et al.* 1999). Questions relating to long-term sustainability with respect to environmental degradation, social functionality and economic prosperity continue to be raised, debated and considered. Given that crime has a major influence on the social and economic performance of an urban region, there is a need to understand the significance of environmental characteristics with respect to crime in a region, particularly south-east Queensland.

The aim of this paper is to detail approaches for better understanding relationships of crime occurrence using GIS and quantitative techniques. The next section discusses research into crime occurrence and the various associated theories. This highlights the importance of developing and integrating quantitative systems-based techniques for synthesizing and relating crime activity. Following this, a necessary framework for information collection efforts, data integration and synthesis is detailed associated with the use of socio-demographic services and crime information. A number of spatial modelling-based approaches will be illustrated for identifying relationships associated with criminal activity. Specifically, the application of pattern detection and spatial statistical analysis techniques for exploring the occurrence of crime will be presented. These utilized approaches demonstrate the importance of geographic space in the analysis of crime in growing and developing regions. The establishment of an analytical and theoretical framework for evaluating the relationship between place and crime will undoubtedly lead to enhanced regional crime prevention strategies. Further, an integrated systems-based framework is essential for evaluating theories of crime and space in a rigorous manner.

GIS and Spatial Analysis in Crime Activity Research

Research into crime occurrence has been substantial and continues to be an important endeavour. The reason for this is that there is a need and desire to better understand criminal activity. Quantitative techniques such as statistical analysis have long been utilized in the analysis of crime. As in many areas in the social sciences, data quality and availability issues have certainly thwarted research progress aimed at understanding criminal behaviour. This is beginning to change as GIS and associated spatial information have become more readily accessible in recent years. A GIS represents a collection of hardware and software for creating/inputting, managing, manipulating, displaying and analysing spatial information (Chrisman 1997). A GIS is seen as a particularly important tool for studying the occurrence of crime (Hirschfield *et al.* 1995; Ratcliffe and McCullagh 1998b). One of the major characteristics of GIS is its relative ease of cartographic display of information. That is, most commercial GISs offer the most novice analyst the capability to quickly and efficiently generate maps and visual displays of crime occurrence that are self-explanatory and illustrative. Another major characteristic of GIS is its capability to integrate crime statistics with census and other spatial information.

One aspect of GIS use, as noted previously, is for displaying the location of crime occurrence. This is generally known as crime mapping and has informed criminological theory and prevention efforts in a somewhat sporadic manner for almost two centuries (Weisburd and McEwen 1997). Crime mapping essentially reflects the use of GIS for displaying crime event occurrences in a more rigorous fashion than the use of hardcopy maps and coloured pins. The use of GIS for crime mapping enables analysts to create custom displays focused on particular types of crime in specific areas in order to search visually for trends or patterns (see Ratcliffe and McCullagh 1998a). However, the display-oriented capabilities are only one aspect of GIS in terms of analysing crime in a region. GIS actually provides an integration and modelling framework for analysing different components of crime activity. As an example, Ratcliffe and McCullagh (1998b) utilize GIS for examining temporal changes in crime events. The integration of information is obviously essential for relating crime occurrence to other spatial distributions as well as investigating the existence of patterns.

What is recognized is that automated or semi-automated approaches and techniques for identifying patterns in crime occurrence continue to be necessary (Hirschfield *et al.* 1995; Weisburd and McEwen 1997). Much of the modelling-oriented research aimed at identifying significant relationships or patterns in criminal activity pre-dates the mainstream use of GIS. Thus, GIS may be able to enhance, supplement and extend quantitative techniques utilized in the past. Further, modelling approaches integrated in or coupled with GIS (developed outside the context of crime analysis) may be valuable for identifying potential patterns of criminal activity.

Buck *et al.* (1991) utilized a modelling approach for explaining the impacts of crime as a gaming industry by-product on property values, but did not utilize crime events as a spatial feature. This may be considered a fairly aggregate analysis of the impact of crime. GIS provides the capability to include the spatial dimension of crime to such an analysis. Brown (1982), in contrast to Buck *et al.* (1991), was interested in exploring the spatial variation of crime and hence patterns of activity. A multiple linear regression approach was structured in Brown (1982) to investigate access to the city (along with other social and environmental variables) on crime rates in suburbs. The work of Brown (1982) is noteworthy because it checked for the existence of spatial autocorrelation, which may alter or weaken the explanatory power and significance of regression results. The need to account for spatial effects is critical in the context of geographic phenomena such as crime (see Anselin and Bao 1997; Anselin 1998). GIS provides the capacity to add further environmental characteristics to an analysis like that of Brown (1982), where proximity to physical barriers or landmarks, roads and other transportation services can be included as well. It is worth noting that only in recent years have we truly reached a point where it is possible to do thorough spatial statistical analysis linked to a GIS environment.

It is clear that pattern detection approaches are critical in the analysis of criminal activity (Hirschfield *et al.* 1995; Weisburd and McEwen 1997). However, there are no readily defined and developed approaches in the literature for finding spatial patterns of criminal occurrence. This is not particularly surprising because most fields of study, in the social sciences in particular, recognize that a range of techniques is typically necessary to reach informed and sound conclusions about the existence of patterns, trends or relationships. GIS certainly provides a number of necessary elements for effectively investigating and relating crime. Further, modelling approaches continue to be important and do benefit from linkage/integration with GIS. This point will be illustrated throughout the paper.

Information

The analysis of criminal activity requires one important item, information on crime occurrence. In Canada crime information is relatively easy to access as a component of census data. In the United States it is possible to obtain information on crime occurrence from major criminal justice agencies such as the Federal Bureau of Investigation or Department of Justice. In Australia, however, the study of crime is a significantly more difficult task. There are no crime attributes reported in the quinquennial census in Australia. Further, there are no major Commonwealth agencies or departments facilitating data collection or integration. This certainly makes research and analysis of crime somewhat challenging. It is therefore the responsibility of state and local governments to organize and manage crime statistics.

