The utility of hotspot mapping for predicting spatial patterns of crime

Spencer Chainey[1], Lisa Tompson[2], and Sebastian Uhlig[3]

[1] Director of Geographical Information Science, Jill Dando Institute of Crime Science, University College London, Brook House, 2-16 Torrington Place London. WC1E 7HN
[2] Research Fellow, Jill Dando Institute of Crime Science, University College London, Brook House, 2-16 Torrington Place London. WC1E 7HN
[3] Project Consultant, Informed Solutions, The Old Bank, Old Market Place, Altrincham, Cheshire. WA14 4PA, UK

Abstract

Hotspot mapping is a popular analytical technique that is used to help identify where to target police and crime reduction resources. In essence, hotspot mapping is used as a basic form of crime prediction, relying on retrospective data to identify the areas of high concentrations of crime and where policing and other crime reduction resources should be deployed.

A number of different mapping techniques are used for identifying hotspots of crime – point mapping, thematic mapping of geographic areas (e.g. Census areas), spatial ellipses, grid thematic mapping, and kernel density estimation. Several research studies have discussed the use of these methods for identifying hotspots of crime, usually based on their ease of use and ability to spatially interpret the location, size, shape and orientation of clusters of crime incidents (Eck et al, 2005, Chainey et al 2002, Jefferis, 1999, Ratcliffe and McCullagh, 1999). Yet surprising, very little research has compared how hotspot mapping techniques can accurately predict where crimes will occur in the future.

This research uses crime data for a period before a fixed date (that has already passed) to generate hotspot maps, and test their accuracy for predicting where crimes will occur next. Hotspot mapping accuracy is compared in relation to the mapping technique that is used to identify concentrations of crime events (thematic mapping of Census Output Areas, spatial ellipses, grid thematic mapping, and kernel density estimation) and by crime type - four crime types are compared (burglary, street crime, theft from vehicles and theft of vehicles).

The results from this research indicate that crime hotspot mapping prediction abilities differ between the different techniques and differ by crime type. Kernel density estimation was the technique that consistently outperformed the others, while street crime hotspot maps were consistently better at predicting where future street crime would occur when compared to results for the hotspot maps of different crime types. The research offers the opportunity to benchmark comparative research of other techniques and other crime types, including comparisons between advanced spatial analysis techniques and prediction mapping methods. Understanding how hotspot mapping can predict spatial patterns of crime and how different mapping methods compare will help to better inform their application in practice.

Key words: crime hotspot mapping, crime prediction, Prediction Accuracy Index

1. Introduction

A key component to tackling crime problems involves the analysis of where crimes take place. This is on the basis of recognising that crime has an inherent geographical quality (Chainey and Ratcliffe, 2005), in that crime takes place at some location. Crime also does not occur randomly. It tends to concentrate at particular places for reasons that can be explained in relation to victim and offender interaction and the opportunities that exist to commit crime (Cohen and Felson, 1979, Brantingham and Brantingham, 1984, Cornish and Clarke, 1986). These concentrations or clusters of crime are commonly referred to as hotspots – geographic locations 'of high crime concentration, relative to the distribution of crime across the whole region of interest' (Chainey and Ratcliffe, 2005, p. 147).

Hotspot mapping has become a popular analytical technique used by law enforcement, police and crime reduction agencies to visually identify where crime tends to be highest, aiding decision-making that determines where to target and deploy resources. Its application has been used to support the operational briefing of police patrols (Home Office, 2005, Osborne and Wernicke, 2003, LaVigne and Wartell, 1999, Goldsmith et al, 2000, Harries, 1999, LaVigne and Wartell, 1998, Hough and Tilley, 1998), inform the generation of intelligence products and problem solving (Home Office, 2005, Chainey and Ratcliffe, 2005, Clarke and Eck, 2003), has been used as a tool to capture the measurement and analysis of crime patterns for crime auditing purposes in the UK's crime reduction partnerships (Chainey and Ratcliffe, 2005, Chainey, 2001), and supports performance analysis (for example in CompStat style performance meetings – for examples see Chainey and Ratcliffe, 2005, Home Office, 2005, Schick, 2004, McDonald, 2002, Walsh, 2001, Harries, 1999). In essence, hotspot mapping is a technique that is used to help determine where crime may happen next, using data from the past to inform future actions. In this sense it acts as a basic technique for predicting where crime may occur, using the premise that retrospective patterns of crime are a useful indicator for future patterns.

There are many different mapping techniques that can be used for identifying and exploring patterns of crime. These techniques could be as straightforward as representing each crime event as a point and observing the geographic distribution of these points; utilising functions within a geographical information system (GIS) for thematically shading administrative areas (e.g. Census zones or police beats); or representing the distribution of crime as a continuous surface that relates to the volumetric densities of the geographic distribution of crime.

Many of these mapping techniques have been subject to several reviews that have considered their utility for hotspot mapping (see Jefferis, 1999, Chainey et al, 2002, and Eck et al, 2005). However, these reviews have been little more than visual comparisons of each method or exercises that have evaluated their ease of use. Importantly, these reviews have demonstrated that different hotspot mapping techniques may produce different results in terms of the location, size and shape of areas that are defined as hotspots, but to date, none of these reviews have determined which of these techniques is best for helping to determine where spatial patterns of crime may occur in the future. In light of this, we consider that it would be useful to determine if there are differences in the ability between the hotspot mapping techniques to predict spatial patterns of crime. This will help practitioners select the hotspot mapping technique that is most suitable for their application and allow them to determine some level of accuracy in the utility of the technique for predicting future crime patterns. We also seek to determine if the ability to predict spatial patterns of crime differs between crime types.

2. Hotspot mapping techniques

Identifying hotspots is the first step a policing or crime reduction agency needs to take when discerning where best to prioritise their resources. Attempting to do this via point mapping has become outdated since the proliferation of GIS software and the increasing sophistication of mapping techniques. In this section we describe four of the most common hotspot mapping techniques, illustrated as well in figure 1.

2.1. Spatial ellipses

One of the earliest crime mapping software applications that became widely available to practitioners for crime analysis was Spatial and Temporal Analysis of Crime (STAC) (Illinois Criminal Justice Information Authority, 1996). STAC is not a GIS, but instead acts as an aid to persons who already have a GIS or desktop mapping capability. STAC is a spatial tool to find and examine hotspot areas within the study area. In concise terms this means that STAC first finds the densest concentration of points on the map (hot clusters), and then fits a 'standard deviational ellipse' to each one. The ellipses themselves indicate through their size and alignment the nature of the underlying crime clusters.

Examples demonstrating the use of STAC include Martin et al (1998) in their study of how to reduce incidents in Detroit's infamous Devil's Night period; Bowers and Hirschfield (1999) who explored links between crime and disadvantage in North West England; Block and Block (2000) when analysing hotspots around rapid transit stations in Chicago, and Langworthy and Jefferis (2000) who examined the influence of the school holiday period on the spatial distribution of crime hotspots in the Bronx, New York. Baltimore County has also used STAC extensively to analyse a range of priority crime types (Block and Perry, 1993).

Reported benefits of using STAC include that it derives hotspots without relying on defined boundaries such as Census units or police administrative boundaries; requires few parameters; and is compatible with most GIS applications (Martin et al., 1998). However, STAC has attracted criticism for several reasons; firstly, it is preferable for the user to be well versed in the routines at work within the software. For the novice, there is little counsel on appropriate parameter values leading to the introduction of ambiguity and increasing variability in the results (Eck et al, 2005). Secondly, crime hotspots do not naturally form into convenient ellipses, thus STAC hotspots do not represent the actual spatial distribution of crime and can often mislead (Eck et al, 2005, Ratcliffe and McCullagh, 2001). Finally, the visualisation of the STAC produced results negates any comparison with events that do not fall into the spatial ellipses (Eck et al, 2005). An example of STAC produced spatial ellipses are shown in figure 1b.

### 2.2. Thematic mapping of geographic boundary areas

A widely used way of representing spatial distributions of crime events is geographic boundary thematic mapping, or choropleth mapping (Home Office, 2001). Boundary areas that are used for this type of thematic mapping are usually arbitrarily defined for administrative or political use, for example they can be police beats, census blocks, wards or districts. Offences as points on a map can be aggregated to these geographic unit areas which can then be shaded in accordance with the number of crimes that fall within them. Williamson et al (2001) assert that maps created in this way are quick to produce and require little technical expertise to interpret. Furthermore, this technique enables the user to quickly determine which areas have a high incidence of crime, and allows further diagnosis of the problem by 'zeroing in' on

those areas. In addition, Census areas can easily be linked with other data sources, such as population, to calculate a crime rate - increasing their versatility for analysis.

Due to the varying size and shape of most geographical boundaries, thematic shading can beguile the map reader in identifying the existence of the highest crime concentrations (Eck et al, 2005). Hence this technique can fail to reveal patterns across and within the geographical division of boundary areas (Chainey and Ratcliffe, 2005). Also, as with all mapping reliant on defined geographical boundaries, the problem of the Modifiable Areal Unit Problem (MAUP - Openshaw, 1984) produces further complications. This is where changes in the boundaries themselves can directly affect the patterns shown on the map.

Thematic mapping of boundary areas continues to see widespread application, from being used for comparing the different volumes of unique and repeat burglaries across a study area's census zones (Ratcliffe and McCullagh, 2001), to comparing vehicle theft in relation to land use in Overland Park, Kansas (Harries, 1999, p47), and for the analysis and presentation of crime patterns and auditing across partnership administrative zones (Home Office, 2005; Chainey, 2001). An example of a thematic map, generated for Census Output Areas, is shown in figure 1c.

