The Influence of Data Classification Methods on Predictive Accuracy of Kernel Density Estimation Hotspot Maps

Nenad Milic[1], Brankica Popovic[2], Sasa Mijalkovic[1], and Darko Marinkovic[1]
[1]Department of Criminalistics, University of Criminal Investigation and Police Studies, Serbia
[2]Department of Informatics and Computer Science, University of Criminal Investigation and Police Studies, Serbia

Abstract: *When it comes to hot spot identification, spatial analysis techniques come to the fore. One of such techniques, that has gained great popularity among crime analysts, is the Kernel Density Estimation (KDE). Small variation in KDE parameters can give different outputs and hence affect predictive accuracy of hotspot map. The influence these parameters have on KDE hotspot output sparked many researches, mostly analyzing the influence of the cell size and bandwidth size. Yet, the influence of different classification methods applied to calculated cell values, including the choice of threshold value, on the KDE hotspot predictive accuracy remained neglected. The objective of this research was to assess the influence of different classification methods to KDE predictive accuracy. In each KDE computation, calculated cell values were divided into five thematic classes, using three the most common (default) classification methods provided by Environmental Systems Research Institute (ESRI) Geographical Information System (Arc GIS) (equal interval classification, quantile classification and natural breaks classification) and incremental multiples of the grid cells' mean. Based upon calculated hit rates, predictive accuracy indices and recapture rate indices and taking into account the necessity that mapping output should satisfy some operational requirements as well as statistical rules, this research suggest that incremental mean approach with hotspot threshold of 3 and above multiples of the grid cell's mean, should be used.*

Keywords: *Crime mapping, hot spot, kernel density, classification methods.*

Received May 20, 2016; accepted May 6, 2018

1. Introduction

Crime distribution is not random. Crimes tend to concentrate at particular geographic locations where favorable opportunities exist. These concentrations or clusters of crime are commonly referred to as hotspots [5]. Proliferation of Geographical Information System (GIS) software [9] contributes to the fast and easy hotspot maps creation, making them a central part of crime analysis and hotspot policing.

Today a variety of techniques are available for hotspots identification and analysis. They can be classed in two broad groups: those applied on point data and those applied on aggregated data [19], but still, there is no agreement among researchers which one of them is the best in terms of accuracy in predicting future crimes.

Among different mapping techniques, Kernel Density Estimation (KDE) is one of the most popular, both with academics and the crime analyst professionals. This is the most suitable technique for visualizing crime data and according to some research its predictive "strength" outperformed other methods [6]. While assessing the predictive accuracy of hotspot mapping techniques, it was found that it depends only to some extent on the predictive crime mapping techniques itself and that there are other factors that must be taken into account, among them the influence of used parameters is the most pronounced [7, 14, 15, 25, 28]. The fact that many of the parameters used to determine a hotspot are subjective threatens to undermine the reliability of the hotspot methods. Consequently, instead of looking for the best hotspot method, research efforts should be concentrated over the issue which parameters are the most appropriate in certain environment (e.g., study area characteristics, different crime types etc.,) in order to provide clear and empirically tested guidance on the parameters selection. Hence, in this research predictive ability of different methods was not evaluated in order to find the best one. Instead, focus was placed on one hotspot method-KDE- in order to analyze the influence of one of its parameters (data classification methods) on its predictive accuracy. In crime mapping literature this influence is still under-researched.

The remainder of this paper is structured as follows: the second section offers a brief discussion of the predictive metrics used to assess hot spot mapping techniques; the third section outlines the parameters utilized in KDE technique; data and methodology used in this research are presented in section four and obtained results and their interpretation are given in the fifth section of this paper. Finally, a conclusion is presented to summarize the main outcomes of this research.

2. Predictive Ability of Hotspot Mapping Techniques

In the essence of hotspot policing is the idea of focusing policing resources to problematic areas in order to identify and "change the underlying conditions, situations, and dynamics that make them attractive to criminals and disorderly persons" [2]. Inaccuracy in hot spots identifications may affect police effectiveness, as well as the citizens' quality of life, or even their rights[1].