In Queensland, it is the responsibility of Queensland Police Service to record and manage crime events. For the research reported in this paper, crime occurrence for 1996 was utilized. This information was made available to us in the form of spreadsheet statistics on crime by suburb. This represents approximately 541 suburbs in the south-east Queensland region and the information is not necessarily complete for some rural areas. There are approximately 64 crime attributes reported by Queensland Police Service. The categories of crime include homicide, assault, sexual offence, robbery, extortion, kidnapping, property damage, breaking and entering, arson, vehicle theft, fraud, etc. For illustration purposes we will utilize the combined category of property crimes, standardized per 1,000 residents, in the analysis presented in this paper.

Information on criminal activity alone is almost meaningless. Rather, it is in relating crime to features and elements in regions that its significance can be appreciated or is of interest. One critical item to relate crime occurrence to is people and their associated characteristics. Such information is typically available from annual censuses. In Australia census information is collected and distributed by the Australian Bureau of Statistics. Census data contain socio-demographic information for areas such as population, age group proportions, occupation, education, income, household size, employment, rent, religion, number of vehicles, etc. at a variety of spatial scales. In our research we have utilized information from the 1996 Census at the collection district level, which is the most spatially disaggregate level of the Australian census hierarchy. This spatial scale averages about 225 dwellings per delineated area. Although collection districts are theoretically supposed to correspond to suburbs, in urban regions such as Brisbane this is not the case. Collection districts in urban regions tend to be substantially smaller than recognized suburbs. The next level of the census hierarchy is the statistical local area and it does typically correspond to urban suburbs. However, given the difficulty in reconciling the differences between statistical local areas and recognized suburbs (as reported by Queensland Police Service), it is more cost effective to work with collection district-based census information.

Other important information for studying crime is associated with geographic features. That is, the capability to relate spatial structures to suburbs, the areas for which we know the levels of criminal activity, is particularly valuable. In our research we have focused on a number of major geographic features: railways, roads, public transport, supermarkets, police stations, fire stations, the city centre, and the river. Information on the main roads in Queensland as well as the public transport stops (including bus, rail and ferry services) was provided by Queensland Transport. There are roughly 20,500 unique main roads (categories 1–3) in the region and approximately 11,000 public transport stops. The supermarket locations in the region had to be digitized. There are 169 major supermarket chain stores (Bi-Lo, Coles, Franklins and Woolworths) distributed throughout the region. Information on the 99 police stations was provided by Queensland Police. Information on the location of the 82 fire stations was provided by Queensland Fire and Rescue Authority. Each of the data sources reflects current information at the time of this research.

Due to data quality issues as well as the need to avoid comparisons based upon urban and rural areas, analysis presented in this paper is limited to the Brisbane Statistical Division, the central sub-region of south-east Queensland containing roughly 1.5 million people in 178 suburbs.

Integration of Spatial Layers

Each of the data items discussed in the previous section may be thought of as individual spatial layers. The suburb boundaries symbolize one layer, with each suburb indicating associated crime statistics. The collection district boundaries, which report the socio-demographic and economic profile of each collection district, represent another layer. Similarly, the road network corresponds to another layer. Still another layer is the public transport stops and so on. While a collection of spatial data layers may be valuable and informative, there is little hope of interpreting relationships between layers unless they share a common referencing system and data formatting which allows them to be incorporated into a particular GIS. That is, the patterns and relationships that are most likely to be of interest will be likely to comprise some combination of the various data layers.

In this research ArcView version 3.1, a commercial GIS package, has been utilized for integrating spatial data layers. Some of the information layers were converted to an appropriate digital format for ArcView. For example, census data is distributed using MapInfo (another commercial GIS package) in Australia, so it was necessary to convert the Brisbane region collection districts to ArcView formatted files. Functionality for converting between these two particular systems is a fairly standard exercise supported in the most recent versions of these two software packages. It was also necessary to utilize Access, a commercial data base management package, for manipulating information (standardization, scaling, etc.).

Suburbs were favoured as the spatial unit to relate the various layers in selected ways. As the crime information is reported at this spatial scale, there was essentially no other choice. There are a number of approaches which may be utilized for relating spatial data layers (see Chrisman 1997). Three worth detailing here are interpolation, proximity and containment. Mentioned previously was the fact that the census information utilized in this research was reported by collection districts (the lowest level of spatial resolution in the Australian census hierarchy). A suburb typically contains a number of collection districts. In order to infer census attributes at the suburb level an areal interpolation process must be used. As an example, one census attribute is total population. In order to determine the total population in the suburb, all collection districts within the suburb must be identified and their population totals summed. Another census attribute is unemployment rates for each collection district. The calculation of the unemployment rate for the suburb involves combining unemployment rates of collection districts based upon their population proportions within the suburb. Thus, the interpolation process in this context consists of aggregating collection district attributes at the suburb level. Such capabilities are commonly provided in most commercial GIS packages.

Proximity analysis is another approach for relating spatial information layers. The use of proximity analysis in this context associates spatial features through their physical relationships. Examples of proximity measures derived in this research include: determining the minimum distance from suburbs to their closest main road and railway; determining the distance of the closest supermarket to a suburb; determining the distance of the closest police station and fire station to a suburb; determining the minimum distance to specific forms of public transport; and, determining the distance from a suburb to the Brisbane CBD (central business district) and the Brisbane River. Figure 1a depicts a proximity-based evaluation of the highlighted suburb of Stafford Heights in Brisbane to main roads. The two closest main roads to Stafford Heights are Pickering St and Gympie Rd, shown in Figure 1. What we are interested in here is the physical distance between Stafford Heights and the closest main road (either Pickering St or Gympie Rd). In this particular case we chose the centre of the suburb as the point from which we would evaluate proximity to main roads using straight line distances. However, we could have utilized shortest path distances using the entire street network for this evaluation. Alternatively, it may be that one is interested in the minimum or maximum physical distance from a suburb to its closest main road, which is another form of proximity analysis. Proximity measurement is characteristic of the relational functions commonly found in GIS, particularly that illustrated in Figure 1a.