(a)   (b)

(c)   (d)
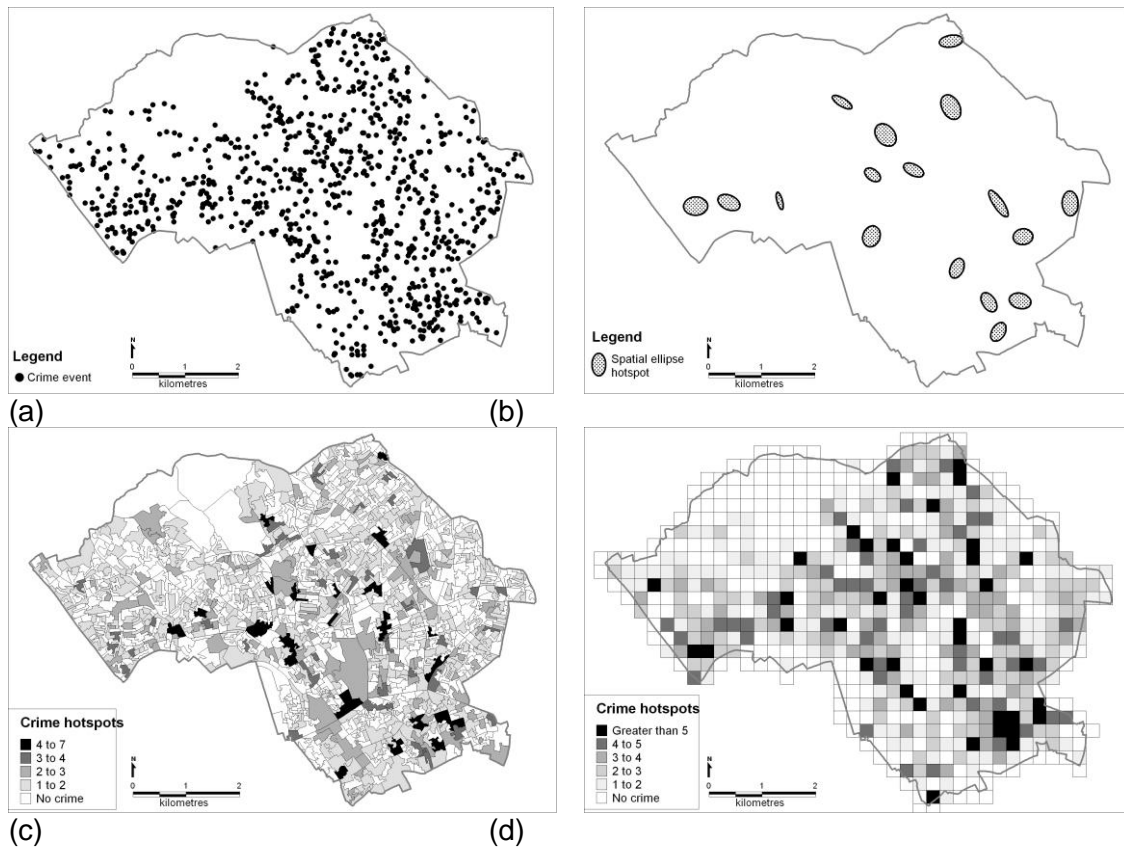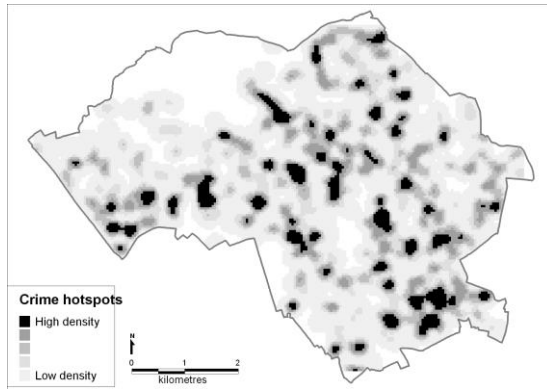
(e)

Figure 1. Common hotspot mapping techniques. (a) Point mapping, (b) standard deviational spatial ellipses, (c) thematic mapping of administrative units, (d) grid thematic mapping, and (e) kernel density estimation

2.3.   Grid thematic mapping

In order to combat the problems associated with different sizes and shapes of geographical regions, uniform grids (or quadrats) can be drawn in a GIS as a layer over the study area and thematically shaded. Therefore all areas used for thematic shading are of consistent dimensions and are comparable, assisting the quick and easy identification of hotspots. Bowers et al (2001) used this method as a component of a GIS-based database application set up to identify vulnerable residences where target hardening was then implemented. LeBeau (2001) also found this technique useful when mapping the volume of emergency calls and violent offences per square mile in North Carolina.

This approach does have some limitations; the usage of grids still restricts how the hotspots can be displayed.   Spatial detail within and across each quadrat is correspondingly lost because the crime events have to conform to one specific quadrat, which can then lead to inaccurate interpretation by the map user. Additionally, many comments have been made about the 'blocky' appearance of this technique (Chainey and Ratcliffe, 2005; Eck et al, 2005; Home Office, 2001), which is affected by grid cell size. The solution, reducing the size of each cell, can destroy the resolution of the thematic map by making it look 'specklely' and can fail to provide any useful information about where crime clusters (Chainey and Ratcliffe, 2005). Finally, grid thematic mapping suffers from the same MAUP problems outlined above (Bailey and Gatrell, 1995). An example of a grid thematic map is shown in figure 1d.

2.4.   Kernel density estimation

Kernel density estimation (KDE) is regarded as the most suitable interpolation technique for visualising crime data (Chainey and Ratcliffe, 2005; Eck et al, 2005; Chainey et al, 2002; Williamson et al 2001; McGuire and Williamson, 1999; Williamson et al, 1999). This is an increasingly popular method due to its growing availability (such as the MapInfo add-on Hotspot Detective), the perceived accuracy of hotspot identification, and the aesthetic look of the resulting map in comparison to other techniques. Point data (offences) are aggregated within a user-specified search radius and a continuous surface that represents the density or volume of crime events across the desired area is calculated. A smooth surface map is produced, showing the variation of the point/crime density across the study area, with no need to conform to geometric shapes such as ellipses. There is flexibility when setting parameters such as the grid cell size and bandwidth (search radius); however despite many useful recommendations (see Chainey and Ratcliffe, 2005; Eck et al, 2005, Ratcliffe 1999),

5

there is no universal doctrine on how to set these and in what circumstances. Examples of the use of KDE are now widespread, with popular crime mapping texts showing many examples of its use (see Chainey and Ratcliffe, 2005, Eck et al, 2005, Goldsmith, et al, 2000, Harries 1999).

KDE is not without its faults. Eck et al (2005) highlight that the choice of thematic range to use still presents itself as a problem as agencies fail to question the validity or statistical robustness of the produced map, instead being caught in its 'visual lure'. This largely affects how hotspots are identified and increases the variation of maps fashioned from the same data. There are also concerns that small amounts of data can misinform the map reader (Home Office, 2001). Nevertheless, the KDE technique is currently in vogue, not only because it is the most visually impactive but also because it has the capability of identifying hotspots through a statistically robust methodology (Chainey and Ratcliffe, 2005; Eck et al, 2005; Chainey et al, 2002; Williamson et al, 1999). An example of a KDE crime hotspot map is shown in figure 1e.

3. Research objectives

This research had two main objectives:

- To identify if the accuracy of hotspot maps for predicting spatial patterns of crime differs between types of crime.

To date, no research has been conducted that measures if a hotspot map that is based on one type of crime is equally as accurate as another hotspot map that is based on a different type of crime for determining where crime (of that type) may occur in the future. For example, those that produce hotspot maps typically assume there to be no difference between a map they generate for showing where burglaries to residential properties occurs to one they generate to show hotspots of thefts of vehicles. That is, they assume each map to be equally as good, without questioning any differences in each map's ability to determine future spatial patterns of crime. This research explores if hotspot maps of certain crime types differ in their ability to predict where crime may occur in the future.

- To perform analysis that compares the differences in the ability of hotspot mapping techniques for determining future spatial patterns of crime.

The hotspot mapping techniques that were explored in this research were ones that are most commonly used by those that generate hotspot maps. We recognise that new techniques that are specifically designed to predict geographic crime patterns are emerging (see Chainey and Ratcliffe, 2005, and Groff and LaVigne 2002 for a summary of these techniques), but these techniques are still someway off as being tools that are commonly used by law enforcement, police and crime reduction practitioners. This research aims to identify if there are differences between certain common hotspot mapping techniques in their ability to predict where crime may occur in the future.

4. Data and methodology

Methodological considerations for this research were centred on how best to explore the range of mapping techniques with data that was suitable for this analysis, using an approach that was consistent (to enable comparisons between results) and typical of that which would be applied by practitioners. The creation of hotspots and the analysis

of their prediction ability were conducted in MapInfo Professional mapping software and included the use of the add-on programme 'Hotspot Detective' (Ratcliffe 2002). CrimeStat software (Levine 2004) was also utilised as this included an up-to-date version of STAC - one of the mapping techniques that was used in this research to create hotspot maps.

