The current issue and full text archive of this journal is available at
www.emeraldinsight.com/1363-951X.htm

Kernel density estimation and hotspot mapping

Examining the influence of interpolation method, grid cell size, and bandwidth on crime forecasting

Received 16 April 2013
Revised 19 July 2013
Accepted 20 July 2013

Timothy Hart
Department of Criminology and Criminal Justice, Griffith University, Gold Coast, Queensland, Australia, and
Paul Zandbergen
Department of Geography, University of New Mexico, Albuquerque, New Mexico, USA

Abstract

Purpose – The purpose of this paper is to examine the effects of user-defined parameters settings (e.g. interpolation method, grid cell size, and bandwidth) on the predictive accuracy of crime hotspot maps produced from kernel density estimation (KDE).

Design/methodology/approach – The influence of variations in parameter settings on prospective KDE maps is examined across two types of interpersonal violence (e.g. aggravated assault and robbery) and two types of property crime (e.g. commercial burglary and motor vehicle theft).

Findings – Results show that interpolation method has a considerable effect on predictive accuracy, grid cell size has little to no effect, and bandwidth as some effect.

Originality/value – The current study advances the knowledge and understanding of prospective hotspot crime mapping as it answers the calls by Chainey *et al.* (2008) and others to further investigate the methods used to predict crime.

Keywords GIS, Hotspots, Predictive policing

Paper type Research paper

1. Introduction

Predicting the location and time of future crime events is of growing interest to both the law enforcement and academic communities. One approach to crime prediction is hotspot mapping, which relies on the assumption that the locations of past events are good predictors of future events. According to the Bureau of Justice Statistics, nearly all large law enforcement agencies in the USA that serve populations of at least 500,000 engage in predictive hotspot mapping (Reaves, 2010). While the use of predictive crime mapping among law enforcement agencies has increased in recent years, the interest in it has also grown within the scientific community. An evolving body of academic literature demonstrates strong theoretical grounding for and widespread application of prospective hotspot mapping.

A wide range of analytic techniques that can be used to characterize crime hotspots has developed from the existing scholarly literature. These approaches generally fall into two broad categories: methods that rely on aggregated crime data and analysis of

This paper is supported in part by Award 2009-IJ-CX-0022, *Effects of Geocoding Quality on Predictive Crime Hotspot Mapping*, from the National Institute of Justice.

Policing: An International Journal of Police Strategies & Management
Vol. 37 No. 2, 2014
pp. 305-323

1363-951X
DOI 10.1108/PIJPSM-04-2013-0039

discrete crime event locations. Although no single technique has emerged as the "best" for predicting crime, recent research suggests that Kernel Density Estimation (KDE) outperforms other approaches (Chainey *et al.*, 2008a). The growing availability of KDE in popular GIS applications, the perceived accuracy of its hotspot identification, and the aesthetically pleasing and easily understandable output are among the driving forces behind its popularity[1].

Several geostatistical software packages include procedures that allow users to produce hotspot maps based on KDE. In crime analysis, this process is relatively straightforward and involves creating a two-dimensional grid surface which covers all crime incident point locations within a given study area, overlaying the grid on the study area, measuring the distance between the center point of each grid cell and each crime incident, and weighting that distance based on a specific method of interpolation (e.g. the kernel function). The end result is a visual representation of discrete crime points "smoothed" over the study area, thereby creating a continuous risk surface.

As with all hotspot-mapping techniques, several user-defined parameters must be established prior to producing a KDE map. For example, the user must decide the type of estimation technique to use (e.g. single vs duel KDE), the method for interpolation (e.g. the kernel function), the bandwidth (e.g. the size of the search radius used by the kernel), and the output format units (e.g. grid cell size) in which the KDE results will be generated. Multiple KDE maps produced from the same data may vary in predictive accuracy simply because parameters used in the interpolation process are changed. And although the criminological literature offers some guidance on how to establish some of the parameters used in KDE, it is less persuasive on which parameter settings yield the most predictively accurate crime hotspot maps.

The current study is designed to address this gap in the current scholarship. The objective of this study is to advance our understanding of prospective crime hotspot mapping techniques by investigating the effects of interpolation method, grid cell size, and bandwidth on the predictive accuracy of KDE. Specifically, the current study produces empirical descriptions of the quality of KDE hotspot maps based on variations in important user-defined settings that are part of the interpolation process. Our analytic strategy includes examining these effects across multiple crime types, including aggravated assault, robbery, commercial burglary, and motor vehicle theft. Our results inform a growing number of criminal justice professionals and academics interested in crime forecasting.

The rest of this manuscript is organized in the following manner: first, an overview of the relevant literature is offered, focussing on hotspot mapping techniques in general and parameter settings used in KDE in particular. After, descriptions of the data and methods used in the current study as well as our analytic strategy are provided. This is followed by the results of our analyses. Our paper concludes with a discussion of the implications of our findings, limitations of the current study, and suggestions for future research.

2. Hotspot identification

Crime hotspots are areas of concentrated crime incident locations that demonstrate a nonrandom pattern in space and/or time. There are different levels of hotspots analysis, depending on the size of the geographic area of concern, from very specific locations or addresses to streets, blocks, and neighborhoods. Each level of analysis corresponds to specific research questions, and a wide range of methods and techniques has emerged from the current literature to characterize crime hotspot locations (Bowers *et al.*, 2004; Chainey and Ratcliffe, 2005; Chainey *et al.*, 2002; Eck *et al.*, 2005; Jefferis, 1999; McLafferty *et al.*, 2000).

Techniques used to identify crime hotspots fall into two generally categories: methods based on aggregated incident locations and analysis of point-patterns. Techniques that rely on aggregated crime counts include grid-based thematic mapping and local tests of spatial association. Two of the most common tests of spatial association used in crime mapping are Local Moran's *I* (Anselin, 1995) and Gi*/Getis-Ord local statistic (Getis and Ord, 1992; Ord and Getis, 1995). Although both of these aggregate tests of statistical association are unique, they generally follow a similar approach: correlation statistics are calculated for each areal unit based on aggregated crime counts, weighted by a spatial matrix that is defined by the locations and distances of neighboring crime counts. For example, Equations (1)-(5) are used to calculate the Local Moran's *I* statistic of spatial association and is given as:

$$I_i = \frac{x_i - \overline{X}}{s_i^2} \sum_{j=1, j\neq i}^{n} w_{i,j}(x_i - \overline{X}) \tag{1}$$

where x_i is an attribute for feature I, $\overline{X}$ is the mean of the corresponding attribute, $w_{i,j}$ is the spatial weight between feature i and j, and:

$$s_i^2 = \frac{\sum\limits_{j=1, j\neq i}^{n} w_{ij}}{n-1} = \overline{X}^2 \tag{2}$$

with n equating to the total number of features.

The z_{I_i}-score for the statistics are computed as follows:

$$z_{I_i} = \frac{I_i - E[I_i]}{\sqrt{V[I_i]}} \tag{3}$$

where:

$$E[I_i] = \frac{\sum\limits_{j=1, j\neq i}^{n}}{n-1} \tag{4}$$

$$V[I_i] = E[I_i^2] - E[I_i]^2 \tag{5}$$

An alternative approach to crime hotspot detection involves point-pattern analysis. Also known as adaptive scanning methods, common techniques that fall within this general category include modal hotspot analysis, spatial and temporal analysis of crime (STAC) (Spring and Block, 1989), nearest neighbor hierarchical clustering (Hartigan, 1975; Ward, 1963), and K-means clustering (Ball and Hall, 1970; McBratney and deBruijter, 1992; Thompson, 1956). Like aggregate methods for determining spatial associations discussed previously, points-pattern techniques share a common approach: identifying spatially clustered discrete crime event locations, based on certain input parameters used to conceptualize a hotspot. Resulting maps produced from these methods visually display hotspots as geometrically defined areas such as ellipses and convex hulls.