Another technique for relating layers of information is spatial containment. The application of spatial containment facilitates the determination of the existence or number of features within an area. An important spatial containment-oriented attribute

Fig. 1 Basic GIS spatial relationship functions: (a) proximity analysis, (b) spatial containment

in this research is the number of public transport stops within a suburb. Population densities may also be considered an important spatial containment-oriented attribute. The nature of spatial containment is depicted in Figure 1b, where the highlighted suburb of Corinda is shown to contain a number of public transport stops (19 in this case). Using GIS it is quite easy to identify which stops are within a particular suburb. Thus, the number of stops in each suburb, in this case, may be readily computed in most commercial GIS packages.

Exploratory Analysis

Given the existence of spatial information layers, the analysis effort may now be directed towards investigating the potential inferences that may be made. This is essentially exploratory spatial data analysis (ESDA) and very much compliments the purposes and functionality of GIS. Numerous authors have discussed the need for tools and techniques to assist in ESDA (Openshaw 1991; Anselin 1998) and various approaches for carrying out such analysis have been detailed elsewhere (Anselin and Bao 1997; Murray and Estivill-Castro 1998).

There are several approaches that may be utilized for ESDA. Three approaches based upon the use of GIS will be discussed and presented in the sections that follow. The first involves only the use of GIS for visual investigation. That is, using the standard mapping and display functionality provided in a commercial GIS, ArcView in this case, the potential for identifying important and meaningful relationships in crime occurrence in or between spatial information layers will be highlighted. The second approach involves the use of GIS coupled with spatial modelling. In this case we will illustrate the use of cluster analysis and GIS for inferring relationships in the spatial distribution of crime. This is accomplished using developed software which integrates MapObjects (a basic GIS module accessed as a dynamic linked library) and specialized clustering approaches. The final approach to be discussed is the use of GIS and statistical techniques for crime analysis. This essentially involves the use of GIS for data integration and association coupled with commercial statistical packages for investigating incorporated attribute variables. This is carried out using S-Plus and SpaceStat for analysis and ArcView to facilitate exploration and interpretation.

Cartographic Display

The various layers of spatial information may be visually presented in a number of different ways using GIS. One approach is to display crime intensity in suburbs. This typically involves the use of a derived choropleth map to represent the number of crimes in each suburb (see Dent 1990). Specifically, this is a classification technique for colouring areas based upon attribute values. An example of this approach is given in Figure 2, where the seven depicted quantile classes are used to reflect the spatial variation of property crimes in Brisbane. A number of patterns appear to exist in Figure 2. For example, high crime occurrence is depicted around Darra and the Brisbane CBD, whereas low crime occurrence is shown in the suburbs surrounding Belmont. Additional topographic features may be helpful for interpreting observed variation in crime levels,

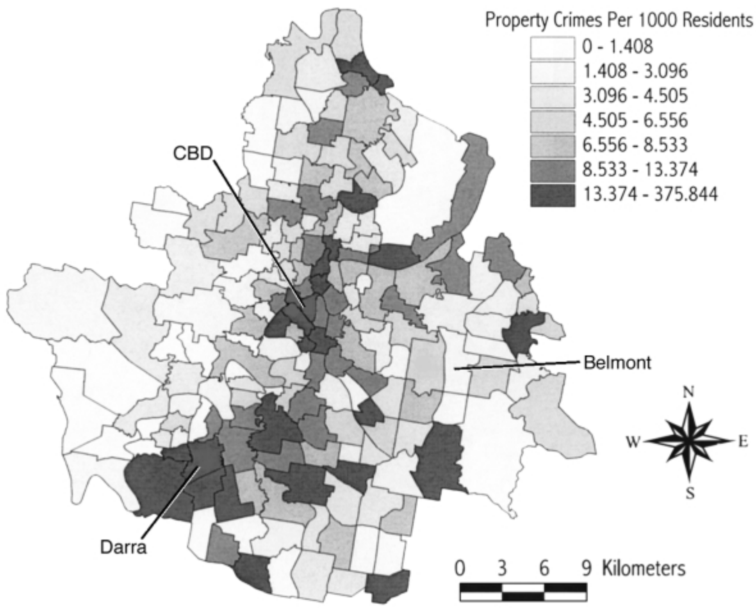

FIG. 2 Property crimes in Brisbane

like rivers, lakes, ocean bodies, etc. or other spatial information layers. For instance, in Figure 3 we have included the main roads information layer. With this additional information, some relationship to certain main roads could be inferred in Figure 3, such as the corridor from Darra to the CBD running south of the Brisbane River. Taking the addition of other information layers a step further, Figure 4 shows the variation in property crimes with respect to main roads, fire stations, police stations, supermarkets and public transport stops. Given the enormous content of information, Figure 4 certainly illustrates that too much information may be a source of confusion. That is, it is difficult to discern any patterns or relationships in this figure due mainly to display clutter and information overload.

There is another aspect to cartographic presentation of information that may also lead to some difficulty in interpreting relationships in spatial variation in crime. There are a number of potential approaches for creating classes in a choropleth display (Dent 1990). Common techniques include the use of quantiles, equal intervals, standard deviations, equal area, etc. In Figures 2–4 quantiles were utilized for class definition. However, if equal intervals are used to define the classes, then an alternative display may be generated. Thus, Figure 5 depicts property crimes using seven equal intervals. The patterning in Figure 5 differs from that shown in Figures 2–4. Only a handful of suburbs appear to have high property crime rates. Class creation is often done with little or no rationale behind a particular approach taken. One of the primary reasons for this is that a choropleth map is easy to specify and create within current GIS environments. The availability of various class interval approaches should be indicative of the need for exploring alternative classifications in order to understand differences in what is being presented. This is highlighted by the information provided in Figure 5 as it clearly enhances what is shown in Figure 2.