4.1. Study area

The area under study is located in Central/North London, and consists of two adjacent local government districts - Camden and Islington within the Metropolitan Police Force area. This area encompasses a wide range of urban geography, including three mainline train stations (Euston, Kings Cross and St. Pancras), a Premiership football stadium (Arsenal FC), popular shopping areas such as Tottenham Court Road, Holborn, Angel and Camden Market and large open parks such as Hampstead Heath. The area contains a synthesis of different land uses (commercial, retail and residential), and a resident population that reflects London's cosmopolitan diversity. The area experiences a high influx of visitors who commute to the area for work, education, are shoppers or tourists or visit the area for its theatres, cinemas, restaurants, bars, music and other attractions. Figure 2 shows a map of the study area with some supporting general statistics about the area.
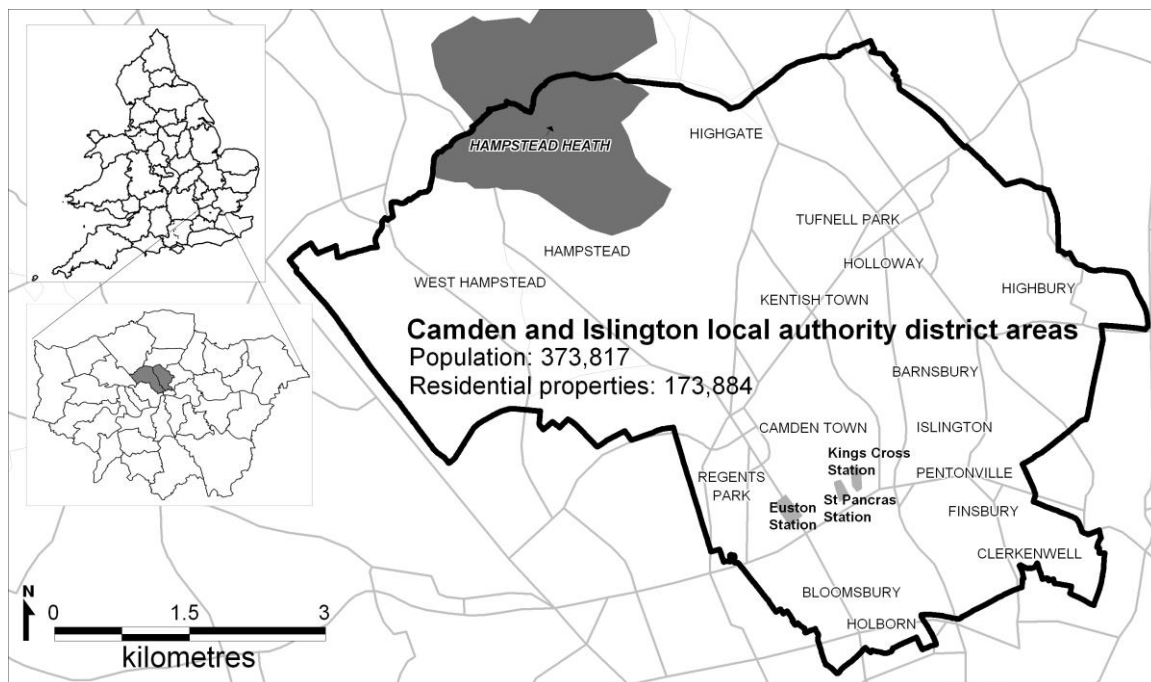

Figure 2. The Camden and Islington study area in Central/North London.

4.2. Crime data

Geocoded crime point data was provided by the Metropolitan Police for a two year period (1st January 2002 to 31st December 2003). For the purposes of applying anonymity to crime data released for academic research, the Metropolitan Police rounded the Easting (x) and Northing (y) coordinates to the nearest ten metres, meaning that a crime could have been displaced by up to five meters in any direction of the original location. It is acknowledged that this reduces the spatial accuracy of the point data; however the small distances involved are thought not to impact the analysis performed in this research. This geocoded crime data was then validated by the researchers conducting this study. This was performed using a methodology for geocoding accuracy analysis as reported in Chainey and Ratcliffe, 2005, p 61-63. This

revealed the crime data to be more than 95% accurate and fit for purpose for this research.

The crime data was then grouped to form four crime types; residential burglary (burglary to domestic properties), street crime (combining robbery of personal property and theft from the person), thefts from vehicles and thefts of vehicles. These crime types were chosen because they are groupings that are regularly analysed by police and crime reduction practitioners; therefore the implications of the research would be accessible and could be more readily translated into policing and crime reduction practice. Table 1 lists each of the crime types and the number of crimes for each calendar year period.

Time period	Residential burglary	Street crime	Theft from vehicle	Theft of vehicle	Total
2002	6300	5249	10792	4243	26584
2003	5671	4911	11536	4142	26260
Total	11971	10160	22328	8385	52844

Table 1. Number of crime events, by type, for each calendar year.

A suitable date had to be chosen within the data time period as the day on which retrospective data was selected to generate hotspot maps against which 'future' events could be compared (we refer to this date as the 'measurement date'). For simplicity, the 1st January 2003 was selected in order to maximise the use of 12 months of retrospective data for generating hotspot maps, and to use the complete set of 12 months of data after this date for measuring the hotspot maps' abilities for predicting future events. The 1st of January is though a rather unusual date in the calendar when the normal routines of people's day to day lives could be quite different to most other days in the year. The three-week period before the 1st of January could also be considered as unusual as it is the period before and during Christmas, when again the routine activities of peoples' lives who live or visit Camden and Islington could be quite different to other dates in the year. For this purpose we also selected the 13th March 2003 as a measurement date. This date was a Thursday, was during school term time and was considered as a normal date on which most people would go about their routine activities in a manner that was similar to many other dates in the year. By choosing two measurement dates this meant that we could compare results for their consistency.

The retrospective time data was sliced into ten time periods and used as 'input data' to generate hotspot maps. This meant that rather than using just one retrospective time period for the research (e.g. the three months prior to the measurement date) which may generate an anomalous result, the use of a number of retrospective time periods could be considered together and would form a more reliable basis on which to draw findings. Retrospective input data was sliced into the time periods shown in table 2, with data for each crime type, for each of the mapping techniques being used to create hotspot maps.

2002 time periods used to create crime hotspots for 01/01/03									
12 months	6 months	3 months	2 months	1 month	2 weeks	1 week	3 days	2 days	1 day
01Jan02-31Dec02	01Jul02 - 31Dec02	01Oct02-31Dec02	01Nov02-31Dec02	01Dec02-31Dec02	18Dec02-31Dec02	25Dec02-31Dec02	29Dec02-31Dec02	30Dec02-31Dec02	31 Dec02

(a)

2002 time periods used to create crime hotspots for 13/03/03

12 months	6 months	3 months	2 months	1 month	2 weeks	1 week	3 days	2 days	1 day
13Mar02-12Mar03	13Sep03-12Mar03	13Dec02-12Mar03	13Jan03-12Mar03	13Feb03-12Mar03	27Feb03-12Mar03	06Mar03-12Mar03	10Mar03-12Mar03	11Mar03-12Mar03	12 Mar03

(b)

Table 2. The temporal slices of input data for generating hotspot maps, for (a) a measurement date of the 1st January 2003 and (b) a measurement date of the 13th March 2003. Crime input data was arranged into ten time periods.

If hotspot maps are generated to help determine where crimes may happen next, the definition of 'next' also needs consideration. For the purposes of this research we were limited to prospecting future crime patterns up to the 31st December 2003. This meant that twelve months of crime data could be used when the measurement date was the 1st January 2003, and nine and a half months when the measurement date was the 13th March 2003. In following a similar argument to the temporal slicing of input data, data after the measurement date (which we refer to from this point as 'measurement data') was sliced into the same arrangement of time ranges that were applied to the input data. This meant that rather than using just one measurement data period for the research (e.g. the three months after the measurement date), the use of a number of measurement data time periods would generate results that could be considered together, from which more reliable conclusions could be made. Measurement data was sliced into the time periods shown in table 3. This meant that hotspot maps that were generated for each period of input data would be measured for their ability to predict spatial patterns of crime, when the prediction period could be the next day, the next two days, the next week, and to the next 12 months.

2003 time periods used to determine the Prediction Accuracy Index for 01/01/03									
1 day	2 days	3 days	1 week	2 weeks	1 month	2 months	3 months	6 months	12 months
01Jan03	01Jan03 - 02Jan03	01Jan03 - 03Jan03	01Jan03 - 07Jan03	01Jan03 - 14Jan03	01Jan03 - 31Jan03	01Jan03 - 29Feb03	01Jan03 - 31Mar03	01Jan03 - 30Jun03	01Jan03 - 31Dec03

(a)

2003 time periods used to determine the Prediction Accuracy Index for 13/03/03									
1 day	2 days	3 days	1 week	2 weeks	1 month	2 months	3 months	6 months	All of 2003
13Mar03	13Mar03 - 14Mar03	13Mar03 - 15Mar03	13Mar03 - 19Mar03	13Mar03 - 26Mar03	13Mar03 - 12Apr03	13Mar03 - 12May03	13Mar03 - 12Jun03	13Mar03 - 12Sep03	13Mar03 - 31Dec03

(b)

Table 3. The temporal slices of measurement data for calculating the ability of hotspot maps to predict spatial patterns of crime, for (a) a measurement date of the 1st January 2003 and (b) a measurement date of the 13th March 2003. Note that the maximum time range of measurement data for analysis based on the 13th March 2003 measurement date was nine and a half months and not twelve months.