In point-pattern analysis, certain parameter settings are required, which can range from simple to very complex. For example, in modal hotspot analysis users need only to define the study area and the features being analyzed (Bruce and Smith, 2011; Levine, 2004). On the other hand, STAC parameter settings include the size of the search radius, the units in which the search radius is based, the minimum number of points that define a cluster, the number of simulations that will be run (if Monet Carlo simulations are utilized), the type of scanning procedure (e.g. rectangular or triangular), and the number of standard deviations used to create the output ellipses (Bates, 1987; Canter, 1993; Spring and Block, 1989). Despite differences in how point-pattern techniques approach the visualization of crime hotspots, they all involve methods that use known volumes of past crime, at specific locations, to identify spatially nonrandom clusters of events.

In order to implement effective crime control strategies, however, hotspot mapping must go beyond retrospective analysis. Several general approaches to forecasting crime have emerged in recent years, including the use of temporally aggregated hotspots, individual-level analysis of near repeat victimization, and various univariate and multivariate analysis of area-level data (Bowers *et al.*, 2004; Groff and LaVigne, 2002; Johnson and Bowers, 2004a, b; Mohler *et al.*, 2011; Ratcliffe and Rengert, 2008; Ratcliffe, 2010). In one of the first comparative analyses involving a range of hotspots methods, Chainey *et al.* (2008a) recently demonstrated that KDE yields results that are more predictively accurate than other approaches[2].

2.1 KDE

Unlike point-pattern analytic techniques discussed above, KDE does more than simply describe known crime events in terms of nonrandom spatial patterns visualized as various geometric shapes (e.g. first- or second-order spatial ellipses and convex hulls). Instead, KDE produces a prospective risk map, which is interpolated from specific crime incident locations found in a defined study area. In other words, KDE generalizes or "smooths" discrete data points so that a continuous surface area is produced and visualized (Bailey and Gatrell, 1995).

A growing literature demonstrates the utility of prospective crime analysis techniques[3], including those that utilize KDE. In Risk Terrain Modeling (RTM), for example, individual map layers associated with theoretically and empirically grounded risk factors are created using KDE (Caplan and Kennedy, 2010; Caplan *et al.*, 2011; Kennedy *et al.*, 2011). Each risk map is then combined to create a composite "risk terrain" that shows the presence, absence, and intensity of all risk factors at every location throughout the landscape. Recent evaluations of RTM indicate that it offers an accurate way to identify micro-level areas of risk so that efficient and effective police resources to combat potential crime problems can be deployed (Caplan *et al.*, 2012; Edmonds and Mallard, 2011; Yerxa, 2013).

Debate over whether KDE consistently outperforms other prospective hotspot techniques is ongoing (see, e.g. Levine, 2008; Pezzuchi, 2008; and Chainey *et al.*, 2008a-c). Nevertheless, past research suggests that KDE can be used to accurately forecast crime and as a result it has become one of the most popular approaches to visualize crime hotspots. Tools used to create KDE maps are available in all of the major software packages used in crime analysis, including leading commercial GIS products (e.g. ArcGIS and MapInfo) and other noncommercial spatial statistics applications (e.g. R and CrimeStatIII). In addition to its availability, visually appealing and easily interpretable results produced from KDE analysis has contributed to its widespread adoption and use. Despite its ability to accurately predict future crimes, its widespread adoption, and growing use among academics and practitioners, however, questions about KDE remain unanswered. For example, it is unclear how variations in parameter settings used in the interpolation process affect the predictive accuracy of KDE hotspot maps.

2.2 Parameter settings used in KDE

The mechanics of KDE are straightforward and were first described by Rosenblatt (1956) and Parzen (1962). In prospective crime hotspot mapping, the process involves estimating the density of crime across an entire two-dimensional study area, based on the known locations of discrete events. KDE begins by overlaying a grid (with n equally sized cells) on top of the study area and calculating a density estimate based on the center points of each grid cell. Each distance between an incident and the center of a grid cell is then weighted based on a specific method of interpolation (e.g. the kernel function) and bandwidth (e.g. search radius). Figure 1 illustrates the KDE process and shows a number of parameters that must be considered before a density estimate can be produced. These parameters include the grid cell size, the method of interpolation, and the bandwidth[4].

As noted above, KDE begins by overlaying a two-dimensional grid (with n equally sized cells) across an entire study area. One of the first parameters that must be established, therefore, is how large to make each cell. Decision rules about grid-cell size or the number of columns into which the grid overlay will be divided[5] are usually based the size of the study area. In general, KDE maps with greater resolution (e.g. less pixilation) are produced from grid overlays with smaller cells (e.g. more columns), whereas maps with less resolution result when grid overlays are comprised of larger sized cells (e.g. fewer columns). Since predictive accuracy measures[6] are based on the proportion of a study area that is identified as hotspots, relative to the overall study area, the choice of grid cell size may have meaningful implications on how successful KDE maps are at predicting future crime patterns.

Current literature offers some guidance to determine grid cell size. Chainey and Ratcliffe (2005), for example, suggest that it should be equal to approximately the extent of the shorter side of a study area, divided by 150. This means that a study area measuring 5×3 miles would use a grid cell overlay consisting of cells that are

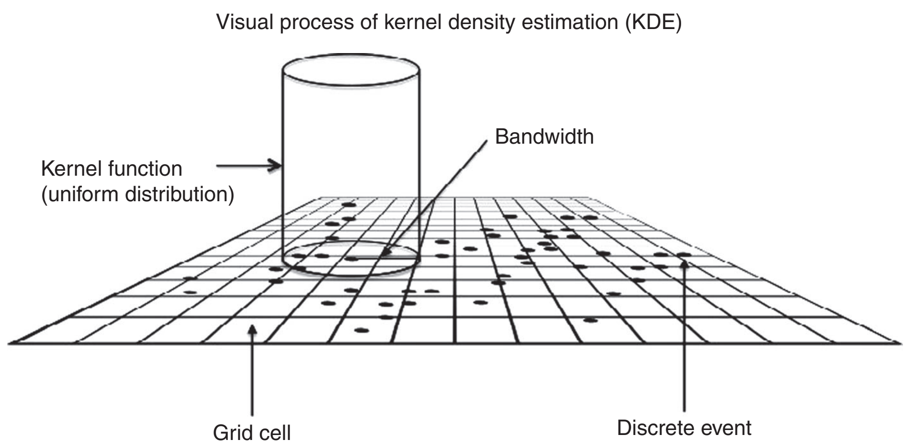

Figure 1.
An illustration of kernel density estimation

approximately 100 × 100 ft[7]. Others suggest that the physical terrain should guide decisions concerning grid cell size. For example, some have argued that grid cells should be between one-half and one-third the length of an average blockface (Caplan *et al.*, 2011, 2012; Kennedy *et al.*, 2011). Caplan and Kennedy (2010) argue that this approach is more appropriate because more "actionable interpretations" of risk are produced when it is used in lieu of larger grid cells.