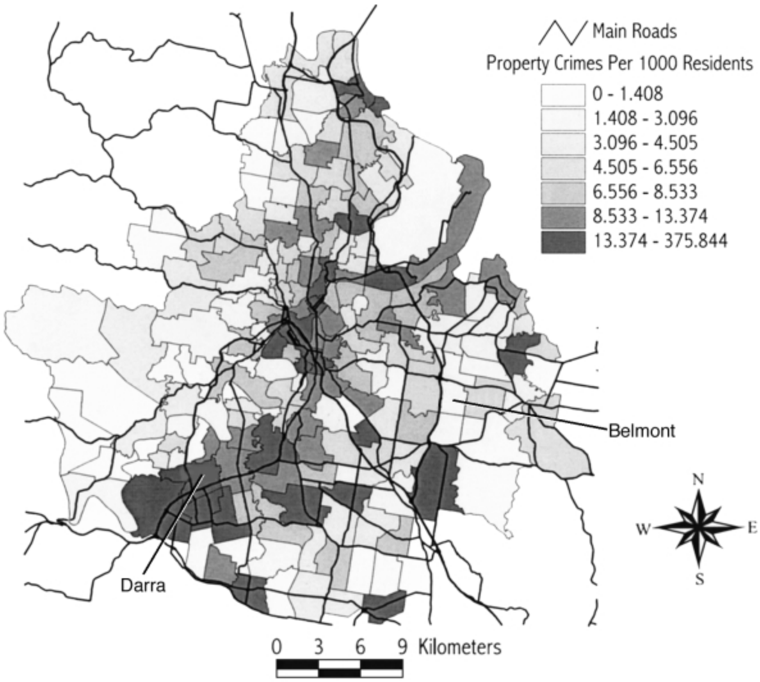

FIG. 3 Property crimes and main roads in Brisbane

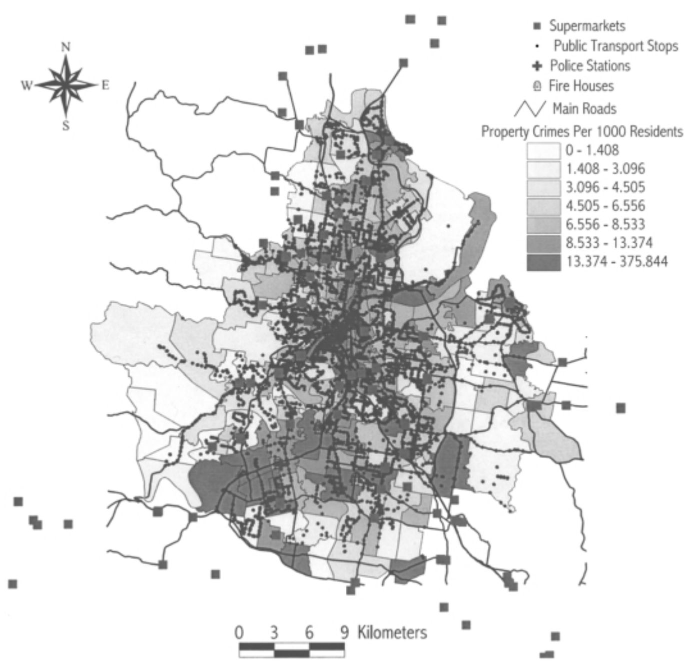

FIG. 4 Property crimes and spatial features in Brisbane

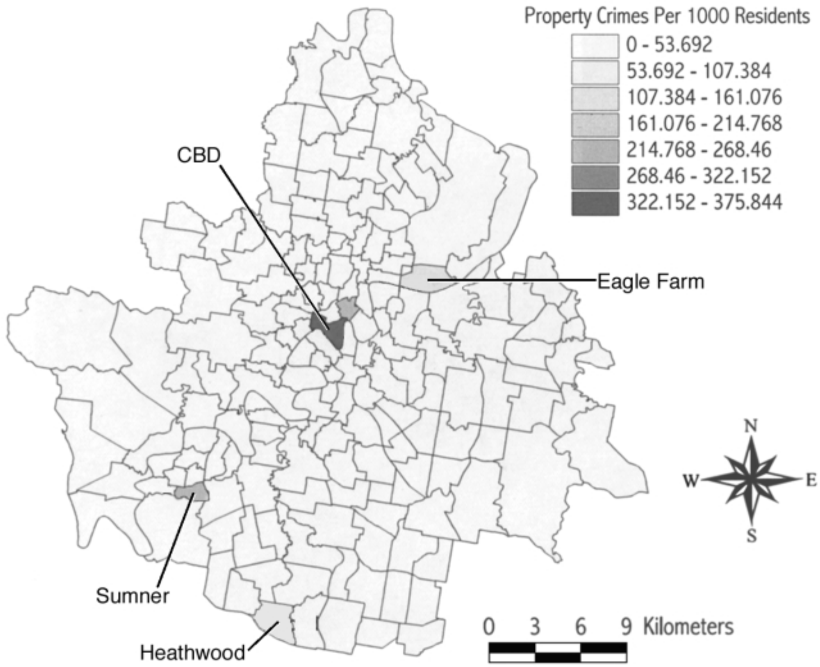

FIG. 5 Property crimes in Brisbane classified using equal intervals

Optimization-based Clustering

The intent of the choropleth map is to convey or illustrate relationships in crime intensity between suburbs. However, this is being done in a very aspatial manner and little possibility exists for establishing any significance to observed patterns. One of the major concerns in crime analysis is the existence of incident clusters. This is an attempt to identify any sub-regional patterns of crime. As an example, we would like to identify suburbs that are strongly related in terms of low, average or high crime occurrence. It would be convenient if such processes could be automated to some degree. Basically we are interested in techniques that are more advanced and sophisticated than visual inspection of information.