4.3. The Prediction Accuracy Index

In recent years, researchers have begun to consider measures that can be used to assess how good a technique is for predicting where crime may occur. One obvious measure would be a hit rate – the percentage of new crimes that occur within the areas where crimes are predicted to occur. While useful and easy to understand, this

measure does not take into account the size of the areas where crimes are predicted to occur. For example, a hit rate could be 100%, but the area where crimes are predicted to occur could cover the entire study area – a result of little use to practitioners who have the need to identify where to target resources. Bowers et al (2004) proposed the Search Efficiency Rate. This measure is the number of events per square kilometre in the areas where crimes are predicted to occur.   This measure works well when considering just a single study area, but does not easily allow for comparisons between study areas of different sizes. This is because that in a calculation of how good a technique may be for predicting spatial patterns of crime, the size of the entire study area should be considered in relation to the areas that have been determined as where crimes are predicted to occur. For example, a study area that is 10km² in area may have determined certain areas where crimes are predicted to occur from which a Search Efficiency Rate of 20 crimes per km² has been calculated. A study area that is 50km² in size may have experienced the same volume of crime as the smaller study area and also have the same Search Efficiency Rate of 20 crimes per km² for the areas where crimes are predicted to occur. Yet in the larger study area there is more space where no crime has been predicted to occur, meaning that the predicted areas that have been identified cover a smaller relative area than the predicted areas determined in the smaller study area, and provide a more useful basis from which to target resources, relative to the entire study area's size.

In this research we introduce the Prediction Accuracy Index (PAI). This index has been devised to consider the hit rate against the areas where crimes are predicted to occur with respect to the size of the study area. The PAI is calculated by dividing the hit rate percentage (the percentage of crime events for a measurement data time period falling into the areas where crimes are predicted to occur determined from input data i.e. the crime hotspots) by the Area Percentage (the percentage area of the predicted areas (the hotspots) in relation to the whole study area (see Equation 1)).

$$\frac{\left(\frac{n}{N}\right)*100}{\left(\frac{a}{A}\right)*100} = \frac{HitRate}{AreaPercentage} = \text{Prediction Accuracy Index}$$

n – number of crimes in areas where crimes are predicted to occur (e.g. hotspots)
N – number of crimes in study area
a – area (e.g. km²) of areas where crimes are predicted to occur (e.g. area of hotspots)
A – area (e.g. km²) of study area

Equation 1. Prediction Accuracy Index equation.

Finding 100% of future events in 100% of the area would give a PAI value of 1. If the Hit Rate and the Area Percentage fall by an equal measure, the value would be computed as 1 also. Finding 25% of future crime events in 50% of the study area would return a PAI value of 0.5; and finding 80% of future crime events in 40% of the area would return a PAI value equal to 2. Thus, the greater the number of future crime events in a hotspot area that is smaller in areal size to the whole study area, the higher the PAI value.

The PAI is easy to calculate, considers the number of crimes that fall into the area determined as hotspots against the size of the hotspot and the size of the study area. We also believe it is a measure that is applicable to any study area, any crime point data, and to any analysis technique that aims to predict spatial patterns of crime. Many practitioners still though find great use in the Hit Rate as a measure to predict how

many crimes they may be able to impact upon by targeting resources to just the hotspot areas. As this measure is calculated as part of the PAI, the Hit Rate can be considered alongside a PAI calculation.

To show if the prediction ability of hotspot maps for different crime types differs, Prediction Accuracy Index measures were aggregated and averaged across the hotspot mapping techniques, for the periods of input data and for the periods of measurement data. This meant that the PAI measures for each crime type could be compared, with any differences being explained in relation to the crime type rather than from the influence of mapping technique, input and measurement data. By following this approach, this equated to averaging,

- 10 PAI measures calculated for each crime type based on the period of input data
- 10 PAI measures calculated for each crime type based on the period of measurement data
- 7 PAI measures calculated for each crime type from the hotspot mapping method that was used (the hotspot mapping techniques are described in section 4.5)

A similar approach was applied for comparing the hotspot mapping techniques; Prediction Accuracy Index measures were aggregated and averaged across the different crime types, for the periods of input data and for the periods of measurement data. This meant that the PAI measures for each hotspot mapping technique could be compared, with any differences being explained in relation to the mapping technique rather than from the influence of crime type, input and measurement data. By following this approach, this equated to averaging,

- Four sets (i.e. the four different crime types) of 10 PAI measures calculated for each hotspot mapping technique based on the period of input data for each crime type
- Four sets of 10 PAI measures calculated for each hotspot mapping technique based on the period of measurement data for each crime type

That is, the PAI measures presented in the results are an average of numerous individual PAI calculations. The standard deviation of the PAI for each crime type across the seven hotspot mapping technique measures was also calculated.

4.4.   Impact of changes in crime patterns on the results

During the data time period (1st January 2002 to 31st December 2003) there could have been police operations and crime reduction initiatives that had an impact on crime. For this study we were comparing analytical techniques against the same data, so any changes in crime patterns would be similarly applied to each of the techniques and would not affect the ability of making comparisons and drawing conclusions on the analyses that were conducted in this research.

4.5.   Hotspot mapping techniques and functional parameters

Four hotspot mapping techniques were chosen and used for generating hotspot maps. These techniques were spatial ellipses, thematic mapping of boundary areas, grid thematic mapping and kernel density estimation. These techniques were chosen as they are the most common techniques used by those generating hotspot maps of crime (Weir and Bangs, 2007). Each of these techniques also requires certain parameters to be set for them to operate or for them to generate a visual map output. The parameters that were set are described below for each technique, with the methodology that was followed being one, from experience, that would be the approach a police or crime reduction analyst would most typically apply.

- **STAC:** CrimeStat (Levine, 2004) was used for generating standard deviational ellipses. The parameters that the user is required to enter are the search radius, the minimum number of points-per-hotspot, the number of standard deviational ellipses used to delineate the hotspots, the number of standard deviations to apply for the generation of spatial ellipses, and the scan type. The triangular scan type was selected due to the irregular road network within the study area (as advised by Levine, 2004). Spatial ellipses were created to one standard deviation. Setting the minimum points-per-hotspot and the number of ellipses proved to be more complex as it is difficult to universally apply a fixed number of points-per-hotspot with the different slices of data using STAC. There is no official direction, therefore a variable minimum points-per-hotspot approach was applied, in which the minimum number of points-per-hotspot producing the highest number of hotspots under 20 was calculated using a trial and error method. This generates ellipses (i.e. hotspots) that reflect the underlying crime events and eliminates those ellipses that could be created to make up the numbers to reach 20.

The only guidance for applying a suitable search radius was experimentation and experience (Levine, 2004). This resulted in the use of three search radii to determine hotspots – 500m, 250m and a search radius equal to the default bandwidth in Hotspot Detective (the rationale for choosing this Hotspot Detective value is explained below in the KDE section). This Hotspot Detective derived parameter determined an alternate search radius size that varied in accordance with the spatial characteristics of the crime input data. The three approaches were called STAC-500, STAC-250 and STAC-HD. Once these parameters had been set and run in CrimeStat, the output was imported into MapInfo for display.

- **Thematic mapping of Output Areas:** This technique requires crime point data to be aggregated to polygonal Census Output Areas to generate a count of crime for each Output Area (Output Areas are the smallest unit of Census geography in England and Wales, each covering approximately 125 households). This was performed using the standard routines in MapInfo. The selection of the thematic range method and values for determining the crime hotspot threshold are explained below as they also apply to grid thematic mapping and KDE.

- **Grid thematic mapping:** MapInfo (and most other GIS software) enables analysts to specify a grid lattice that can be positioned across the study area. The parameter the user must decide is the size of each square grid cell. There is little guidance on which cell size to select; if the cells are too large, the resulting hotspot map will only show coarse geographic patterns of crime, and if too small it can be difficult to discern any spatial patterns from the hotspot map (Chainey and Ratcliffe, 2005). Experimentation and experience are again the best advice, but for novices to hotspot mapping using this technique, Chainey and Ratcliffe (2005) suggest as a starting point to choose a grid cell size that is approximately the distance in the longest extent of the study area, divided by 50. For this research a grid cell size of 250m was chosen (if the Chainey and Ratcliffe approach was chosen it would have suggested a similar cell size of 200m), and for comparability, the calculations for the grid thematic mapping technique were repeated using the bandwidth metric determined from the Hotspot Detective default (the applicability of this technique is discussed in the KDE section below). This determined an alternate grid cell size that varied in accordance with the spatial characteristics of the crime input data.

Once the grid had been calibrated, functions in MapInfo were used to calculate a count of the number of crime points within each grid cell. The selection of the thematic range

method and values for determining the crime hotspot threshold are explained in detail below.

- **Kernel density estimation (KDE):** KDE hotspot maps were generated using Hotspot Detective (Ratcliffe, 2002). This technique requires two parameters to be entered by the user – the cell size and the bandwidth (also commonly referred to as the search radius). Hotspot Detective determines default settings for these parameters after performing an analysis of the input data. Following the default settings is an approach that most analysts take, and indeed are encouraged to take if they are not experts in spatial analysis.

The analysis of the input data for determining default cell size and bandwidth values uses a methodology that considers the spatial characteristics of the input data. KDE cell size is calculated by dividing the shorter side of the minimum bounding rectangle around the crime input data by 150 (Ratcliffe, 2004). Bandwidth selection is more complicated (see Chainey and Ratcliffe, 2005). In Hotspot Detective the calculation of the default bandwidth value is not divulged to users, but it is known to be a function of the shorter side of the minimum bounding rectangle surrounding the crime data, divided by a number that provides a suitable enough cell resolution without requiring a significant number of iterations to generate a representative KDE surface (Ratcliffe – personal communication). Experience with using these defaults suggests they are appropriate in most cases for determining cell size and bandwidth settings applied to crime data. As this approach considers in turn the spatial characteristics of the crime input data (i.e. the bandwidth varies in relation to the volume and spatial distribution of the input data) it was considered as useful to use the bandwidth measure determined for each input dataset of crime data as a search radius parameter for STAC and as a cell size for the grid thematic mapping technique. Table 4 lists the Hotspot Detective cell size and bandwidth default parameters applied to KDE, and also lists the bandwidth parameter with respect to its use for STAC and grid thematic mapping.