Assuming that locations of past events may be good predictors of the time and place of future events, hotspot techniques are often evaluated by their ability to predict future crimes based on historic crime data. Answer to the question to what extent identified hotspot is able to predict future crime incidents could indicate predictive accuracy of the used hotspot technique. For that purpose different predictive measures can be utilized. Three most frequently used are:

- Hit Rate (HR), defined as the proportion of new crimes that occurs within the areas where crimes were predicted to occur:

$$HR = \frac{n}{N} \tag{1}$$

Where n is the number of crimes in areas where crimes are predicted to occur (hotspots) and N is the number of crimes in the whole study area:

- Predictive Accuracy Index (PAI) [6], described as the ratio of the hit rate to the proportion of the study area that consists of hotspots:

$$PAI = \frac{HR}{proportionon\_of\_hot\_spot\_area} = \frac{n/N}{a/A} \tag{2}$$

Where a is the total area occupied by hotspots, and A is the size of entire study area:

- Recapture Rate Index (RRI) [18], defined as a ratio of predicted (for time period 2 and labelled as n_2) and historic (for time period 1 and labelled as n_1) hotspot densities, standardized for changes of the total area density in each year (for time period 2 labelled as N_2 and for time period 1 labelled as N_1):

$$RRI = \frac{hot\_spot\_crime\_ratio}{total\_crime\_ratio} = \frac{n_2/n_1}{N_2/N_1} \tag{3}$$

It is often neglected that pursuance of high predictive accuracy may undermine operational utility of produced hotspot output. This means that hotspot output having great scores in applied predictive metrics may at the same time be of limited use in (police) practice [26]. In line with this it was noted that large hotspot areas are "not only too big to treat, they are also too large in size and too small in number to evaluate the effects of treatments" [22]. On the other side, too small hotspots may indicate that crime problem is not (fully) identified, which means that some areas continue to produce crime incidents, endangering the citizen's lives and property and police, unaware of existing crime problem, fails to protect them. Therefore, those who are striving to develop "predictive excellence" in their hotspot mapping outputs must not neglect their operational utility. As a consequence, produced mapping output (hotspots) will not be actionable, limiting the ability of police decision maker to effectively focus resources. In that case what we get is "a good hotspot and bad hotspot policing".

Today the field of predictive policing is in expansion. New ideas are popping out faster than ever before, while the older ones receive increasing empirical support. For example, predictive ability of repeat and near repeat patterns [1] was tested for different crime types [12, 16, 17, 21], holding the promise of being the cornerstone of successful crime prevention initiatives, in particular those intended to solve burglary problems. Also, Risk Terrain Modeling (RTM) is today one of the most prominent ideas in the field of predictive policing, basing its risk assessment not on prior crime histories, but on the environmental risks [3]. Nevertheless because of its functionality and popularity, KDE remains unavoidable 'player' in predictive policing 'arena'. This justifies the necessity of further research in order to improve its predictive accuracy.

3. KDE Hotspot Identification

The KDE technique produces a continuous surface, representing the density of crime events over the study area. This is created by 'overlaying a grid (with n equally sized cells) on top of the study area and calculating a density estimates 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 (the kernel function) and the bandwidth (search radius)' [12]. The grid cells are then shaded according to their density scores to create a density map with a smooth surface.

The predictive 'strength' of KDE was evaluated by many researchers. It was suggested that that 'KDE method predicts future crime incidents the best if the

[1]For example, placing the label "high crime area" on safe area may cause stigmatizing effect, which may hinder economic development of the particular neighbourhood. In the USA this label can be a factor for "reasonable suspicion" to stop a suspect.

PAI and the RRI are applied' [10], and that 'as a longer time period was used for the prediction base, the KDE methods became stronger' [25]. On the other side, there are findings that challenged the growing consensus that KDE produces the most accurate hotspot predictions. It was suggested that KDE 'is useful for a citywide or regional view of crime concentrations, but is not adequate for defining the specific boundaries of hotspot' [18], leading to conclusion that 'KDE does produce hotspot maps with the highest PAI, but only for analysis that:

- Uses a small search radius.
- Rely data geocoded against an address point reference layer.
- For certain study areas'[14].

KDE requires three parameters to be determined before it can be used on crime data:

1. The grid cell size.
2. Bandwidth (also known as the search radius).
3. Calculation (interpolation) method. They affect calculation of the density surface, and, together with.
4. Classification scheme (thematic class).

They influence how final output (density map) will look like.