There is a field of research in geographical analysis known as spatial data mining (see Murray and Estivill-Castro 1998). The intent of developed spatial data mining techniques is to spot patterns in spatial information in a semi-automated fashion. Differing from choropleth mapping, spatial data mining reflects an interest in both attribute values (number of crime events) and spatial proximity between suburbs. What has primarily been developed thus far is non-hierarchical clustering approaches for spatial application. Murray and Estivill-Castro (1998) review several approaches which may be utilized for spatial data mining. The classic k-means statistical clustering technique represents a basic approach that can be (and has been) applied for spatial data mining. Empirical comparisons of some of the various techniques may be found in Murray (2000). It is worth pointing out the use of clustering as a technique for exploring spatial information is somewhat widely advocated as a basic step in any analysis process (see Griffith and Amrhein 1997) and certainly coincides with the spirit of spatial data mining.

A key to effective pattern spotting in spatial information is that suburb attributes (crime statistics) and spatial proximity (distance between suburbs) is appropriately structured (Murray and Shyy 1999). That is, we would like to identify clusters that have spatial relationships as well as similarity in crime occurrence. Approaches for doing this are detailed in Murray and Shyy (1999). Utilized here is the median-oriented clustering approach. It is interesting to note that integrating attribute characteristics and spatial relationships extends traditional cartographic capabilities of choropleth display such as those discussed in Cromley (1996). It should be obvious that we could use this clustering approach for examining any of the spatial or aspatial attributes associated with the suburbs, like crime rates, population density, public transport density, unemployment rates, etc.

Applying the optimization-based clustering extension developed in Murray and Shyy (1999) to suburb crime statistics identifies a number of interesting patterns of crime occurrence. Depicted in Figure 6 are 21 clusters.[1] The individual clusters are represented as a unique colour. Further, the clustered suburbs have similar property crime levels and tend to be spatially near the other suburbs in their cluster. What is most interesting in Figure 6 is that the highest property crime rate areas appearing in Figure 2 are not necessarily grouped together. Thus, the patterns shown in Figure 6 are quite different to

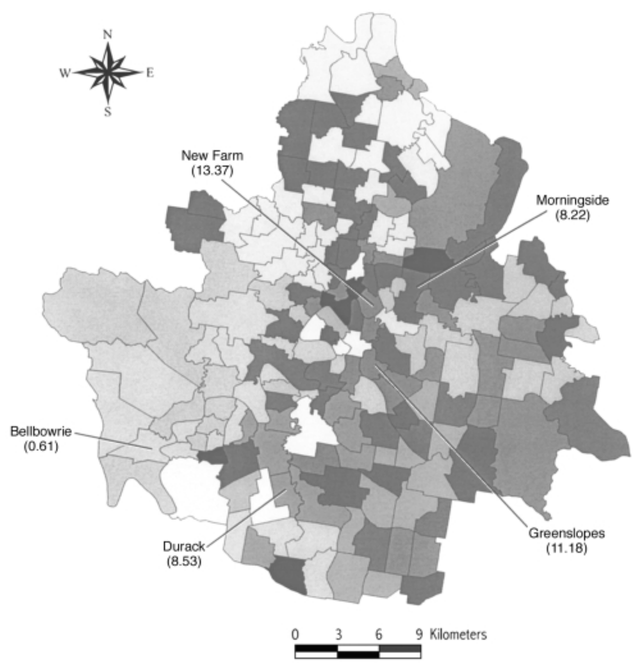

FIG. 6 Twenty-one spatial clusters of property crime

[1] A distance weight of 0.25 (distance measured in kilometres between suburbs) and attribute weight of 1.0 was utilized.

those in Figure 2 (and Figure 5). The primary reason for this is that the higher property crime rate suburbs in Figure 2 around Darra and the CBD, as an example, are not particularly similar to each other, although they are all above the regional mean. Alternatively, there are some pattern similarities to be found in Figures 2 and 6. For example, the group containing Belmont is rather obvious as is the Bellbowrie cluster (both clusters of low property crime suburbs). These clusters clearly demarcate recognized middle class/elite suburbs. An interesting cluster in Figure 6 is the one that includes Morningside, which has a property crime rate of approximately 8.22 per 1,000 residents (indicated in brackets). This cluster is fairly industrial with shipping and ship building and maintenance establishments. On the higher side, and somewhat circling the CBD, is a cluster that includes New Farm, with a property crime rate of approximately 13.37 per 1,000 residents. It is worth noting that these highlighted clusters are consistently identified across a range of cluster values (the number of generated clusters), which suggests further significance to their existence (in addition to the fact that there is implicit significance associated with an optimization technique). These homogeneous clusters are consistent with the relative neighbourhood stabilities in crime rates noted in Bottoms and Wiles (1986), among others.

Spatial Statistical Analysis

Statistical analysis continues to be a popular and important quantitative tool for analysing the occurrence of crime. Given the spatial nature of crime, there are certainly potential problems associated with the use of classical statistical techniques (see Griffith and Amrhein 1997). Thus, spatial statistical analysis is particularly important for exploring crime in urban regions. Anselin and Bao (1997), Zhang and Griffith (1997) and Anselin (1998) have detailed a number of spatial statistical techniques for exploring spatial information, particularly within a GIS environment.

Building upon the previous sections, we will continue to explore the spatial distribution of property crime. In addition we will also examine the following power transformation of the property crime per 1,000 residents variable:

$$\text{new-variable} = (\text{property-crime})^{0.195}$$

This transformation is the closest approximation to a normally distributed variable. The analysis which follows suggest that both variables (property crime per 1,000 residents and its power transformation) are informative in the statistical analysis process. Along with ArcView, S-Plus version 4.0 and SpaceStat version 1.90 (Anselin 1999) are utilized for statistical analysis and exploration. Both of these commercial packages have capabilities for spatial statistical analysis linked directly or indirectly with GIS (using the S+SpatialStats module with S-Plus and the ArcView extension with SpaceStat). Results presented in this paper are generated using SpaceStat and its ArcView extension (Anselin and Smirnov 1998) unless otherwise noted.