Two other parameters that users of KDE may be required to enter are the kernel type and if a weighting attribute is to be applied. The common kernel type is the quartic measure (Chainey and Ratcliffe, 2005, Eck et al, 2005, Chainey et al, 2002, Williamson et al, 1999), and as each retrospective crime event is applied with equal weight, no weighting scheme was applied to the input data in this research.

Hotspot Detective default values for all crime types and input data time periods (cell size / bandwidth) for 01/01/03								
Input data time period	Residential Burglary		Street Crime		Theft from vehicle		Theft of vehicle	
	C.S.	B.	C.S.	B.	C.S.	B.	C.S.	B.
12 months	49	245	49	245	49	245	48	240
6 months	49	245	49	245	49	245	46	230
3 months	48	240	48	240	49	245	46	230
2 months	48	240	48	240	48	250	46	230
1 month	46	230	46	230	46	230	46	230
2 weeks	44	220	45	225	45	225	45	225
1 week	44	220	44	220	45	225	37	185
3 days	38	190	44	220	44	220	34	170
2 days	38	190	31	155	40	200	34	170
1 day	38	190	2	10	40	200	27	135

(a)

| Hotspot Detective default values for all crime types and input data time periods (cell size / bandwidth) for 13/03/03								
Input data time period	Residential Burglary		Street Crime		Theft from vehicle		Theft of vehicle	
	C.S.	B.	C.S.	B.	C.S.	B.	C.S.	B.
12 months	46	230	46	230	46	230	45	225
6 months	45	225	46	230	46	230	45	225
3 months	45	225	44	220	45	225	45	225
2 months	45	225	43	215	45	225	45	225
1 month	43	215	43	215	44	220	44	220
2 weeks	41	205	43	215	42	210	42	210
1 week	39	195	42	210	41	205	42	210
3 days	39	195	39	195	40	200	36	180
2 days	30	150	39	195	40	200	31	155
1 day	25	125	38	190	32	160	27	135

(b)

Table 4. Hotspot Detective KDE default values (C.S. – cell size and B. – bandwidth) for each crime type and each period of input data, for (a) a measurement date of the 1st January 2003 and (b) a measurement date of the 13th March 2003. Bandwidth measures were used to determine alternate search radii for spatial ellipses, and alternate grid thematic mapping cell sizes. For example, a grid thematic map was generated from one month of street crime data, for the 01/01/2003 measurement date, using a grid cell size of 230m.

4.6. Determining a thematic threshold for hotspots

A final parameter to consider in hotspot map generation is a threshold value for determining which areas are 'hot'. For spatial ellipses this is straightforward as it is simply the area drawn by each ellipse. Thematic mapping of Output Areas, grid thematic mapping and KDE produce areal values across a continuous range (e.g. for grid thematic mapping, each grid cell has a value representing the number of crimes located within the cell). A threshold value must be determined that specifies that any values above this threshold can be classified as an area that is a hotspot. The same thematic range approach was applied to each of these three hotspot mapping techniques for purpose of simplicity and consistency in methodology (for more details on thematic mapping ranges applied to crime data see Chainey and Ratcliffe, 2005 and Harries, 1999). Five thematic classes were used and default values generated from using the quantile method in MapInfo were applied to the calculations that thematic mapping of Output Areas, grid thematic mapping and KDE generate. This approach was used as the number of classes falls within the upper and lower settings specified by Dent (1999) and Harries (1999), and the quantile method was chosen because it distributes the data in an approximately equal balance between the classes, resulting in a visually balanced map pattern (Monmonier 1996). This approach is also a common approach that many practitioners apply.

'Hot' was then determined by the top thematic class. Figure 3 illustrates this approach. It shows a hotspot map generated using KDE where the thematic ranges were grouped into five classes and arranged following the quantile range method default in MapInfo. Cells with values in only the top thematic class were then selected, with these areas determined as the hotspots.

(a) Hotspot map

(b) Top thematic class of hotspot map
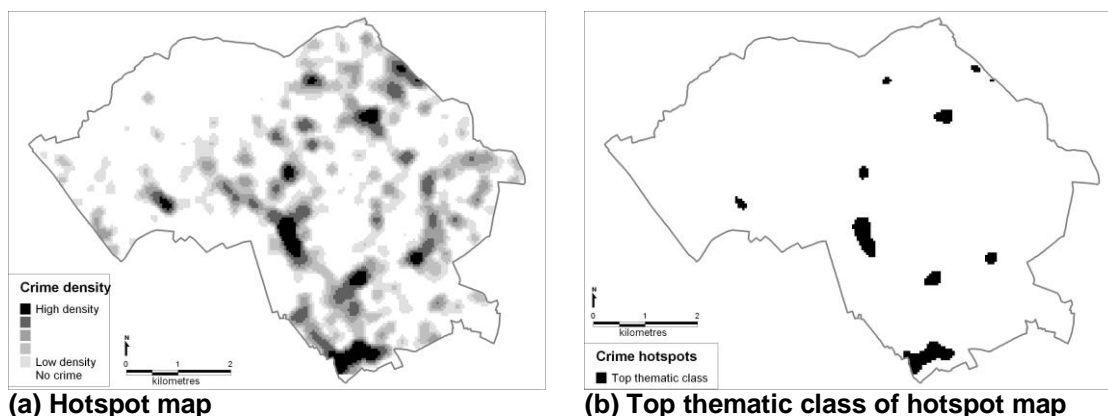
Figure 3. Hotspots were determined by selecting the uppermost thematic class calculated using the five classes and the default values generated from applying the quantile thematic range method in MapInfo.

5. Results

5.1. A comparison between hotspot maps for different crime types for predicting where crimes may occur

Hotspot maps are typically regarded as comparable in their ability to predict future crime patterns regardless of any differences that may exist in the prediction ability between crime types. In this research we have explored if differences exist in the ability to predict future crime patterns from hotspot maps for four groupings of crime; residential burglary, street crime, theft from motor vehicles and theft of motor vehicles. For example, we used retrospective data on residential burglaries to generate hotspot maps, and used these hotspot maps to predict future patterns of residential burglary.

Table 5 shows the PAI results for the four different crime types. These are presented in the table as two sets of results - when the measurement date was the 1st January 2003 and when it was the 13th March 2003. These results show that there are differences between crime types in their ability to predict future patterns of crime. Hotspot maps of street crime consistently produced the highest PAI values, and were clearly higher than PAI values for the other crime types. Theft from vehicles was the crime type that recorded PAI values of the next highest level. PAI values for residential burglary and theft of vehicles were similar to each other. The standard deviation values between the four crime types indicated there to be some degree of variability in the results generated by hotspot mapping techniques. This is further explored below.

Crime type	Average PAI (01/01/2003)	Standard deviation of PAI	Average PAI (13/03/2003)	Standard deviation of PAI
Residential Burglary	1.56	0.39	1.59	0.42
Street Crime	3.24	1.17	3.18	1.07
Theft from Vehicle	1.89	0.41	2.56	0.65
Theft of Vehicle	1.53	0.52	2.14	0.60

Table 5. Prediction Accuracy Index values for residential burglary, street crime, theft from vehicles and theft of vehicles. The higher the PAI, the greater the ability of the hotspot map to predict where future crimes (of the same respective type) will occur. Grey cells indicate the highest PAI values and values in bold italics indicate the lowest PAI values.

These results counter an assumption that many practitioners place in their use of hotspot maps. That is, hotspot maps for different crime types that are generated from retrospective data do differ in their ability to predict future patterns of crime. These results show that hotspot maps of street crime have a greater ability to predict future patterns of street crime in comparison to residential burglary, theft from vehicles and theft of vehicles.

5.2. A comparison of hotspot mapping techniques and their ability to predict where crimes may occur

As previous studies have shown, hotspot mapping techniques differ in the output they generate (see Eck et al, 2005, and Jefferis, 1999). In this research we have explored each of the common hotspot mapping techniques to determine if differences exist in their ability to predict future patterns of crime.

Table 6 shows the PAI results for the different hotspot mapping techniques. These are presented in two columns to show differences between results for the two measurement dates. These results show that there are differences between hotspot mapping techniques in their ability to predict future patterns of crime. Kernel density estimation proved to be the best hotspot mapping technique for predicting where crimes may occur in the future and spatial ellipses appeared to be the worst. Thematic mapping of Output Areas and the thematic mapping of grids produced similar PAI values.

Hotspot mapping technique	Average PAI (01/01/2003)	Average PAI (13/03/2003)
Spatial ellipses 250m	1.74	2.25
Spatial ellipses 500m	1.24	1.52
Spatial ellipses HSD	1.69	2.03
Thematic mapping of Output Areas	1.91	2.38
Thematic mapping of grids 250m	2.00	2.34
Thematic mapping of grids HSD	2.06	2.63
Kernel density estimation	2.90	3.41

Table 6. Prediction Accuracy Index values for different hotspot mapping techniques. Grey cells indicate the highest values and values in bold italics indicate the lowest PAI values. Results are presented for each of the dates when hotspot maps were generated. These results show that KDE consistently produced the best hotspot maps for predicting future events.

Table 7 shows the average PAI values for each hotspot mapping technique for each crime type (these are calculated from averaging the PAI values for each crime type – for each hotspot mapping technique – for all periods of input and measurement data). These results were generated to see if differences in the prediction abilities of hotspot mapping techniques were consistent, and to further explore differences between crime types. These results show that for each crime type, KDE consistently proved to be the best technique for predicting future patterns of crime. The spatial ellipses technique was not though the poorest performer for each crime type. Thematic mapping of Output Areas generated the lowest PAI values for residential burglary and in one case for thefts from vehicles.