Cell sizes should be aligned to the scale at which the output will be viewed. Large cell sizes will result in more coarse or blocky-looking maps but are fine for large-scale output, while smaller cell sizes add to the visual appeal of the continuous surface produced but may create large file sizes [8].

Bandwidth is the parameter that will lead to most differences in output when it is varied [7]. The larger the bandwidth, the more generalized the patterns in the density surface will be. With a larger bandwidth, the GIS consider more incidents when calculating the value of each cell. A smaller bandwidth usually shows more local variation. However, if the search radius is so small that most cells have very low density values, broader patterns in the data may not be visible. Considering the bandwidth size, one should consider recommendation that if the sample size is large, a smaller bandwidth is appropriate whereas a larger bandwidth is more desirable for smaller samples [19].

According to crime mapping literature, the parameter which attracts the least research interest in applying KDE is the choice of kernel function (interpolation method). The reasons for that may be that different kernel functions make small differences in kernel outputs [19]. However recent research showed that kernel function affects predictive accuracy of the hot spot output, so variations of this parameter should not be neglected [15].

The last, but not the least important parameter that can influence KDE output is the selection of thematic range settings or classification system used by map maker to classify calculated (density) values, as different choices can produce vastly different map outputs. This parameter will be further discussed below.

3.1. Role of Classification Methods and Hotspot Threshold

One consistent criticism with the hotspot methods is that many of the parameters used to determine a hotspot are subjective. Having in mind that small variation in KDE parameters can give different outputs and hence affect predictive accuracy of hotspot map, the influence these parameters have on KDE hotspot output sparked many researches who mainly paid attention to the influence of the cell size and bandwidth size [7, 15], but the influence of different classification methods applied to calculated data value range, including the choice of threshold value, on the KDE hotspot predictive accuracy remained neglected.

Unlike selecting appropriate cell size, bandwidth and kernel function, parameters that precede kernel density calculation, classification and visualization of the grid cell's values follow after and also require some subjective decisions. The user should decide:

- How many classes data should be categorized into.
- What the value ranges of those classes should be.
- What is the threshold value which separates 'hot' values from other values (values above threshold are used to determine hotspot areas).

Like the others, these parameters can strongly influence KDE output.

The classification method is almost always more problematic than the number of groups. Generally, in the field of crime mapping no more than six and not less than four classes is used [13, 20]. Classification of data is usually done applying default classification methods[2] provided in the most GIS software:

- *Equal interval classification* (divides the range of values into equal-sized classes).
- *Quantile classification* (each class contains an equal number of features).
- *Natural breaks classification* (class breaks maximize the differences between classes and minimize the within-class differences).
- *standard deviation classification* (finds the mean value, and then places class breaks above and below the mean at intervals of either 0.25, 0.5 or one standard deviation until all the data values are contained within the classes) etc.

For example, in the analysis of robberies reported to the police in the City of Roanoke, VA for the years

[2] Besides the default classification methods, in the most GIS software the user has possibility to manually define class ranges.

2004 through 2007, the range of density values was divided into five equal sized groups, each with 20% of the data range. Hot spots are then defined as the highest valued class, or the densities at or above the 80% [25]. Similarly, applying KDE to thefts from motor vehicles in Lincoln, NE in 2007, it was found that five classes, using the equal interval classification tends to result in the best display [23]. While researching the spatial-temporal prediction of various crime types in Houston, TX based on hot-spot techniques, Fan used standard deviation classification scheme and the class greater than three standard deviations are defined as hotspots [10]. Chainey *et al*. used five thematic classes and quantile classification scheme and values (cells) in the top thematic class were determined as ‘hot’ [6]. In their research of the patterns of reported crime over a time and space within the Rochester, NY, from 2005 through 2011, Ghosh et al. applied natural breaks classification method. Cell values were sorted from lowest value to highest and large gaps or ‘natural breaks’ were used to define class breaks [11].

Unlike previously stated research examples, Chainey *et al*. [4] suggested the application of incremental multiples of the grid cells' mean to help standardize thematic thresholds. Calculations for the mean are applied only to grid cells that have a value of greater than 0 and that are within the study area boundary. Grid cell thematic thresholds should be set in the following way: 0 to 1 mean, 1 mean to 2 mean, 2 mean to 3 mean, 3 mean to 4 mean, 4 mean to 5 mean and greater than 5 mean since the mean is an easier value for the novice map reader to grasp.