Two additional parameters that must be defined prior to KDE maps being produced are interpolation method and search radius (also referred to as bandwidth) (see Figure 1). A normal distribution kernel, uniform (flat) distribution, quartic (spherical) distribution, and triangular (linear) distribution are examples of different interpolation methods that can be used in crime analysis[8] (Levine, 2004). When the kernel function is placed over a grid cell center point, the number of crime incidents within the function's bandwidth is used to determine the density estimate assigned to each cell. Values are weighted if a nonuniform kernel function is selected, but not weighted when a uniform (flat) distribution is chosen. Although multiple kernel functions can be used in the interpolation process, some GIS applications limit the number of interpolation methods available or require the user to develop syntax-based routines if they are to be utilized.

Current scholarship offers no guidance on which interpolation method is the most appropriate in prospective crime hotspot mapping. The selection of bandwidth, on the other hand, has received relatively more academic attention. For example, Bailey and Gatrell (1995) suggest a bandwidth of 0.68 times the number of points raised to the −0.2 power, scaled to the areal extent of the study area. This value can be adjusted by multiplying it by the square root of the study area size.

More recently, decisions concerning bandwidth length have been informed by criminological literature that suggests risky places are typically within a few street blocks of a crime incident (Felson, 1995; Taylor, 1997). Similarly, patterns observed in near-repeat offending have been used to assist in the determination of specific bandwidths (Ratcliffe, 2009). More commonly, however, users rely on an application's default setting, which lack scientific justification. If the default setting in ArcGIS is not modified by the user, for example, the search radius in KDE is calculated as the shortest of the width or height of the study area (e.g. a bounding rectangle that encompasses all crime incident locations), divided by 30. The software manufacturer offers no justification or rationale for this calculation.

2.3 Current study

Application of hotspot mapping techniques in crime analysis has grown considerably in recent years. In response, the predictive accuracy of different methods has begun to be evaluated in the criminological literature (Chainey *et al.*, 2008a). Several questions related to one of the most popular techniques, however, have not been adequately answered. First, does interpolation method affect the predictive accuracy of KDE hotspot maps? Second, do variations in grid cell size influence the predictive accuracy of KDE hotspot maps? Third, does the predictive accuracy of KDE hotspot maps change across bandwidths? And finally, are the effects of parameter settings used for KDE hotspot maps consistent across different crime types? Findings from this investigation answer each of these questions and provide researchers and practitioners with valuable guidance and insight into one of the most often used ways of analyzing crime patterns. A more detailed discussion of the data and methods used in the current study are provided in the next section.

3. Data and methods

In order to answer our research questions, we analyzed four specific crime types that occurred in Arlington (Texas) Police Department's (APD) jurisdiction between 2007 and 2008. Our analysis was restricted to only those incidents known to law enforcement and recorded as an aggravated assault ($n = 1{,}056$), robbery ($n = 827$), commercial burglary ($n = 2{,}144$), or motor vehicle theft ($n = 2{,}757$). These crimes were chosen because they represent crimes that are regularly analyzed by law enforcement personnel. They are also qualitatively different offenses in terms of context, which we expect would require different parameter settings in the KDE process.

Using the FBI's hierarchy rule, only the most severe crime was included in the analysis for events involving more than one crime type. Finally, incidents were geocoded to an address point reference layer with an average match rate of 93 percent, which exceeds the generally accepted level of 85 percent established by Ratcliffe (2004). Table I shows the number of crimes recorded by APD, incidents matched, and match rates for both the 2007 and 2008 crime data.

3.1 Analytic strategy

Our analytic approach was a two-step process. First, NIJ's CrimeStatIII (v3.3) software was used to create 36 KDE maps for each of the four crime types. Each surface area represented a unique combination of three different parameters settings used in interpolation process: variation in grid cell size (three column settings), search radius (three bandwidths), and interpolation method (four kernel functions) ($3 \times 3 \times 4 = 36$). Second, each KDE surface area was imported as a shapefile into ArcGIS 10.1 and analyzed to determine the predictive accuracy of the hotspots. Figure 2 provides an example of a KDE map produced for the current study and depicts a KDE surface area interpolated from motor vehicle theft locations identified by APD in 2007. Details of each parameter setting, how crime hotspots were conceptualized, and how predictive accuracy was measured follow.

Grid cell size. The effect of three different grid cell sizes was examined. The first two grid cell size settings were determined by dividing the shorter side of the 2007 crime type's bounding rectangle (e.g. the short of the length or the width of the rectangle that encompassed all crime incident locations) into 250 and 150 columns, respectively[9]. The first setting reflects the default setting used in ArcGIS, whereas the second setting is guided by the recommendations made by Williamson *et al.* (2001). The third grid cell size setting is based on the approach used in RTM (Caplan and Kennedy, 2010; Caplan *et al.*, 2011). In RTM, the grid cell size is typically set so that it is equal to

	2007			2008			Total		
Crime types	Incidents	Matched	Match rate	Incidents	Matched	Match rate	Incidents	Matched	Match rate
Aggravated assault	575	516	89.7	481	423	87.9	1,056	939	88.9
Robbery	453	392	86.5	374	325	86.9	827	717	86.7
Commercial burglary	1,160	1,133	97.7	984	957	97.3	2,144	2,090	97.5
Motor vehicle theft	1,440	1,345	93.4	1,317	1,204	91.4	2,757	2,549	92.5
Total	3,628	3,386	93.3	3,156	2,909	92.2	6,784	6,295	92.8

Table I. Crimes recorded by Arlington (Texas) Police Department, incidents matched, and match rates