Hotspot mapping technique	Residential burglary	Street crime	Theft from vehicle	Theft of vehicle
Spatial ellipses 250m	1.38	2.36	2.18	1.65
Spatial ellipses 500m	1.34	1.46	1.54	0.82
Spatial ellipses HSD	1.43	2.45	2.12	1.29
Thematic mapping of Output Areas	1.10	4.20	1.17	1.18
Thematic mapping of grids 250m	1.70	4.04	1.82	1.37
Thematic mapping of grids HSD	1.68	3.46	2.12	2.06
Kernel density estimation	2.31	4.68	2.29	2.32

(a)

Hotspot mapping technique	Residential burglary	Street crime	Theft from vehicle	Theft of vehicle
Spatial ellipses 250m	1.32	2.59	2.15	2.93
Spatial ellipses 500m	1.31	1.40	1.55	1.82
Spatial ellipses HSD	1.29	2.63	2.63	1.59
Thematic mapping of Output Areas	1.25	3.32	2.93	2.01
Thematic mapping of grids 250m	1.67	3.58	2.43	1.66
Thematic mapping of grids HSD	1.95	4.14	2.55	1.89
Kernel density estimation	2.33	4.59	3.66	3.05

(b)

Table 7. Prediction Accuracy Index values for different hotspot mapping techniques, by crime type. (a) PAI values calculated from the 1st January 2003 measurement date and (b) PAI values calculated from the 13th March 2003 measurement date. Grey cells indicate the highest values and values in bold italics indicate the lowest PAI values. These results show that KDE consistently produced the best hotspot maps for predicting spatial patterns of crime for all crime types, and that in some cases STAC was not the worst performer. Instead, thematic mapping of Output Areas generated the lowest PAI values for residential burglary, and in one case for theft from vehicles.

The results shown in tables 6 and 7 also indicate the influence of parameter settings in the ability of hotspot maps to predict where crimes may occur in the future. For example, for grid thematic mapping the size of each grid needs to be determined and for spatial ellipses and KDE a bandwidth (or search radius) needs to be entered. In this research the Hotspot Detective for MapInfo default parameter settings were used to generate KDE hotspot maps as these appear to be determined with some thought and consideration on the data that is being analysed. However, for spatial ellipses there is no such default setting and instead it requires the user to determine their own values. In this research, to avoid coming to conclusions that could be determined as bias because of the selection of parameters for certain techniques, different settings for the parameters for hotspot mapping of spatial ellipses and grid thematic mapping were chosen.

Table 6 shows that PAI results varied within each technique due to the parameters that were entered. For spatial ellipses a 250m search radius parameter on average generated the best PAI results for this technique. However, inspection of table 7 shows that this was not always the case, with Hotspot Detective determined STAC search radius settings occasionally generating the best PAI values for the spatial ellipse method. The larger search radius setting of 500m for the spatial ellipse technique generated the worst PAI values in all but one case, and for this occasion, the PAI result was very similar to both the 250m and Hotspot Detective determined parameter settings. For grid thematic mapping, a grid cell size determined from

Hotspot Detective parameter settings produced the best PAI values on most occasions. It is important to note however that in not one case did any technique generate PAI values that were higher than those generated by KDE.

5.3. How many crimes can a hotspot map predict?

The PAI provides a useful comparative measure, but more useful perhaps to practitioners is a measure that compares how many crimes a hotspot map can most likely predict. To do this it requires the area that the mapping technique determines as 'hot' to be controlled for in size. That is, there would be little use in saying that a technique can predict 100% of all future crimes if the area it determines as hot is the entire study area. For the purposes of demonstrating how many crimes a hotspot mapping technique can predict we controlled the area determined as 'hot' to be 3% of the entire study area and as KDE consistently produced the highest PAI values, this technique was selected to provide these comparisons for each of the different crime types.

Table 8 shows PAI results for KDE when three months of input data was used for generating hotspot maps for each crime type to determine the number of crimes that it could predict, when the measurement date was the 1st January 2003 and the prediction period was for the next one month. These results again highlight the differences between crime types in their PAI values and show the relatively high PAI results that KDE tends to generate. Interestingly, these PAI values are higher than the previous average values (table 6) indicating that the currency of input data and the measurement period may influence differences in PAI results. Further research to explore this is discussed in the next section. When the size of the area that was determined as hot was controlled to 3% of the total study area, it was possible to compare the number of crimes for each crime type that occurred in the next month in the areas determined as hotspots. The results from this analysis show that in all cases at least 8% of crime that took place after the measurement date was in an area representing only 3% of the study area, validating the role of hotspot mapping for helping to determine where to focus police and crime reduction resources. This value was even higher for theft from vehicles (12%) and street crime (20%), indicating that for street crime, if police and crime reduction resources are targeted to just those areas determined as hotspots, there is the chance of tackling a fifth of all crime that is committed across the whole study area.

Crime type	PAI	Crimes committed in January 2003	Number of crimes in hotspots	Percentage of crimes in hotspots
Residential Burglary	2.77	470	39	8%
Street Crime	6.59	460	91	20%
Theft from Vehicle	3.98	962	115	12%
Theft of Vehicle	3.26	307	30	10%

Table 8. PAI and actual crimes predicted using kernel density estimation to generate a hotspot map from the previous three months of crime data and determine where crimes in the next month may occur (using a measurement date of the 1st January 2003). The area determined as 'hot' was controlled to cover only 3% of the study area's total area.

6. Discussion and implications

The results from this research provide a measure that compares hotspot mapping techniques in their ability to perform a function that they most usually are designed to

do – identifying where future crimes may occur. Kernel density estimation is increasingly becoming the technique of choice by those that generate hotspot maps, partly based on the findings from previous reviews that have profiled the technique's ability to outperform others in accurately identifying the location, size, orientation and spatial distribution of the underlying point data, and the visual appeal in the output that it generates (Chainey and Ratcliffe, 2005, Eck et al, 2005). This research shows that KDE is also the best of the common hotspot mapping techniques for predicting spatial patterns of crime. Grid thematic mapping proved to be slightly better than thematic mapping of Output Areas, whilst spatial ellipses generated using STAC were the worst at predicting spatial patterns of crime. As an example, figure 4 shows hotspot maps generated for each of these techniques from three months of residential burglary input data when the measurement date was the 1st January 2003. The figures show that each technique identified similar areas, but in terms of the ability to predict future spatial patterns of residential burglary over the next one month, KDE was better at more acutely defining hotspots and predicting where future residential burglaries may occur.
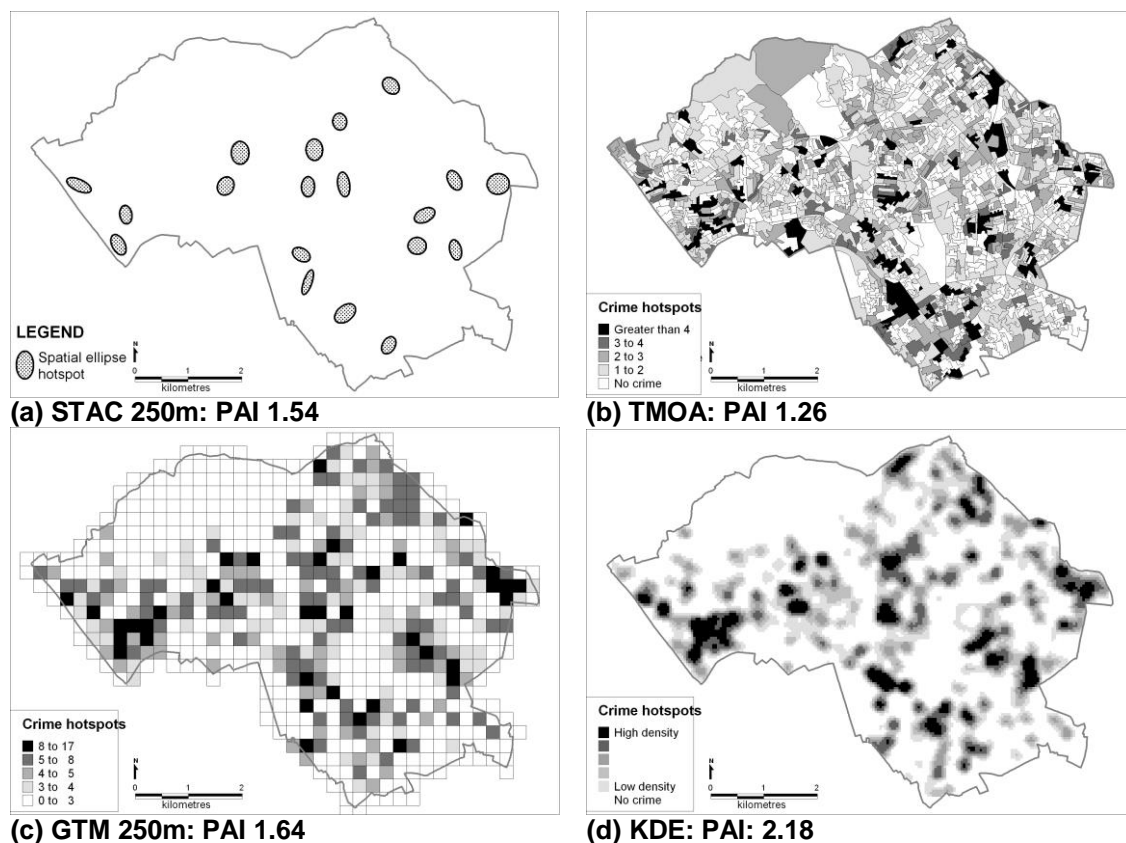
Figure 4. Hotspot maps generated from three months of residential burglary input data (measurement date of the 1st January 2003) using (a) STAC, (b) thematic mapping of Output Areas, (c) grid thematic mapping, and (d) kernel density estimation. Each map is shown with its PAI value, based on one month of measurement data.