The influence different classification methods may have on mapping output is often taken for granted by map users and neglected by map makers. In line with this, Tompson *et al*. [24] warn that careful range setting is vitally important to the credibility of the map and that crime analyst responsible for making KDE maps are sometimes not aware of this nor alert to the problems it can cause. Automatic choice of class breaks, called a ‘default option’, may not give a good map, and the issue about which thematic range to choose remains a serious problem [8].

4. Methodology and Data

The objective of this research was not to compare different mapping techniques for the sake of assessing their predictive accuracy. Instead only one technique (KDE) was applied, paying attention to only one of its parameter – the choice of classification scheme.

Hence, in each KDE computation, calculated cell values were divided into five thematic classes, using three the most common default classification method provided by ESRI ArcGIS® Spatial Analyst (hereafter ESRI ArcGIS) (*equal interval classification*, *quantile classification* and *natural breaks classification*)[3] and incremental multiples of the grid cells’ mean as recommended by Chainey *et al*. [4].

In the ArcGIS classification schemes, ‘hot’ values were determined by the top thematic class and in the case of the incremental mean approach, following Chainey *et al*. [4] recommendation, value above 3 multiples of the grid cells' mean (hereafter incremental mean 3+) was used to describe when a concentration reaches hotspot level.

In order to better assess the utility of incremental multiples of the grid cells’ mean approach, threshold value above 5 multiples of the grid cells' mean (hereafter incremental mean 5+) was used too.

In order to assess the influence of these classification methods on predictive accuracy of kernel density estimation hotspot maps, for each KDE output three the most common measures of predictive accuracy were calculated: HR, PAI and RRI.

However, prior to applying different classification schemes, the other three KDE parameters (the cell size, the bandwidth and kernel function) have to be determined. In this research the existing guidelines from crime mapping literature were followed, because those guidelines, with additional research in the field, should lead us towards establishing universal (‘gold’) standards on how to set/find appropriate parameters in particular setting (e.g., taking into account peculiarities of study area, crime type etc.,). Having in mind that ‘guidelines for parameter setting are not widely known among crime analysts, so they are inclined to use the GIS default options for KDE, unaware of the influence these settings may have on final output’ [24], while choosing appropriate KDE parameters, ESRI’s ArcGIS default methodology for parameter determination was also used in this research.

- Cell Size: In order to choose appropriate cell size two recommendations often stated in crime mapping literature were taken into account:
  - Taking the shortest side of the Minimum Bounding Rectangle (MBR) of the study area, and dividing this distance by 150, as suggested in [5].
  - Taking the shorter of the width or height of the output extent in the output spatial reference, divided by 250, which represents the default way of calculating cell size in the ESRI ArcGIS software.

Although both 150 and 250 are arbitrary values, they were used as starting point. Having in mind the small size of study area (the shortest side of study area’s MBR is 1974 m) these values produced very small cell sizes (13,16 m and 7,9 m), so the another arbitrary

[3] Having in mind that calculated KDE values forms skewed distribution, standard deviation classification was not considered as appropriate to be applied in this research

value was used – 100. This gave the value of 19,74 (rounded to 20m) which is more appropriate for given study area, but still very close to the values calculated following above mentioned recommendations. For the reason of consistency, the cell size of 20 m was used in every KDE calculation.