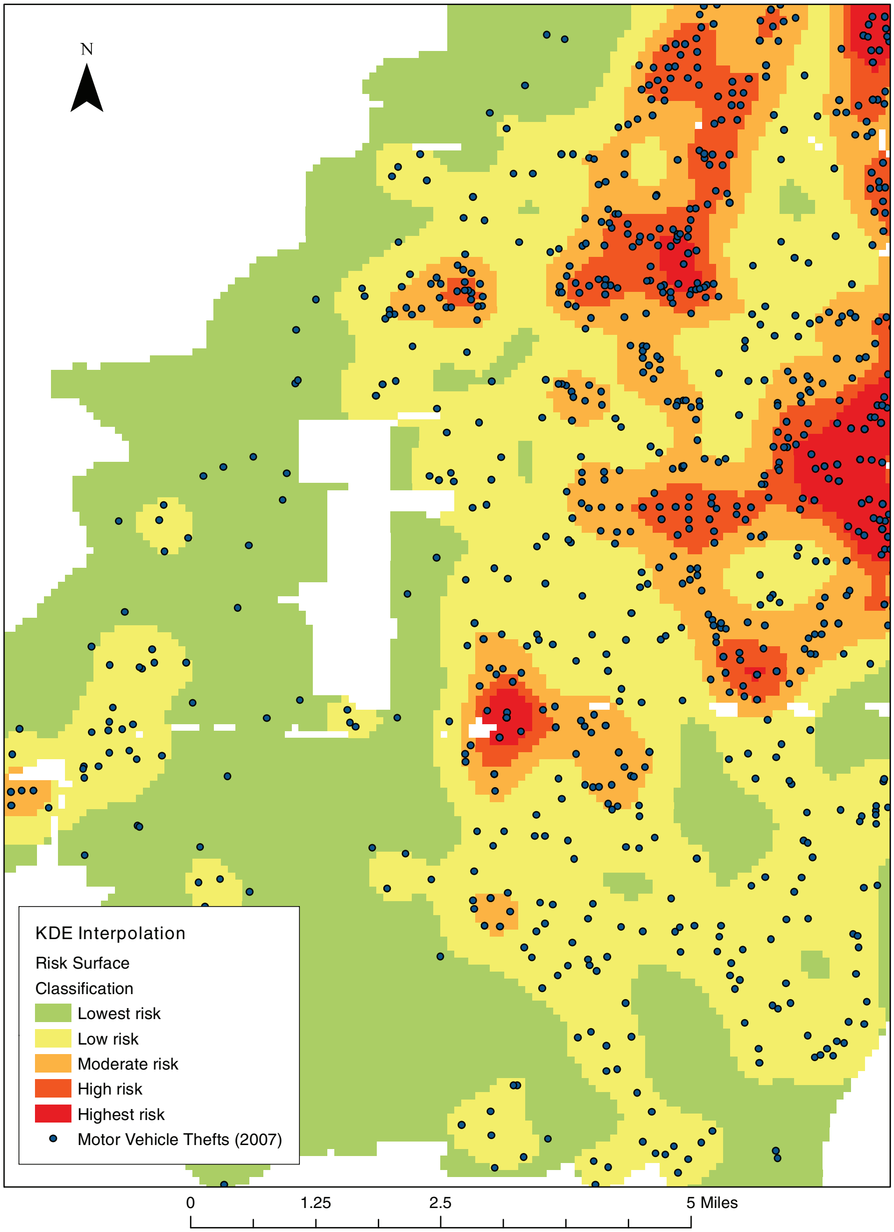

Figure 2.
KDE risk surface in Arlington, Texas produced from known and reported locations of motor vehicle thefts in 2007

Notes: The risk surface is classified by standard deviation breaks. Areas on the map that are red represent the greatest risk of future motor vehicle thefts

approximately one-third the length of the average blockface. The number of columns by which the study area's grid overlay is divide is based on this value[10].

Search radius. Search radius or bandwidth refers to the "spread" of the kernel function used to estimate grid-cell density. As with the grid cell size, the effect of three different search radius settings was examined in the current study. Search radii were set at $\frac{1}{4}$-, $\frac{1}{2}$-, and 1-mile intervals. These settings were intended to capture the qualitatively different contexts in which interpersonal violence (e.g. aggravated assault and robbery) and property crime (e.g. commercial burglary and motor vehicle theft) occur. For our study area, the default radius setting in ArcGIS would have been between $\frac{1}{4}$ and $\frac{1}{2}$ mile depending on crime type.

Interpolation method. Interpolation methods considered in the current study include the normal, quartic, triangular, and uniform distributions. Although not exhaustive of every kernel density function available, these interpolation methods were chosen because they represent the default functions used in popular GIS applications[11] and both weighted and unweighted methods for creating density estimations.

Hotspots and predictive accuracy. A consistent conceptual definition of a hotspot was required in order to assess the predictive accuracy across each KDE map. Therefore, we defined hotspots in the current study as density estimations for grid cells that exceeded the average density score, plus 1.96 times the standard deviation of density scores. In other words, hotspots were defined in the current study as a significantly higher than normal ($p < 0.05$) concentration of crime events within the study area[12].

Finally, the predictive accuracy of KDE maps was evaluated using three different metrics: hit rate, Predictive Accuracy Index (PAI), and Recapture Rate Index (RRI). Hit rate, which the simplest of the three measures, is the percentage of all Time 2 crimes that are contained by predicted hotspots created from Time 1 data. In this study, for example, the hit rate is the percentage of all 2008 crimes contained by the hotspots created from the 2007 events. Although the hit rate is a popular approach to determining predictive accuracy of hotspot maps, it is strongly influenced by the size of the study area and the number of crimes observed at Time 2. Given these limitations, a PAI was also used to assess the quality of predictions.

First introduced by Chainey *et al.* (2008a), PAI is the ratio of the hit rate to the area percentage, or the percentage of the study area that is defined as a hotspot; and is given in the following equation:

$$\mathrm{PAI} = \frac{\left(\frac{n_h}{N_c}\right)}{\left(\frac{a_h}{A_s}\right)} = \frac{\text{Hit rate}}{\text{Area percentage}} \tag{6}$$

where n_h is the number of present crimes observed in the predicted hotspots, N_c is the total number of present crimes, a_h is the area of hotspots, and A_s is the area of the entire extent (e.g. study area). Higher PAI scores indicate greater predictive accuracy.

Levine (2008) suggests that the PAI provides a meaningful way to compare the accuracy of hotspot methods, but argues that it should be accompanied by a measure of precision. He offers the RRI as that metric. RRI is the ratio of crime density at Time 2 to the crime density at Time 1, standardized by the ratio of the total area density from Time 2 to Time 1. RRI scores of 1 indicate a proportionate change in crimes that fall into predicted hotspots, relative to the overall change in the volume of crime within a study area. RRI scores < 1 indicate a decrease in hotspot precision. In short, RRI serves as a consistency check for predictive accuracy. RRI has been used in past research, in conjunction with PAI (Van Patten *et al.*, 2009); and we have taken a similar approach in the current study by including RRI as a way to contextualize PAI scores. Results of our analysis are as follows.

4. Results

Our assessment of the effects of KDE parameter settings on prospective hotspot maps began by examining how interpolation methods influenced predictive accuracy. Table II presents the average hit rate, PAI, and RRI across the four kernel density functions used in the current study. Results show that the predictive accuracy of KDE maps varies considerably by interpolation method. In terms of hit rate, for example, KDE maps that use a weighted, triangular distribution function produces the highest average predicted accuracy (39 percent), whereas the unweighted, uniform function produces the lowest (32 percent). This general pattern is also observed among PAI scores. However, RRI scores suggest that increased accuracy of the hotspot predictions come at the expense of precision.

Unlike the effects of kernel function, our findings indicate that grid cell size has little influence on the predictive accuracy of KDE hotspot maps. Table III presents the average predictive accuracy score for the three different grid cell size settings, across four interpolation methods. Results show little to no variation in predictive accuracy or precision, based on the number of cells contained in the grid overlay. For example, on an average, no predictive accuracy score changed more than 1.5 percent as grid cell size decreased (e.g. as the number of columns used in the grid overlay increased from 150 columns to over 300 columns).