One of the critical items of information in carrying out spatial statistical analysis is the spatial weights matrix, which reflects the spatial influence of neighbours. One approach for constructing a spatial weights matrix is to base it on connectivity, where a row and column entry of one in the square matrix corresponds to the row and column suburbs being neighbours. This may be based upon two suburbs sharing a common border or that two suburbs are within some threshold distance of each other. In this section we have utilized the queen criterion for defining connectivity between suburbs (any two suburbs sharing at least one point in common are considered neighbours). Although no results will be discussed or presented, we also evaluated results using contiguity defined as suburbs being within 7 and 10 kilometres of each other. All weights matrices were row standardized in SpaceStat.

There are essentially two ways of analysing the distribution of property crime in Brisbane using spatial statistical techniques. One is regionally (or globally) and the other is locally (Zhang and Griffith 1997). From a global perspective, a useful technique is the box map (see Anselin 1998) shown in Figure 7, which depicts the quartiles and associated outliers (1.5 times the interquartile range above the third quartile and 1.5 times the interquartile range below the first quartile). Figure 7a depicts the box map for property crime per 1,000 residents, whereas Figure 7b illustrates the box map for the power transformation of the property crime variable. One major difference between Figures 7a and 7b is that there are no lower outliers in Figure 7a. Thus, Figure 7b is what we would expect from an appropriately transformed variable. The patterns shown in Figure 7 are not too different from those shown in Figure 2, which is expected, but the inclusion of outliers in Figure 7 is particularly valuable. In some sense we are actually provided with more relationship characterization in Figure 7 as compared to both Figures 2 and 5, especially Figure 7b, due to the representation of lower outliers and fewer upper outliers.

Another globally based summary is the Moran's I statistic, which indicates the degree of spatial association between areas (see Griffith and Amrhein 1997). In this case, Moran's I provides some insight regarding the global level of spatial autocorrelation between suburbs in terms of observed rates of property crime (essentially testing for the random patterning of similar values in suburbs). The Moran's I statistic for property crime was 0.102 with a standard normal z-value of 2.37 ($p<0.02$) and the Moran's I statistic for the power transformation of property crime is 0.152 with a standard normal z-value of 3.49 ($p<0.0005$). This is indicative of positive spatial autocorrelation associated with the distribution of property crime in Brisbane, which means that we would expect to see a certain degree of grouping of suburbs with like property crime rates. In fact, the figures presented thus far support this to some extent for the suburbs around the CBD and perhaps around Bellbowrie (see Figures 2 and 6). More detailed analysis is required in order to confirm this assertion.

Local approaches for analysing spatial association attempt to identify whether there are any spatial patterns of similar occurrence rates in the spatial distribution of property crimes. One local approach is the Moran scatterplot (Anselin and Bao 1997; Zhang and Griffith 1997). The Moran scatterplot characterizes suburbs using their own crime rates in relation to crime rates in neighbouring suburbs (based upon the specification of connectivity in the weights matrix).[2] Figure 8 depicts the Moran scatterplots for property

[2] Technically the Moran scatterplot utilizes a standardized variable, y, plotted against its spatial lag, Wy, where W is the contiguity weights matrix (Anselin and Bao 1997). This places observations in one of the four planar graph quadrants centred at (0,0). Given this, an observation may be characterized as being a high value around high values (upper right quadrant), a low value around low values (lower left quadrant), a high value around low values (lower right quadrant), or a low value around high values (upper left quadrant). This characterization may then be (and is) illustrated in a map-based format.

Fig. 7 Box map of property crime: (a) property crimes per 1,000 residents, (b) power transformation of (a)

Fig. 8 Moran scatterplot map of property crime: (a) property crimes per 1,000 residents, (b) power transformation of (a)

crime and its power transformation. Figure 8a is rather effective at distinguishing between the significantly high property crime rates in Sumner, Heathwood and Eagle Farm in relation to surrounding suburbs. In contrast, Figure 8b using the power transformation of property crime does not illustrate the noteworthy differences in these suburbs. Certainly the inner city stands out due to neighbouring suburbs with relatively high property crime rates (High-High) as does the grouping of high rates south of Sumner. A particularly interesting relationship depicted in Figure 8b is that the low crime rate suburbs surrounded by low crime rate suburbs (Low-Low) correspond well with middle class/elite areas.

Another local approach is the use of LISA (local indicators of spatial association) statistics (Anselin and Bao 1997). The local Moran LISA statistic is utilized here for analysis and assessed using a normal distribution approximation. The local Moran statistic is analogous to the global Moran's I statistic, but suburb specific in this case. This is obviously useful for identifying spatial patterns or groupings. Figure 9 identifies suburbs with significant local Moran statistics. In Figure 9a the local Moran statistic using property crime per 1,000 residents is shown. The significant suburbs are around the Brisbane CBD and each has a positive value, which supports the previous assertion of positive spatial autocorrelation in the city centre. The significant local Moran values using the power transformation of property crime rates are presented in Figure 9b. There are certainly more significant suburbs in Figure 9b as compared to Figure 9a, but only the inner city area represents a grouping of high rates of property crime. The other significant suburbs in Figure 9b correspond to a clustering of low rates (Anstead and Burbank) and area outliers such as Heathwood and Rochedale (significantly higher property crime rates than surrounding suburbs) as well as Brisbane Airport (significantly lower than surrounding suburbs). Much of this patterning (Brisbane CBD, Anstead and Burbank) is also present using the New Gi* statistic (another LISA statistic option available in SpaceStat—see Anselin 1999), but these results will not be presented here.

Discussion and Conclusions

A notable omission in the analysis presented in the previous sections is the use of linear regression, which is probably the most widely applied technique for analysing information and relationships. In fact, quantitative studies of criminal activity typically utilize linear regression in their analysis (e.g. Brown 1982). It is important in spatial applications of statistical techniques (including multivariate linear regression) that the analysis is sensitive to the assumptions upon which it is based (Griffith and Amrhein 1997). The analysis of crime in Brown (1982) was particularly thorough in this respect. The issues of importance in linear regression are spatial lag and error dependencies as well as heteroskedasticity[3] (Griffith and Amrhein 1997; Anselin 1999). One feature of SpaceStat is that it facilitates the process of carrying out detailed regression analysis using spatial information by providing a range of diagnostics and alternative regression model

[3] Heteroskedasticity is a lack of constant variance in the linear regression error term.