- *Bandwidth Size*: Crime mapping literature offers different recommendations on choosing appropriate bandwidth size [5, 7, 15, 19, 27]. Because of the fact that there is no enough research evidence which of these is more (or less) appropriate to follow in order to maximize KDE predictive accuracy, the most common of them, including default values determined by ESRI ArcGIS[4], were applied. This was the reason why four bandwidth sizes were calculated and used in this research. Calculation of the bandwidth sizes were performed using:
  1. Default methodology used in ESRI ArcGIS (version 10.2.1 and above). For ArcGIS 10.2.1, the default search radius (bandwidth) is calculated based on the spatial configuration and number of input points. It is computed using a spatial variant of Silverman's Rule of Thumb that is robust to spatial outliers[5]. Applying this methodology calculated bandwidth size was 428,6 m (hereafter the ArcGIS I bandwidth).
  2. Default methodology used in ESRI ArcGIS (versions previous to 10.2.1). Applying this methodology (the default search radius is calculated by taking the smaller of the width or height of the extent of the input, divided by 30), calculated bandwidth size was 164,58 m (hereafter the ArcGIS II bandwidth).
  3. Calculated value of average K-th nearest neighbor distance among crime incidents [27]. Having in mind the size of distribution, value of k=8 was chosen and bandwidth size of 271,54 m was calculated (hereafter the K bandwidth).
  4. Calculation of Moran's I for a series of increasing distances, and the distance at which intensity of clustering (z-score) peaks may be used as appropriate bandwidth size. Using ESRI ArcGIS Incremental Spatial Autocorrelation Tool (number of distance bands was set at 30, beginning distance at 200 m and distance increment at 75 m) reveals two peaks, the first at 350 m and the second (maximal) peak at 575 m. Taking into account the scale of analysis, the first peak (350 m) was selected (hereafter the M bandwidth).

[4]Users often use default parameter settings offered by used GIS software, so they must be part of any research which asses accuracy of particular crime mapping technique. Since ESRI's ArcGIS today is the most used GIS software in law enforcement community, its default parameter settings are used throughout this research.

[5] More in: ArcGIS Resources, ArcGIS Help file.

- *Kernel function*: In this research quartic kernel function was used. This is in accordance with research findings which, among different kernel functions used in popular GIS applications (normal, quartic, triangular, and uniform kernel functions) recommend using either quartic or triangular kernel function, as they, according to authors, were the methods that produced consistently high predictive accuracy scores, despite being slightly less precise [15].

The research was conducted on the robbery dataset committed in one of the Belgrade (Serbia) municipalities (Cukarica). The territorial coverage of this research referred only to the urban area of Cukarica (urban area of 8 km^2), which is a part of the Belgrade's urban area. In this area during the period of 2008-2010, 140 robberies were committed.

Using ESRI ArcGIS, kernel density was calculated on robbery dataset which was obtained from Belgrade Police Department. Having set cell size to 20 m, four different bandwidths were used (428,6 m, 164,58 m, 271,54 m and 350m), and for each KDE output, cell values were divided into five thematic classes, using four different classification methods. In the case of equal interval, quantile and natural breaks classification methods hotspot areas were determined by the top thematic class and in case of the fourth classification method (incremental mean approach) values above 3 and 5 multiples of the grid cells' mean were used. This produced 20 different KDE outputs (Figure 1).

Using robbery crime dataset for the years 2011–2013, predictive accuracy was assessed using three most common measures of predictive accuracy: HR, PAI and RRI. Each of them was calculated for each year respectively and comparisons between different classification methods and bandwidth sizes were made using average values for these years (Table 1).

5. Results and Discussion

Across all bandwidth sizes, the Hit Rate (HR) values were the highest within quantile classification method (Table 1). The lowest HR values were calculated using equal interval classification method, with the exception of the ArcGIS I bandwidths where incremental mean 5+ classification method produced the minimum HR value. The average HR values for ArcGIS I bandwidth varied between 0,193 to 0,570, for ArcGIS II bandwidth from 0,074 to 0,529, for K bandwidth from 0,106 to 0,626 and for M bandwidth from 0,106 to 0,570.

Equal interval classification produced the highest average HR using ArcGIS I bandwidth (0,213), the quantile classification produced the highest HR using K bandwidth (0,626), natural breaks classification using ArcGIS I bandwidth (0,420), incremental mean 3+ classification using ArcGIS I bandwidth and M bandwidth (0,453 and 0,455), and incremental 5+ classification using ArcGIS I bandwidth (0,119). As it can be seen all classification methods (except quantile classification) produced the highest HR values using the largest bandwidth size (ArcGIS I). Among different classification methods, the HR results were much higher for quantile classification method.