We continued to investigate of the effects of KDE parameter settings on prospective hotspot maps by examining the influence of bandwidth on predictive accuracy. Table IV presents the average predictive accuracy scores for three different search radius settings (e.g. $\frac{1}{4}$ mile, $\frac{1}{2}$ mile, and 1 mile). As with previous results, findings are presented across the four kernel density functions examined. Results show that changing the search radius affects the predictive accuracy of KDE maps considerably; and is most pronounced in the PAI when the search radius was increased from $\frac{1}{4}$ mile to 1 mile. This increase in bandwidth resulted in a decline of more than 43 percent when the triangular

Table II. Average hit rate, predictive accuracy index (PAI), and recapture rate index (RRI) by kernel density function

Functions	Hit rate	PAI	RRI
Normal	33.12	2.98	0.87
Quartic	37.88	4.19	0.78
Triangular	38.51	4.18	0.78
Uniform	32.01	3.34	0.85

Note: Figures represent the average score for all crimes and all variations in KDE parameter settings tested

Table III. Average hit rate, predictive accuracy index (PAI), and recapture rate index (RRI) for three grid cell settings, by kernel density function

	150 columns			250 columns			RTM columns[a]		
Functions	Hit rate	PAI	RRI	Hit rate	PAI	RRI	Hit rate	PAI	RRI
Normal	33.24	2.99	0.87	33.10	2.98	0.87	33.03	2.97	0.86
Quartic	37.85	4.19	0.78	37.84	4.19	0.78	37.94	4.20	0.78
Triangular	38.57	4.18	0.78	38.49	4.17	0.78	38.48	4.17	0.78
Uniform	31.91	3.32	0.85	31.84	3.33	0.84	32.27	3.37	0.86

Notes: Figures represent the average score for all crimes and all search radii tested. [a]The column count for the RTM-based method varied from 316-324 columns distribution function was used, 44 percent when a normal distribution function was used, and nearly 50 percent when a uniform function was used.

Finally, we examined whether findings varied across crime type. Table V shows the highest and lowest (high/low) average predictive accuracy score for three different bandwidths, by interpolation method[13]. Results show that except for when the smallest search radius is used and when described in terms of a hit rate, KDE maps produced from robbery incidents were always associated with the highest predictive accuracy. These findings are constant across each interpolation method considered. Similar findings were observed when predictive accuracy was defined in terms of the PAI. With only one exception (e.g. quartic distribution function with a one-mile search radius), KDE robbery hotspot maps always had the highest accuracy score. Interestingly, with only one exception (e.g. quartic distribution function with a $\frac{1}{4}$-mile search radius), KDE motor vehicle hotspot maps always had the lowest PAI score. Finally, RRI scores indicate that predictive accuracy was most precise for aggravated assault, but least precise for motor vehicle theft.

Collectively, findings suggests that the predictive accuracy of KDE crime hotspot maps are fairly stable across crime type, generally producing higher scores – when the search radius exceeds $\frac{1}{4}$ mile – for interpersonal crimes like robbery than for property offenses like motor vehicle theft. The consistency of predictions was also observed to become more stable as bandwidth increased. That is, findings are constant across interpolation method and search radius, once it exceeds $\frac{1}{4}$ mile. A discussion of these findings is presented in Section 5.

5. Discussion and conclusion

Using crime event location data from Arlington, Texas, the current study informs a growing predictive policing literature by answering four questions related to the accuracy of crime hotspot maps produce by KDE. We began by investigating whether

Table IV. Average hit rate, predictive accuracy index (PAI), and recapture rate index (RRI) for three search radius settings, by kernel density function

	$\frac{1}{4}$ mile			$\frac{1}{2}$ mile			1 mile		
Functions	Hit rate	PAI	RRI	Hit rate	PAI	RRI	Hit rate	PAI	RRI
Normal	38.76	3.89	0.83	33.01	2.87	0.86	27.59	2.17	0.91
Quartic	40.48	5.03	0.65	39.03	4.26	0.81	34.13	3.29	0.88
Triangular	41.52	5.31	0.66	39.42	4.20	0.81	34.60	3.02	0.87
Uniform	35.68	4.55	0.79	32.66	3.12	0.90	27.68	2.34	0.86

Note: Figures represent the average score for all crimes and all grid sizes tested

Table V. Highest/lowest average hit rate (HR), predictive accuracy index (PAI), and recapture rate index (RRI) for three search radius settings, by kernel density function

	$\frac{1}{4}$ mile			$\frac{1}{2}$ mile			1 mile		
Functions	HR	PAI	RRI	HR	PAI	RRI	HR	PAI	RRI
Normal	R/MVT	R/MVT	AA/R	R/MVT	R/MVT	AA/MVT	R/MVT	R/MVT	AA/MVT
Quartic	CB/R	R/AA	CB/R	R/MVT	R/MVT	AA/MVT	R/MVT	AA/MVT	AA/MVT
Triangular	CB/R	R/MVT	CB/R	R/MVT	R/MVT	AA/MVT	R/MVT	R/MVT	AA/MVT
Uniform	CB/MVT	R/MVT	AA/MVT	R/CB	R/MVT	AA/R	R/MVT	R/MVT	AA/MVT

Notes: AA, aggravated assault; R, robbery; CB, commercial burglary; MVT, motor vehicle theft. Figures represent the average score for all grid sizes tested interpolation method affects the predictive accuracy of KDE hotspot maps. Depending on the interpolation method employed, considerable differences in the ability of KDE hotspot maps to predict future crime events – based on past events – were evident. Second, we observed that regardless of interpolation method, variation in grid cell size had almost no affect on the predictive accuracy of KDE hotspot maps or the reliability of those predictions. However, we observed considerable variation in predictive accuracy when bandwidth varied. Finally, we found the effects of parameter settings used for KDE hotspot maps were constant across different crime types. In most cases, KDE hotspot maps were most successful at predicting personal violence like robbery or aggravated assault and least successful at predicting property crimes such as commercial burglary and motor vehicle theft. For example, 12 combinations of search radius settings and kernel density functions were considered for each crime type. The highest hit rate for nine of 12 combinations (75 percent) and the highest PAI for 11 of the 12 combinations (92 percent) were associated robbery. The precision of KDE maps followed similar patterns.

Based on results from the current study, academics and practitioners involved in prospective hotspot mapping and that utilize KDE as their analytic approach should consider the following recommendations. First, since kernel density function affected predictive accuracy, we recommend using either a quartic or triangular distribution for predicting crime events as they were the methods that produced consistently high predictive accuracy scores, despite being slightly less precise. In contrasts, normal and uniform distribution interpolation methods should be avoided, as they tended to underperform other kernel functions.

Second, we recommend setting the cell size of the grid overlay to approximately one-third the length of the average blockface of the study area. While it will likely make little difference in terms of predictive accuracy if this recommendation is adopted in lieu of leaving the default setting for grid cell size unchanged, we feel it is a recommendation that is more consistent with the current literature on the micro-level dynamics of victimization. From a practical standpoint, increasing the default grid cell size setting may increase the visual quality of the KDE map (e.g. make it look less pixelated); but the predictive accuracy will not be enhanced in a meaningful way. Keep in mind, however, analysis will be slower as a result of the additional processing time that will required to compute additional density calculations for the additional cells added to the grid overlay when the default grid cell size is decreased.

Third, we recommend using a small bandwidth, as current findings suggest that successful prediction of future crime events generally declined as search radius increases. A similar general pattern has been observed in previous research (Chainey *et al.*, 2008a). In our study, a $\frac{1}{4}$-mile bandwidth out performed the other two settings considered. The default bandwidth setting in ArcGIS is determined by dividing the smaller of the length and width of the study area by 30 and consistently produced a search radius of slightly more than a $\frac{1}{4}$ mile. Modifying this formula so that the denominator 50 produced a search radius of slightly less than a $\frac{1}{4}$ mile. Therefore, we suggest that a standard search radius, equal to the smaller of the length or width of a study area, divided by between 30 and 50 be used for determining an appropriate KDE bandwidth.

Finally, although crime predictions were relatively stable across parameter settings, successful crime prediction was crime-type dependent. Robbery was consistently predicted more successfully than the other crimes considered, although with lower reliability. Motor vehicle theft, on the other hand, was consistently predicted less accurately. This observation is not surprising, given similar results found in past research (Chainey *et al.*, 2008a). Therefore, it is recommended that KDE be used with caution when attempting to predict property crimes like motor vehicle theft and burglary.