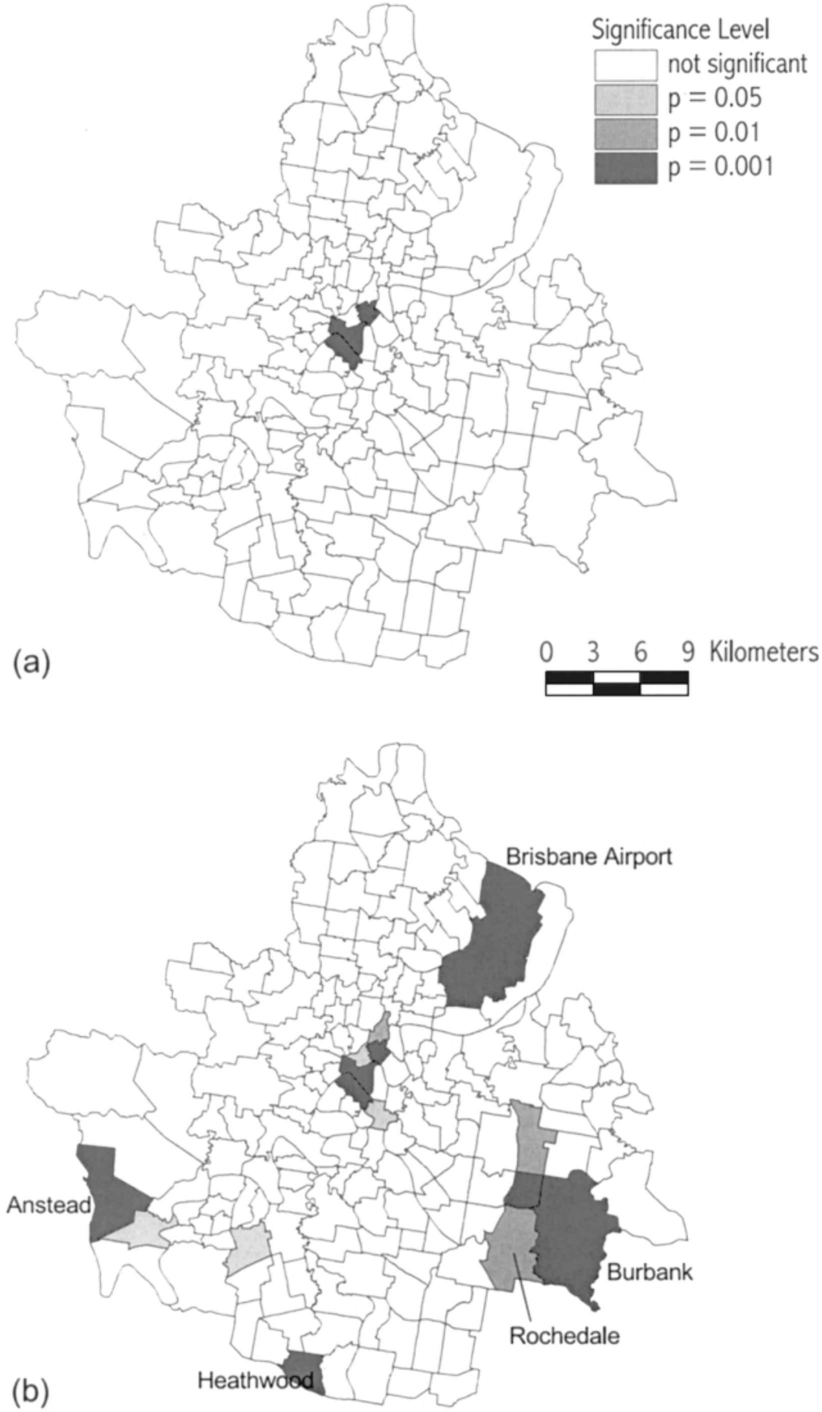

Fig. 9 Local Moran statistic for property crime: (a) property crimes per 1,000 residents, (b) power transformation of (a)

specifications. Although the presentation of multivariate regression results associated with property crime in Brisbane is beyond the scope of this paper, it is worth mentioning that the spatial variables generated, using the detailed GIS functionality discussed previously, are in fact quite significant. For property crime per 1,000 residents significant spatial variables included:

– Density of public transport stops in suburbs.
– Distance to closest police station.
– Distance to closest ferry platform.
– Distance to the Brisbane River.

For the power transformation of property crime per 1,000 residents significant spatial variables included:

– Number of public transport stops in a suburb.
– Distance to closest bus stop.
– Distance to a rail line.
– Distance to the Brisbane River.

These are interesting variables because of their potential to provide insights into social and environmental processes influencing or influenced by criminal activity. Coupling these environmental structures with socio-economic, demographic and other land use characteristics not only ensures that omitted variables are less likely, but also increases the explanatory power of linear regression (provided that assumptions are properly dealt with).

One message emphasized throughout this paper has been that GIS is more than a mapping tool in the analysis of crime. This is in contrast to the impression that one gets from crime mapping research, perhaps unjustly. The basic capabilities developed (and provided) in GIS for information integration are quite powerful. For example, the ability to internalize information, such as relationships to the physical environment, is essential for studying patterns of criminal activity. Further, visualizing crime occurrence in relation to various selected spatial layers of information using adjustable spatial scales is not only valuable, but indispensable. This certainly is suggested in the summaries depicted in Figures 2–5.

Another point stressed in this paper is that current commercial GIS packages provide the linkage capabilities for integrating sophisticated exploratory techniques for analysing the spatial distribution of crime. One approach shown was the use of a spatial clustering technique (Figure 6). This technique makes use of attribute information (e.g. crime rates) as well as explicit spatial proximity between suburbs. An additional approach was the use of spatial statistical methods to identify outliers and significant groupings of similar suburbs (Figures 7–9). Both of these approaches were integrated in some way with a commercial GIS package.