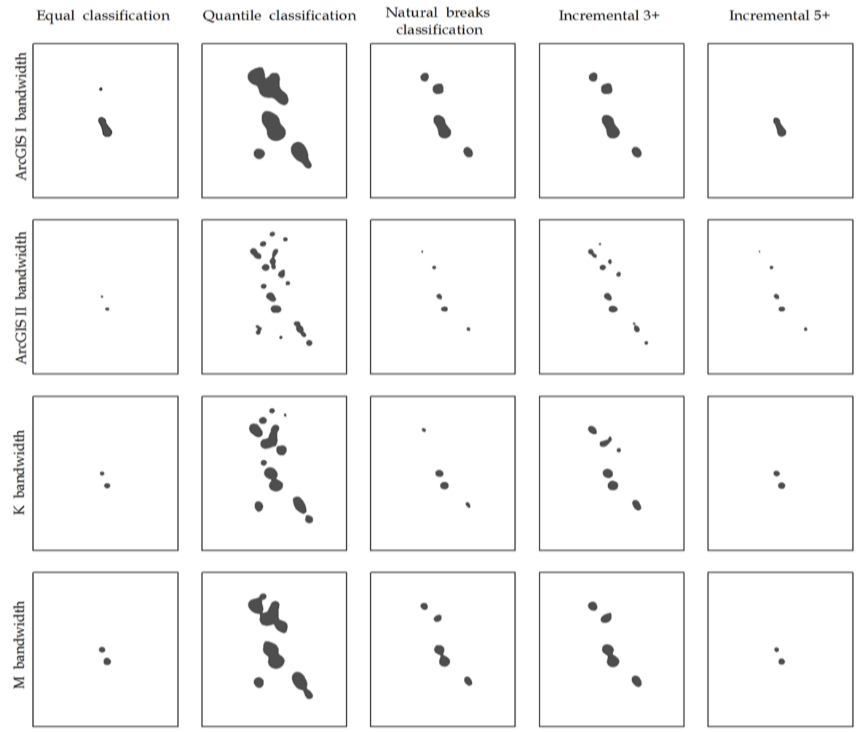

Figure 1. KDE outputs produced using different classification methods and bandwidths.

Table1. Calculated Hit Rate (HR), Predictive Accuracy Index (PAI) and Recapture Rate Index (RRI) values using different classification methods and KDE bandwidths for the years 2011, 2012 and 2013.

	HITRATE					**PAI**					**RRI**				
	EI	Q	NB	I3+	I5+	EI	Q	NB	I3+	I5+	EI	Q	NB	I3+	I5+
	ArcGIS I bandwidth (428,6 m)					ArcGIS I bandwidth (428,6 m)					ArcGIS I bandwidth (428,6 m)				
2011	0,138	0,586	0,448	0,480	0,140	6,468	2,372	6,493	6,580	7,511	0,660	0,750	0,870	0,900	0,690
2012	0,188	0,625	0,313	0,380	0,125	8,803	2,530	4,536	5,210	6,706	0,910	0,810	0,620	0,730	0,625
2013	0,313	0,500	0,500	0,500	0,313	14,671	2,024	7,246	6,850	16,793	1,510	0,650	0,990	0,960	1,563
average	0,213	**0,570**	0,420	0,453	*0,193*	9,981	*2,309*	6,092	6,213	**10,337**	**1,027**	*0,737*	0,827	0,863	0,959
	ArcGIS II bandwidth (164,58 m)					ArcGIS II bandwidth (164,58 m)					ArcGIS II bandwidth (164,58 m)				
2011	0,035	0,586	0,172	0,380	0,170	19,167	6,816	17,958	11,180	19,816	0,400	0,670	0,520	0,710	0,536
2012	0,125	0,500	0,125	0,310	0,125	69,444	5,814	13,021	9,120	14,617	1,460	0,580	0,380	0,590	0,389
2013	0,063	0,500	0,063	0,250	0,063	34,722	5,814	6,510	7,350	7,285	0,730	0,580	0,190	0,470	0,194
average	*0,074*	**0,529**	0,120	0,313	0,119	**41,111**	*6,148*	12,496	9,217	13,906	**0,863**	0,610	*0,363*	0,590	0,373
	K bandwidth (271,54 m)					K bandwidth (271,54 m)					K bandwidth (271,54 m)				
2011	0,069	0,690	0,172	0,380	0,069	13,800	4,421	10,141	7,760	8,103	0,430	0,810	0,610	0,710	0,371
2012	0,125	0,500	0,250	0,310	0,125	25,000	3,205	14,706	6,330	14,680	0,800	0,590	0,900	0,590	0,673
2013	0,125	0,688	0,188	0,310	0,125	25,000	4,407	11,029	6,330	14,680	0,800	0,820	0,670	0,590	0,673
average	*0,106*	**0,626**	0,203	0,333	*0,106*	**21,267**	*4,011*	11,959	6,807	12,487	0,677	**0,740**	0,727	0,630	*0,573*
	M bandwidth (350 m)					M bandwidth (350 m)					M bandwidth (350 m)				
2011	0,069	0,586	0,173	0,410	0,069	7,188	2,867	3,833	6,400	10,598	0,370	0,730	0,460	0,800	0,402
2012	0,125	0,625	0,250	0,310	0,125	13,021	3,058	5,556	4,840	19,199	0,670	0,790	0,670	0,620	0,729
2013	0,125	0,500	0,438	0,500	0,125	13,021	2,446	9,722	7,810	19,199	0,670	0,630	1,180	0,990	0,729
average	*0,106*	**0,570**	0,287	0,455	*0,106*	11,077	*2,790*	6,370	6,350	**16,332**	*0,570*	0,717	0,770	**0,803**	0,620