Our four recommendations are considered collectively and summarized in Table VI. Information contained in this table shows the combination of all parameter settings considered in the current study and the highest corresponding hit rate, PAI, and RRI for each crime type. Researchers and practitioners should use this information as a starting point when investigating the effects of parameter settings on KDE hotspot maps within their specific study areas.

Although findings from the current study are informative, additional research in this area is warranted, especially in light of some of the study's limitations. First, data used in this investigation are limited to a single geographic area. APD is a large law enforcement agency, employing approximately 600 full-time sworn law enforcement officers. The agency serves a population of slightly more than 350,000, spread over approximately 100 square miles. Although many agencies share similar characteristics, the use of a single agency limits the generalizability of our findings. Therefore, future research should seek to replicate results observed in the current study in order to determine the extent to which urban morphology affects the predictive accuracy of KDE hotspot maps.

Second, KDE maps produced in the current study were based on a single predictor: crimes recorded in 2007. Developing composite KDE maps that include multiple risk factors would have likely yielded different predictive accuracy scores. In addition, the predictive accuracy scores reported in the current study may reflect underestimations of the effectiveness of KDE maps if APD engaged in intervention strategies that focussed on crime clusters within their jurisdiction that were observed between 2007 and 2008[14].

Third, although considerable effort went into geocoding crime events against positionally accurate reference data, not all crime events recorded by APD were successfully matched. Hart and Zandbergen (2013) recently demonstrated the effects of geocoding on the predictive accuracy of crime hotspot mapping and current findings should be considered in light of this information. Specifically, it is unclear whether the positional error associated with the ungeocoded crime data provided by APD is normally distributed. If it is not, then current findings may not be unbiased.

Finally, although the current study considered several parameter settings in the KDE process, the influence on predictive accuracy related to other aspects of the analytic technique were not investigated. For example, duel KDE is used when two

	Highest score		
	HR	PAI	RRI
Aggravated assault	RTM/T/0.25	RTM/T/0.25	RTM/U/0.50
Robbery	RTM/T/0.50	RTM/Q/0.25	250/Q/1
Commercial burglary	150/T/0.25	250/Q/0.25	150/N/1
Motor vehicle theft	RTM/T/0.25	RTM/Q/0.25	150/U/0.50

Notes: Number of columns (150 = 150 columns; 250 = 250 columns; and RTM, average streetline segment divided by 0.33). Density function (N, normal; Q, quartic; T, triangular; and U, uniform). Search radius (0.25 = $\frac{1}{4}$ mile, 0.50 = $\frac{1}{2}$ mile, and 1 = 1 mile)

Table VI. Combination of parameter settings (i.e. number of columns/density function/search radius) that yield the highest hit rate (HR), predictive accuracy index (PAI), and recapture rate index (RRI), by crime type interpolated maps are expressed as a ratio of each other and is the desired approach to normalizing risk (see Bruce and Smith, 2011; Levine, 2004). In the current study, for example, the number of commercial properties within the study area could have been used to normalize KDE predictions for commercial burglary. Normalizing risk of commercial burglary could have affected KDE predictions. However, the use of the duel KDE technique was beyond the scope of the current study. Future research in this area is therefore warranted and should focus primarily on property crimes, given current findings.

In conclusion, the current investigation advances our knowledge and understanding of prospective hotspot crime mapping as it answers the calls by Chainey *et al.* (2008a) and others to further investigate methods used to forecast crime. Unfortunately, the speed at which GIS is being embraced and utilized by both academics and practitioners is outpacing the scientific literature on important methodological issues related to popular techniques. Research must therefore strive to keep pace in order to provide scientific guidance to an ever-growing audience.

Notes

1. Skepticism over the utility of "predictive policing" techniques like prospective hotspot mapping exists among both researchers and practitioners. However, the goal of the current study is not to test the merits of predictive policing; rather, the current study: first, quantifies the effects of parameters settings on the predictive accuracy of KED; and second, offers empirically based recommendations to those currently using or planning to use KDE to forecast crime. Reviews of the effectiveness of predictive hotspot mapping have been offered by Braga (2007) and Braga *et al.* (2012).
2. Evidence exists that the predictive accuracy of crime hotspot maps maybe crime-type dependent. For example, Johnson *et al.* (2009) successfully developed a predictive individual-level model based on optimal foraging behavior of repeat offenders, but the approach was specifically designed for burglaries and may not apply to other types of offenses. Similarly, Johnson *et al.* (2008) determined substantial variation in the stability of crime hotspots, which they expect vary by type of crime.
3. Braga (2007) and Braga *et al.* (2012) provide comprehensive reviews of existing research on the effects of hotspot policing on crime. They conclude that focussing police efforts in crime hotspot areas can effectively prevent and reduce crime as well as calls for police service. However, their investigations do not identify the particular analytic techniques used in each study to identify crime hotspots.
4. Bandwidths are defined as either "adaptive" or "fixed interval". Adaptive bandwidths are typically used when a sample of point locations are used in lieu of all points located within the study area, which is generally not the case in crime analysis. If adaptive bandwidths are used, the minimum sample size is an additional parameter that must be defined. Fixed interval bandwidths, on the other hand, are used when the entire population of events is analyzed (e.g. all crimes within a given timeframe). When a fixed interval bandwidth is selected, the size of the bandwidth must be defined (see Brunsdon, 1995).
5. Instead of asking users to define the actual cell size, most GIS applications ask the user to indicate how many columns the grid overlay should be divided into. This input value is then used to calculate the actual size of each grid cell.
6. The current study uses three measures of predictive accuracy: hit rate, PAI (Chainey *et al.*, 2008a), and RRI (Levine, 2008). Each is discussed in greater detail in Section 3.
7. That is to say $(5{,}280\,\text{ft} \times 3)/150 = 105.6\,\text{ft}^2$. Given the shortest side of the extent is three miles, the number of columns by which the grid overlay would be divided is 50 ($15{,}840\,\text{ft}/105.6\,\text{ft}^2 = 150$ columns).

8. Other functions exist, including negative exponential, Epanechnikov, triweight, tricube, cosine, and Gaussian.
9. The overall area of each grid cell is decreased as the bounding area is divided into more columns.
10. Analysis of the street centerline file for Arlington, Texas indicated that the average blockface in our study area was approximately 480 ft; and the corresponding number of columns varied based on the extent: 316 for aggravated assault, 319 for commercial burglary, 324 for robbery, and 348 for motor vehicle theft.
11. ArcGIS, for example, uses the quartic distribution when conducting KDE. Although the default interpolation method can be changed, this type of modification requires more advanced programming knowledge.
12. Different classification methods can be used to define hotspots (e.g. equal intervals, natural breaks, manual intervals, and standard deviations). Classifying data as either in the top 5 percent of crime clusters or not and breaking the data on that dichotomy allowed us to make comparisons across each KDE map because maps were standardized.
13. Appendix presents the data contained in Table V disaggregated for each crime type and kernel function.
14. According to J. Mallard, former Manager of APD's Crime Analysis Unit, prospective crime mapping techniques were not used between 2007 and 2008 specifically to allocate police resources designed to target crime hotspots (personal communication on July 12, 2013).

References

Anselin, L. (1995), "Local indicators of spatial association – LISA", *Geographical Analysis*, Vol. 27 No. 2, pp. 93-115.