Understanding the dynamics of crime is particularly important to regional planners, policy makers, and policing agencies. Geographic information systems (GIS) and spatial analysis techniques are increasing what we know about criminal activity. Further, they are making the response time and analysis effort more efficient. This paper has discussed and illustrated many of the most relevant capabilities of GIS coupled with recent spatial analysis approaches for examining the distribution of crime in Brisbane, Australia. These techniques will undoubtedly enhance monitoring and policing capabilities.

References

Anselin, L. (1998). 'Interactive Techniques and Exploratory Spatial Data Analysis', in P. Longley, M. Goodchild, D. Maguire and D. Rhind, eds., *Geographical Information Systems: Principles, Techniques, Management and Application*, 251–64. New York: John Wiley & Sons.

——(1999), *SpaceStat: A Software Program for the Analysis of Spatial Data*, version 1.90. Ann Arbor: SpaceStat.

Anselin, L. and Bao, S. (1997), 'Exploratory Spatial Data Analysis Linking SpaceStat and ArcView', in M. Fischer and A. Getis, eds., *Recent Developments in Spatial Analysis*, 35–59. Berlin: Springer-Verlag.

Anselin, L. and Smirnov, O. (1998), *The SpaceStat Extension for ArcView*. Allen, TX: SpaceStat.

Bottoms, A. and Wiles, P. (1986), 'Housing Tenure and Residential Community Crime Careers in Britain', in A. Reiss and M. Tonry, *Communities and Crime*, 101–62. Chicago: University of Chicago Press.

Brantingham, P. and Brantingham, P. (1981), *Environmental Criminology*. Beverly Hills: Sage.

Brown, M. (1982), 'Modelling the Spatial Distribution of Suburban Crime', *Economic Geography*, 58: 247–61.

Buck, A., Deutsch, J., Hakim, S., Spiegel, U. and Weinblatt, J. (1991), 'A Von Thunen Model of Crime, Casinos and Property Values in New Jersey', *Urban Studies*, 28: 673–86.

Byrne, J. and Sampson, R. (1986), *The Social Ecology of Crime*. New York: Springer-Verlag.

Chrisman, N. (1997), *Exploring Geographic Information Systems*. New York: John Wiley & Sons.

Cromley, R. (1996), 'A Comparison of Optimal Classification Strategies for Choroplethic Displays of Spatially Aggregated Data', *International Journal of Geographical Information Systems*, 10: 405–24.

Dent, B. D. (1990), *Cartography: Thematic Map Design*, 2nd ed. New York: WC Brown.

Fattah, E. (1997), *Criminology: Past Present and Future*. Basingstoke, England: Macmillan.

Greenburg, S. and Rohe, W. (1984), 'Neighborhood Design and Crime', *Journal of the American Planning Association*, 50: 48–61.

Griffith, D. and Amrhein, C. (1997), *Multivariate Statistical Analysis for Geographers*. New Jersey: Prentice-Hall.

Harries, K. (1980), *Crime and the Environment*. Springfield, IL: Charles C. Thomas.

Hirschfield, A., Brown, P. and Todd, P. (1995), 'GIS and the Analysis of Spatially-Referenced Crime Data: Experiences in Merseyside, UK', *International Journal of Geographical Information Systems*, 9: 191–210.

Murray, A. (2000), 'Spatial Characteristics and Comparisons of Interaction and Median Clustering Models', *Geographical Analysis*, 32: 1–19.

Murray, A. and Estivill-Castro, V. (1998), 'Cluster Discovery Techniques for Exploratory Spatial Data Analysis', *International Journal of Geographical Information Science*, 12: 431–43.

Murray, A. and Shyy, T. (2000), 'Integrating Attribute and Space Characteristics in Choropleth Display and Spatial Data Mining', *International Journal of Geographical Information Science*, 14: 649–67.

Murray, A., Davis, R., Stimson, R. and Ferreira, L. (1998), 'Public Transportation Access', *Transportation Research D*, 3: 319–28.

Openshaw, S. (1991), 'Developing Appropriate Spatial Analysis Methods for GIS', in D. Maguire, M. Goodchild and D. Rhind, eds., *Geographical Information Systems: Principles and Applications*, 389–402. New York: Longman.

——(1998), 'A New Human Geographic Research Agenda for GIS', *Environment and Planning A*, 30: 383–4.

Ratcliffe, J. and McCullagh, M. (1998a), 'Identifying Repeat Victimization with GIS', *British Journal of Criminology*, 38: 651–62.

——(1998b), 'Aoristic Crime Analysis', *International Journal of Geographical Information Science*, 12: 751–64.

Smith, D. (1995), 'The New Urban Sociology Meets the Old: Rereading Some Classical Human Ecology', *Urban Affairs Review*, 30: 432–57.

Stimson, R. and Taylor, S. (1999), 'City Profile: Brisbane', *Cities*, 16: 285–95.

Stimson, R., Western, J., Mullins, P. and Simpson, R. (1999), 'Urban Metabolism as a Framework for Investigating Quality of Life and Sustainable Development in the Brisbane-Southeast Queensland Metro Region', in L. Yuan, B. Yuen and C. Low, eds., *Urban Quality of Life*, 143–68. Singapore: National University of Singapore.

Weisburd, D. and McEwen, T., eds. (1997), *Crime Mapping and Crime Prevention*. New York: Criminal Justice Press.

Wright, D., Goodchild, M. and Proctor, J. (1997), 'GIS: Tool or Science?', *Annals of the Association of American Geographers*, 87: 346–62.

Zhang, Z. and Griffith, D. (1997) 'Developing User-Friendly Spatial Statistical Analysis Modules for GIS: an Example Using ArcView', *Computers, Environment and Urban Systems*, 21: 5–29.