The research has also revealed the differences in hotspot maps between crime types in their ability to predict future patterns of crime. For example, while figures 5a and 5b may look similar in their visual appeal for identifying hotspots of crime and their ability to suggest where crime may occur in the future, the hotspot map of street crime (5b) is over twice as good at predicting future events (capturing 20% of street crimes in the areas defined as hotspots that took place in the 'next' month) than the hotspot map of

residential burglary (capturing 8% of residential burglaries that took place in the 'next' month).

(a) PAI: 2.77 Hit Rate: 8%   (b) PAI: 6.59 Hit Rate: 20%
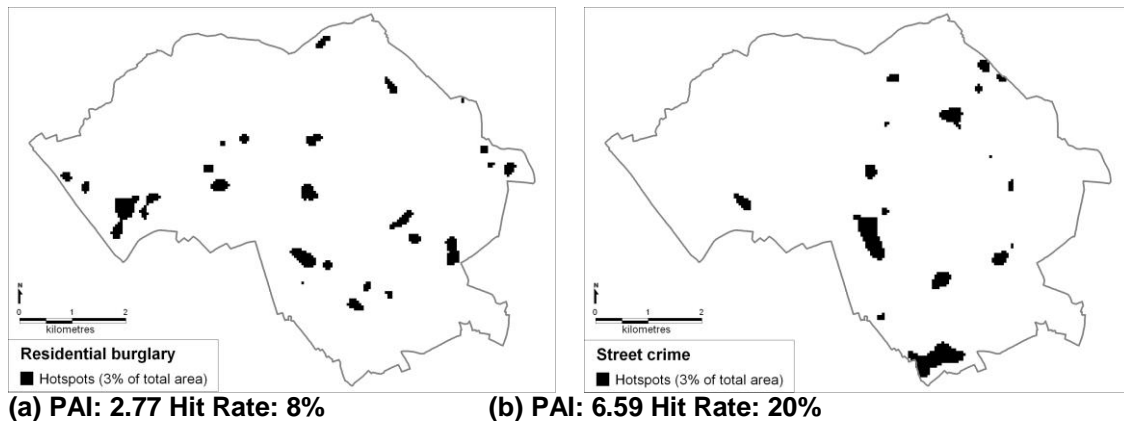
Figure 5. KDE hotspot maps of (a) residential burglary and (b) street crime, generated from three months input data, and where the hotspot area in each is controlled to represent 3% of the total area. Each figure is presented with its PAI value and its hotspot hit rate for predicting where crimes in the next month occurred. The street crime hotspot map is over twice as a good as the residential burglary hotspot map for predicting where crimes of the respective crime type may occur in the future.

Street crime hotspots maps were consistently better in their ability to predict spatial patterns of street crime than any of the other crime types in their ability to predict crimes of their respective type. Closer examination of hotspot maps revealed street crime predominantly occurred in areas where shops, bars, restaurants, markets and other forms of retail and entertainment concentrate - places that are prone to the opportunities to commit street crime. This type of land use tends to be clustered at particular localities, meaning that the opportunity for street crime is similarly highly concentrated. These types of land use also tend to be static, in that they do not shift around the urban landscape but instead become a stationary part of the area's environmental fabric. Crime patterns tend to follow that of opportunities to commit crime (Cohen and Felson, 1979; Cornish and Clarke, 1986), and hence as the opportunities to commit street crime remain fairly static in geographic space, it is likely that retrospective data on where street crimes have occurred previously would be a good indicator of where street crime may occur in the future. This is most probably the reason why hotspot maps of street crime generated the highest PAI values.

In comparison, opportunities to commit thefts from vehicles and thefts of vehicles, while they may be concentrated to certain places, tend to be more transient. This is particularly the case in Camden and Islington where parking of vehicles is often limited to owners parking on the street due to few houses in this area possessing a garage. Car ownership continues to increase in London and parking spaces in this densely populated region of London are at a premium. This means that many residents in Camden and Islington find it difficult to park in close proximity to their house, and visitors to Camden and Islington often have to search extensively for an on-street parking place. Car parks do exist towards to the south of the two boroughs and in the main town centres, but parking tends to be a lottery, with no guarantee that a space will be found in close proximity to the desired destination. This means that the opportunity to commit crimes against vehicles is more evenly spread across Camden and Islington than the opportunities that exist for street crime. While hotspots of vehicle crime may occur, influenced by other socio-economic and physical characteristics, the high density of vehicles parked on the street is a feature across the entire landscape of the

study area, meaning that crime patterns for vehicle crime would be more dispersed. This means that retrospective data on where vehicle crime occurs would most likely be less reliable than retrospective data on street crime for predicting spatial patterns of future events. Similarly, residential properties are spread across Camden and Islington, with very few areas being devoid of any resident population. While certain types of property may be more prone to burglary than others, the opportunity to commit this type of offence is not as heavily concentrated as opportunities to commit street crime. This difference in opportunities is again most likely to be the reason why hotspot maps of residential burglary generated PAI values that were lower than those for street crime.

The influence of opportunities to commit crime on predicting where crime may occur in the future, over the ability of retrospective data to help determine future spatial patterns of crime raises debate on the practical utility of past events for determining future actions. This is a discussion that other researchers have also begun to consider (Bowers et al, 2004, Gorr and Olligschaelger, 2002, Groff and LaVigne, 2002). It raises the need for more systematic evaluation of indicators that can be used to best inform future events, and includes identifying if a weighting scheme can be applied to retrospective crime data, with recent incidents having a greater influence over events that occurred in the distant past.

In this research only two measurement dates were used for the calculation of PAI values. While the results were consistent between the measurement dates, analysis on different measurement dates would add further value to the findings. This may involve designing a software programme to run iterations of hotspot map generation for multiple dates. Repeating the analysis for a different study area would also be of value to compare research findings, plus there is interest in applying the methodology to other crime types, particularly violent crime.

This research has also shown the influence that parameters can have on mapping output. KDE requires two parameters – the cell size and the bandwidth – and previous research has identified how the bandwidth in particular can have an influence on hotspot maps that are generated (see Chainey and Ratcliffe, 2005, and Eck et al, 2005). The results also showed the influence that the currency or volume of input data may have on the ability to predict future events and for how long into the future a hotspot map continues to offer value in predicting where crimes may occur (i.e. considering if the PAI value of a hotspot map generated from retrospective data decays over time, requiring the map to be updated with more recent retrospective data). This calls for more research to investigate the impact of KDE parameters, the influence of input data on PAI values, and whether a hotspot map has a 'shelf-life' that suggests it should be replaced with more up-to-date retrospective data that will more accurately predict future events. This notion of the shelf-life and influence of the currency of data may also differ between crime types, therefore influencing an agency's production line of analysis and intelligence creation.

Introducing the Prediction Accuracy Index has made it possible to make comparisons between hotspot mapping techniques and crime types in their ability to predict future crime events. Its calculation also allows the analyst to generate a statistic (the hit rate) that indicates the possible impact that targeted resource activity could have if it was focused to those areas defined as hotspots. Value in the PAI though is that it can be applied to point data of any crime type, any mapping technique that attempts to predict where spatial patterns of crime may occur, and for any study area. In this research, common hotspot mapping techniques and familiar crime types have been analysed, providing useful results that allow others to benchmark against. Indeed, as a methodology, it offers potential as a standard measure against which more advanced

spatial analysis and predictive crime mapping techniques can be measured and evaluated in their ability to identify where crime may occur in the future.

7. Conclusion

Hotspot mapping is used by many policing and crime reduction practitioners to identify spatial patterns of crime. By using data from the past, hotspot mapping identifies where crime most densely concentrates, starting a decision-making process that considers where best to target enforcement and prevention resources. There are many mapping techniques that can be used for generating hotspot maps, but to date none have been researched to consider if they are accurate in predicting where crime may occur in the future.

This research has introduced the Prediction Accuracy Index as a measure to compare prediction abilities of crime mapping techniques and identify if the most common hotspot mapping techniques used by practitioners differ in their ability to predict spatial patterns of crime. This research has also considered if hotspot maps for different crime types differ in their prediction abilities.

Kernel density estimation has become the technique of choice by many practitioners, supported in part by other reviews that look favourably upon this technique. This paper has shown that KDE is also better at predicting future spatial patterns of crime in comparison to the other most common hotspot mapping techniques; standard deviational ellipses, thematic mapping of boundary areas, and grid thematic mapping.

An assumption that many practitioners make in their use of hotspot maps, is that their utility does not differ between crime types. In this research, comparisons between crime types revealed that hotspot maps of street crime were consistently better at predicting where street crime will occur in the future, than hotspot maps of other crime types were at predicting their respective future occurrences of crime. When this result was considered further it identified the link between the stability of opportunities for crime to occur and the use of retrospective data for predicting future spatial patterns - identifying that street crime opportunities tended to be more stable than opportunities for other crimes, and hence retrospective data on street crime was a better predictor of future street crimes in comparison to crime types where the opportunities were more transient or geographically dispersed.