Bailey, T. and Gatrell, A. (1995), *Interactive Spatial Data Analysis*, Longman Scientific and Technical, Essex.

Ball, G.H. and Hall, D.J. (1970), "A clustering technique for summarizing multivariate data", *Behavioral Science*, Vol. 12 No. 2, pp. 153-155.

Bates, S. (1987), *Spatial and Temporal Analysis of Crime*, Research Bulletin, Illinois Criminal Justice Information Authority, Chicago, IL.

Braga, A., Papachristos, A. and Hureau, D. (2012), *Hot Spots Policing Effects on Crime*, Campbell Systematic Reviews, The Campbell Collaboration, Oslo.

Braga, A.A. (2007), *The Effects of Hot Spots Policing on Crime*, Campbell Systematic Reviews, The Campbell Collaboration, Oslo.

Bowers, K.J., Johnson, S.D. and Pease, K. (2004), "Prospective hot-spotting: the future of crime mapping?", *British Journal of Criminology*, Vol. 44 No. 5, pp. 641-658.

Bruce, C.W. and Smith, S.C. (2011), "Spatial statistics in crime analysis: using CrimeStatIII", available at: www.icpsr.umich.edu/CrimeStat/workbook/CrimeStat_Workbook.pdf (accessed September 1, 2012).

Brunsdon, C. (1995), "Estimating probability surfaces for geographical point data: an adaptive kernel algorithm", *Computers & Geosciences*, Vol. 21 No. 7, pp. 877-894.

Canter, P. (1993), "Hot spot area analysis in Baltimore County", STACNews, Vol. 1 No. 1, available at: www.icjia.state.il.us/public/index.cfm?metastory=1628

Caplan, J.M. and Kennedy, L.W. (2010), *Risk Terrain Modeling Manual: Theoretical Framework and Technical Steps of Spatial Risk Assessment*, Rutgers Center on Public Security, Newark, NJ.

Caplan, J.M., Kennedy, L.W. and Baughman, J. (2012), "Kansas City's violent crime initiative: a place-based evaluation of location-specific intervention activities during a fixed time period", *Crime Mapping: A Journal of Research and Practice*, Vol. 4 No. 2, pp. 9-37.

Caplan, J.M., Kennedy, L.W. and Miller, J. (2011), "Risk terrain modeling: brokering criminological theory and GIS methods for crime forecasting", *Justice Quarterly*, Vol. 28 No. 2, pp. 360-381.

Chainey, S. and Ratcliffe, J.H. (2005), *GIS and Crime Mapping*, Wiley, San Francisco, CA.

Chainey, S., 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.), *Socio-Economic Applications of Geographic Information Science*, Taylor & Francis, London, pp. 22-36.

Chainey, S., Tompson, L. and Uhlig, S. (2008a), "The utility of hotspot mapping for predicting patterns of crime", *Security Journal*, Vol. 21 No. 1, pp. 4-28.

Chainey, S., Tompson, L. and Uhlig, S. (2008b), "Response to Pezzuchi", *Security Journal*, Vol. 21 No. 4, pp. 293-294.

Chainey, S., Tompson, L. and Uhlig, S. (2008c), "Response to Levine", *Security Journal*, Vol. 21 No. 4, pp. 303-306.

Eck, J.E., Chainey, S., Cameron, J.G., Leitner, M. and Wilson, R.E. (2005), *Mapping Crime: Understanding Hotspots*, National Institute of Justice Special Report, National Institute of Justice, Washington, DC.

Edmonds, W. and Mallard, J. (2011), "Using risk terrain modeling to analyze holiday robberies in Arlington, TX", Paper presented at the Training Conference of the International Association of Crime Analysts, Hyannis, MA, September.

Felson, M. (1995), "Those who discourage crime", in Eck, J.E. and Weisburd, D. (Eds), *Crime and Place: Crime Prevention Studies, Vol. 4*, Willow Tree Press, Monsey, NY, pp 349-359.

Getis, A. and Ord, J.K. (1992), "The analysis of spatial association by use of distance statistics", *Geographical Analysis*, Vol. 24 No. 3, pp. 189-206.

Groff, E.R. and LaVigne, N.G. (2002), "Forecasting the future of predictive crime mapping", *Crime Prevention Studies*, Vol. 13, pp. 29-57.

Hart, T.C. and Zandbergen, P.A. (2013), "Reference data and geocoding quality: examining completeness and positional accuracy of street geocoded crime incidents", *Policing: An International Journal of Police Strategies & Management*, Vol. 36 No. 2, pp. 263-294.

Hartigan, J.A. (1975), *Clustering Algorithms*, John Wiley & Sons, New York, NY.

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

Johnson, S.D. and Bowers, K.J. (2004a), "The burglary as a clue to the future: the beginnings of prospective hot-spotting", *The European Journal of Criminology*, Vol. 1 No. 1, pp. 237-255.

Johnson, S.D. and Bowers, K.J. (2004b), "The stability of space-time clusters of burglary", *British Journal of Criminology*, Vol. 44 No. 1, pp. 55-65.

Johnson, S.D, Bowers, K.J., Birks, D.J. and Pease, K. (2009), "Predictive mapping of crime by ProMap: accuracy, units of analysis and the environmental backcloth", in Weisburd, D., Bernasco, W. and Bruinsma, G.J.N. (Eds), *Putting Crime in its Place: Units of Analysis in Geographic Criminology*, Chapter 8, Springer, New York, NY, pp. 171-198.

Johnson, S.D., Lab, S.P. and Browers, K.J. (2008), "Stable and fluid hotspots of crime: differentiation and identification", *Built Environment*, Vol. 34 No. 1, pp. 32-45.

Kennedy, L.W., Caplan, J.M. and Piza, E. (2011), "Risk clusters, hotspots, and spatial intelligence: risk terrain modeling as an algorithm for police resource allocation strategies", *Journal of Quantitative Criminology*, Vol. 27 No. 3, pp. 339-362.

Levine, N. (2004), *CrimeStatIII: A Spatial Statistics Program for Analyzing Crime Incident Locations*, National Institute of Justice, USGPO, Washington, DC.

Levine, N. (2008), "The 'hottest' part of a hotspot: comments on 'the utility of hotspot mapping for predicting spatial patters of crime' ", *Security Journal*, Vol. 21 No. 4, pp. 295-302.

McBratney, A.B. and deBruijter, J.J. (1992), "A continuum approach to soil classification by modified fuzzy k-means with extragrades", *Journal of Soil Science*, Vol. 43 No. 1, pp. 159-175.

McLafferty, S., Williamson, D. and Maguire, P.G. (2000), "Identifying crime hot spots using kernel smoothing", in Goldsmith, V. *et al.* (Eds), *Analyzing Crime Patterns: Frontiers of Practice*, Sage, Thousand Oaks, CA, pp. 77-85.

Mohler, G.O., Short, M.B., Brantingham, P.J, Schoenberg, F.P. and Tita, G.E. (2011), "Self-exciting point process modeling of crime", *Journal of the American Statistical Association*, Vol. 106 No. 493, pp. 100-108.

Ord, J.K. and Getis, A. (1995), "Local spatial autocorrelation statistics: distributional issues and an application", *Geographical Analysis*, Vol. 27 No. 4, pp. 286-306.

Parzen, E. (1962), "On estimation of a probability density function and mode", *Annals of Mathematical Statistics*, Vol. 33 No. 3, pp. 1065-1076.

Pezzuchi, G. (2008), "A brief commentary on 'the utility of hotspot mapping for predicting spatial patterns of crime'", *Security Journal*, Vol. 21 No. 4, pp. 291-292.

Ratcliffe, J.H. (2004), "Geocoging crime and a first estimate of a minimum acceptable hit rate", *International Journal of Geographical Information Science*, Vol. 18 No. 1, pp. 61-72.

Ratcliffe, J.H. (2009), *Near Repeat Calculator (Version 1.3)*, Temple University, Philadelphia, PA and the National Institute of Justice, Washington, DC.

Ratcliffe, J.H. (2010), "The spatial dependency of crime increase dispersion", *Security Journal*, Vol. 23 No. 1, pp. 18-36.

Ratcliffe, J.H. and Rengert, G.F. (2008), "Near repeat patterns in Philadelphia shootings", *Security Journal*, Vol. 21 Nos 1-2, pp. 58-76.

Reaves, B.A. (2010), *Local Police Departments, 2007*, Bureau of Justice Statistics, USGPO, Washington, DC.

Rosenblatt, M. (1956), "Remarks on some nonparametric estimates of a density function", *Annals of Mathematical Statistics*, Vol. 27 No. 3, pp. 832-837.

Spring, J.V. and Block, C.R. (1989), *STAC Users Manual*, Illinois Criminal Justice Information Authority, Chicago, IL.

Taylor, R.B. (1997), "Order and disorder of streetblocks and neighborhoods: ecology, microecology and the systemic model of social organization", *Journal of Research in Crime and Delinquency*, Vol. 34 No. 1, pp. 113-155.

Thompson, H.R. (1956), "Distribution of distance to nth neighbour in a population of randomly distributed individuals", *Ecology*, Vol. 37 No. 2, pp. 391-394.

Van Patten, I.T., McKeldin-Coner, J. and Cox, D. (2009), "A microspatial analysis of robbery: prospective hotspot mapping in a small city", *Crime Mapping: A Journal of Research and Practice*, Vol. 1 No. 1, pp. 7-32.

Ward, J.H. (1963), "Hierarchical grouping to optimize an objective function", *Journal of the American Statistical Association*, Vol. 58 No. 301, pp. 236-244.

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*, Taylor & Francis, London, pp. 187-202.

Yerxa, M. (2013), "Evaluating the temporal parameters of risk terrain modeling with residential burglary", *Crime Mapping: A Journal of Research and Practice*, Vol. 5 No. 1, pp. 7-38.

Further reading

Anselin, L., Cohen, J., Cook, D., Gorr, W. and Tita, G. (2000), "Spatial analyses of crime", *Criminal Justice*, Vol. 4 No. 2, pp. 213-262.

Brantingham, P.J. and Brantingham, P.L. (1991), *Environmental Criminology*, 2nd ed., Waveland Press, Prospect Heights, IL.

Brantingham, P.L. and Brantingham, P.J. (1999), "A theoretical model of crime hot spot generation", *Studies on Crime and Crime Prevention*, Vol. 8, pp. 7-26.

Groff, E.R. and LaVigne, N.G. (2001), "Mapping an opportunity surface of residential burglary", *Journal of Research in Crime and Delinquency*, Vol. 38 No. 3, pp. 257-278.

Hickman, M.J. and Reaves, B.A. (2006), *Local Police Departments, 2003*, Bureau of Justice Statistics, USGPO, Washington, DC.

Miethe, T.D. and Deibert, G.R. (2006), *Fight Time: The Normative Rules and Routines of Interpersonal Violence*, Waveland, Long Grove, IL.

Rossmo, D.K. (2000), *Geographic Profiling*, CRC Press, Boca Raton, FL.

Rossmo, D.K. (2004), "Geographic profiling", in Thurman, Q. and Zhao, J. (Eds), *Contemporary Policing: Controversies, Challenges, and Solutions*, Roxbury Publishing Co, Los Angeles, CA, pp. 274-284.

Rossmo., D.K. (2011), "Evaluating geographic profiling", *Crime Mapping: A Journal of Research and Practice*, Vol. 3 No. 1, pp. 42-65.

Appendix

	$\frac{1}{4}$ mile			$\frac{1}{2}$ mile			1 mile		
Crime type and function	Hit rate	PAI	RRI	Hit rate	PAI	RRI	Hit rate	PAI	RRI
Aggravated assault									
Normal	39.32	4.26	0.94	33.02	2.88	0.94	28.92	2.20	1.01
Quartic	41.29	3.89	0.65	38.77	4.57	0.90	33.10	4.15	0.97
Triangular	42.16	5.19	0.65	40.11	4.66	0.92	33.49	3.02	0.96
Uniform	34.04	4.63	0.88	31.28	3.35	1.07	27.97	2.32	0.94
Robbery									
Normal	41.95	4.79	0.79	39.49	3.50	0.85	33.13	2.66	0.85
Quartic	37.64	6.38	0.59	42.46	5.34	0.78	41.54	3.86	0.87
Triangular	38.46	6.32	0.60	42.97	5.27	0.79	40.72	3.73	0.83
Uniform	36.82	6.05	0.75	39.69	3.79	0.83	32.00	2.76	0.84
Commercial burglary									
Normal	39.67	3.58	0.84	31.14	2.67	0.88	26.61	2.05	0.96
Quartic	44.51	5.18	0.70	39.88	3.81	0.81	31.03	2.65	0.88
Triangular	45.49	5.12	0.70	39.85	3.72	0.81	32.39	2.74	0.89
Uniform	37.83	4.02	0.78	29.57	2.77	0.88	29.61	2.39	0.92
Motor vehicle theft									
Normal	34.11	2.94	0.77	28.41	2.43	0.79	21.71	1.78	0.82
Quartic	38.48	4.64	0.68	34.99	3.31	0.74	30.84	2.48	0.80
Triangular	39.95	4.59	0.69	34.75	3.16	0.74	31.81	2.58	0.79
Uniform	34.03	3.52	0.74	30.09	2.58	0.83	21.12	1.90	0.75

Table AI.
Hit rate, predictive accuracy index (PAI), and recapture rate index (RRI) for three search radius settings, by crime type and kernel density function

About the authors
Dr Timothy Hart is a Senior Lecturer in the School of Criminology and Criminal Justice at the Griffith University. He received his PhD in Criminology from the University of South Florida in 2006. His areas of interest include survey research, applied statistics, geographic information systems, and victimization. Prior to joining the faculty at the Griffith University, Dr Hart worked

as a Statistician for the Bureau of Justice Statistics, a Program Analyst for the Drug Enforcement Administration, and a Research Analyst for the Hillsborough County Sheriff's Office. In addition, he is the former Statistical Analysis Centre Director for the state of Nevada. Dr Timothy Hart is the corresponding author and can be contacted at: t.hart@griffith.edu.au

Dr Paul Zandbergen is an Associate Professor in the Department of Geography at the University of New Mexico. He obtained his PhD in Resource Management and Environmental Studies in 1998 at the University of British Columbia in Vancouver, Canada. His research interests include both the fundamentals of GI science as well as the applications of geospatial technologies to several fields, including water resources, spatial ecology, environmental health, and criminal justice. His current research focusses on issues of scale, error, and uncertainty in spatial analysis, as well as on the robustness of spatial analytical techniques.