All hotspot mapping techniques require parameters to be set and require data for analysis. Findings from this research indicate that consideration of the currency (or volume) of input data and careful parameter selection may optimise the ability of hotspot mapping techniques to predict spatial patterns of crime, and indeed influence the development of new techniques in predictive crime mapping. Results from this research provide an important estimate on the ability of hotspot maps to target resources, and the possible volumetric impact that this targeting may have. The research also benchmarks how existing techniques perform in their ability to predict future occurrences of crime against which more advanced and new techniques can be compared.

References

Bailey, T.C., and Gatrell, A.C. (1995) *Interactive Spatial Data Analysis*. Reading, Massachusetts: Addison-Wesley.

Block, R. and Block, R.B. (2000) The Bronx and Chicago – Street Robbery and the Environs of Rapid Transit Stations. In Goldsmith, V., McGuire, P.G., Mollenkopf, J.H. and Ross, T.A. (Eds.) *Analysing Crime Patterns: Frontiers and Practice*. Sage: Thousand Oaks, pp 137-152

Block, R. and Perry, S. (1993) *STAC News.* Vol. 1, No. 1. Illinois Criminal Justice Information Authority: Statistical Analysis Center.
Available online at: www.icjia.state.il.us/public/index.cfm?metasection=publicationsandmetapage=STACNEWS_01_W93

Bowers, K.J., Johnson, S. and Pease, K. (2004) Prospective Hotspotting: The Future of Crime Mapping? *British Journal of Criminology*, Vol. 44, No. 5, pp 641-658.

Bowers, K., Newton, M. and Nutter, R. (2001) A GIS-linked Database for Monitoring Repeat Domestic Burglary. In Hirschfield, A. and Bowers, K. (Eds.) *Mapping and Analysing Crime Data – Lessons from Research and Practice*. London: Taylor and Francis.

Bowers, K.J. and Hirschfield, A. (1999) Exploring Links Between Crime and Disadavantage in North-West England – An Analysis Using Geographical Information Systems. *International Journal of Geographic Information Science.* Vol. 13, No. 2, pp 159-184.

Brantingham, P.J. and Brantingham, P.L. (1984) *Patterns In Crime*. New York: Macmillan.

Chainey, S.P. (2001) Combating Crime Through Partnership; Examples of Crime and Disorder Mapping Solutions in London, UK. In Hirschfield, A. and Bowers, K. (Eds.) *Mapping and Analysing Crime Data – Lessons from Research and Practice*. London: Taylor and Francis.

Chainey, S.P. and Ratcliffe, J.H. (2005) *GIS and Crime Mapping*. London: Wiley.

Chainey, S.P., Reid, S. and Stuart, N. (2002) When is a Hotspot a Hotspot? A Procedure for Creating Statistically Robust Hotspot Maps of Crime. In Higgs, G. (Ed.) *Innovations in GIS 9 Socio-economic Applications of Geographic Information Science*. London: Taylor and Francis.

Clarke, R.V. and Eck, J. (2003) *Become a Problem-Solving Crime Analyst in 55 Small Steps.* London: Jill Dando. Institute, University College.
Available online at http://www.jdi.ucl.ac.uk/publications/other_publications/55steps

Cohen, L.E. and Felson, M. (1979) Social Change and Crime Rate Trends: A Routine Activity Approach. *American Sociological Review*. Vol. 44, pp 588-605.

Cornish, D. and Clarke, R. (1986) The Reasoning Criminal: Rational Choice Perspectives on Offending. New York: Springer-Verlag.

Dent, B. (1999) *Cartography – Thematic Map Design*. London: McGraw-Hill

Eck, J.E., Chainey, S.P., Cameron, J.G., Leitner, M. and Wilson, R.E. (2005) *Mapping Crime: Understanding Hot Spots*. USA: National Institute of Justice. Available online at www.ojp.usdoj.gov/nij

Goldsmith, V., McGuire, P.G., Mollenkopf, J.H. and Ross, T.A. (2000) *Analysing Crime Patterns: Frontiers of Practice.* New York: Altamira Press.

Gorr, W. and Olligschlaeger, A. (2002). Crime hotspot forecasting: Modelling and comparative evaluation. *Final report to the National Criminal Justice Reference Service (NCJRS).*

Groff, E.R. and La Vigne, N.G. (2002) Forecasting the Future of Predictive Crime Mapping. In Tilley, N. (Ed.) *Analysis for Crime Prevention* (Crime Prevention Studies Vol. 13). Monsey NY: Criminal Justice Press.

Harries, K. (1999) *Mapping Crime: Principle and Practice*. United States National Institute of Justice.
Available online at http://www.ojp.usdoj.gov/nij/maps/pubs.html

Home Office (2005) Crime Mapping: Improving Performance, A Good Practice Guide for Front Line Officers. London: Home Office.
Available   online   at:
http://www.jdi.ucl.ac.uk/downloads/publications/other_publications/crime_mapping_guide.pdf

Home Office (2001). Crime Reduction Toolkits Focus Area and Hotspots. Crime Reduction Unit.
Available online at www.crimereduction.gov.uk/toolkits/fa00.htm

Hough, M. and Tilley, N. (1998) *Getting the Grease to the Squeak: Research Lessons for Crime Prevention.* Crime Prevention and Detection Paper 85. London: Home Office

Illinois Criminal Justice Information Authority (1996) *Spatial and Temporal Analysis of Crime*. State of Illinois.

Jefferis, E. (1999) *A Multi-Method Exploration of Crime Hot-spots: A Summary of Findings*. Crime Mapping Research Centre Intramural project. Washington, DC: National Institute of Justice.

Langworthy, R.H. and Jefferis, E. (2000) The Utility of Standard Deviation Ellipses for Evaluating Hotspots. In Goldsmith, V., McGuire, P.G., Mollenkopf, J.H. and Ross, T.A. (Eds.) *Analysing Crime Patterns: Frontiers and Practice*. Thousand Oaks, California: Sage.

LaVigne, N. and Wartell, J. (Eds.). (1998) *Crime Mapping Case Studies: Successes in the Field,* Vol. 1. Washington, DC: PERF.

LaVigne, N. and Wartell, J. (Eds.). (1999) *Crime Mapping Case Studies: Successes in the Field,* Vol. 2. Washington, DC: PERF.

LeBeau, J.L. (2001) Mapping Out Hazardous Space for Police Work. In Bowers, K. and Hirschfield, A. *Mapping and Analysing Crime Data – Lessons from Research and Practice*. London: Taylor and Francis.

Levine, N. (2004) *CrimeStat III: A Spatial Statistics Program for the Analysis of Crime Incident Locations*. Houston, TX: Ned Levine and Associates. Washington, DC: National Institute of Justice.
Available online at http://www.icpsr.umich.edu/nacjd/crimestat.html

Martin, D., Barnes, E. and Britt, D. (1998) The Multiple Impacts of Mapping it Out; Police, Geographic Information Systems (GIS) and Community Mobilization During Devil's Night in Detroit, Michigan. In La Vigne, N. and Wartell, J. (Eds.) *Crime Mapping Case Studies: Successes in the Field*. USA: Police Executive Research Forum

McDonald, P.P. (2002) Managing Police Operations: Implementing the New York Crime Model – CompStat, Belmont, CA: Wadsworth

McGuire, P.G. and Williamson, D. (1999) Mapping Tools for Management and Accountability. Paper presented to the Third International Crime Mapping Research Center Conference, Orlando, Florida, December 11-14, 1999.

Monomier, M. (1996) *How to Lie with Maps*. Chicago: University of Chicago Press.

Openshaw, S. (1984) *The Modifiable Areal Unit Problem*. Concepts and Techniques in Modern Geography 38. Norwich, UK: Geobooks

Osborne, D.A. and Wernicke, S.C. (2003) Introduction to Crime Analysis: Basic Resources for Criminal Justice Practice. New York: Haworth Press.

Ratcliffe, J., (2002) *HotSpot Detective 2.0 for MapInfo Professional 7.0*. Available online at http://jratcliffe.net/hsd/

Ratcliffe, J. (2004) HotSpot Detective for MapInfo Helpfile Version 2.0

Ratcliffe, J.H. and McCullagh, M.J. (1999) Hotbeds of Crime and the Search for Spatial Accuracy. *Journal of Geographical Systems*, Vol. 1, No. 4, pp 385-398.

Ratcliffe, J. and McCullagh, M. (2001) Crime, Repeat Victimisation and GIS. In Bowers, K. and Hirschfield, A. *Mapping and Analysing Crime Data – Lessons from Research and Practice*. London: Taylor and Francis. Pp 61-92.

Schnick, W. (2004) CompStat in the Los Angeles Police Department. *Police Chief*, Vol. 71, No. 1, pp 17-23.

Walsh, W. (2001) Compstat: An Analysis of an Emerging Police Paradigm. *Policing: An International Journal of Police Strategies and Management,* Vol. 24, No. 3, pp 347-363.

Weir, R. and Bangs, M. (2007). *The Use of Geographic Information Systems by Crime Analysts in England and Wales*. Home Office Online Report Series. London: Home Office.

Williamson, D., McLafferty, S., McGuire, P., Ross, T., Mollenkopf, J., Goldsmith, V. and Quinn, S. (2001) Tools in the Spatial Analysis of Crime. In Hirschfield, A. and Bowers, K. (Eds.)*, Mapping and Analysing Crime Data: Lessons from Research and Practice*, London: Taylor and Francis, pp 187-202.

Williamson, D., McLafferty, S., McGuire, P., Goldsmith, V. and Mollenkopf, J. (1999) A Better Method to Smooth Crime Incidence Data,' *ArcUser Magazine,* January - March, 1999, 1-5. Available online at http://www.esri.com/news/arcuser/0199/crimedata.html