Legend: EI ─ equal interval, Q ─ quantile, NB ─ natural breaks, I3+ ─ incremental 3 +, I5+ ─ incremental 5 +

To summarize, the highest HR values were produced using the largest bandwidths and quantile classification method. Both of these parameters influence HR through the hotspots size. Namely, greater bandwidth size means greater smoothing which leads to larger hotspot sizes. If we take into account that greater hotspots sizes, means greater probability to "catch" more crimes, this bandwidth size's influence on HR is expected. This was in accordance with Hart&Zandbergen's[6] conclusion that "there is a corresponding increase in hit rate with an increase in kernel bandwidth" [14]. Similar "effect" could be noticed with quantile classification method

[6]These authors also concluded that the increase is not proportionate to the increase in the search radius. For example, doubling the bandwidth from 50m to 100m or from 100m to 200m does not result in a corresponding two-fold increase in hit rate.

which, comparing to others produced the largest hotspot sizes (Figure 2).

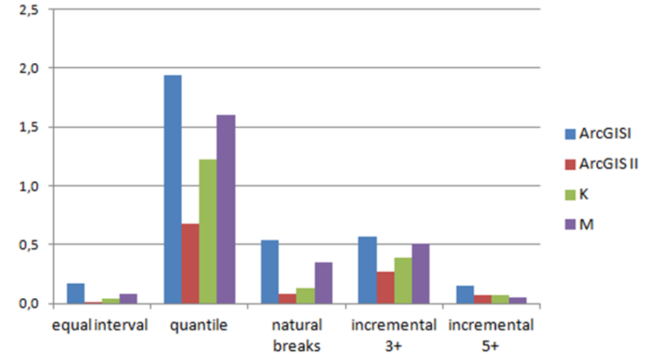

Figure 2. The total size of hotspots (in square kilometers) produced using different classification methods and bandwidth sizes.

Because of that, while interpreting HR, attention should be paid to the size of the hotspot area, since a larger hotspot size means greater probability of catching more crimes. This becomes obvious after the calculation of PAIs which takes the study area and the hot spot sizes into consideration. Across four different bandwidths the PAIs were the lowest for the largest ArcGIS I bandwidth (which confirms Chainey's conclusion that "spatial prediction ability of the KDE hotspot map is degrading as bandwidth size increases" [7]) and quantile classification method (the largest hotspot sizes and the highest HRs). This was expected having in mind accuracy vs. precision trade-off

Knowing the way the bandwidth size and classification method influence PAI, it was expected that the highest average PAI value (41,111) was calculated using the smallest bandwidth (ArcGIS II) and equal interval classification method (the smallest hot spot sizes).

Besides bandwidth size and chosen classification method, the predictive ability of KDE maps may be affected by hot spot threshold. After applying different classification methods, hot spots were defined using the highest "top" class. As kernel density values were not normally distributed, the top (hotspot) classes, depending on chosen classification method, took different data value ranges (Figure 3). The quantile classification's top (hotspot) class encompasses the largest portion of calculated data value range and produced the largest hotspots sizes[7]. At the other side, the top classes of equal interval classification method (which has predefined classes of 20% of data value range) and incremental mean 5+ classification method produced the smallest data value ranges (Figure 3).

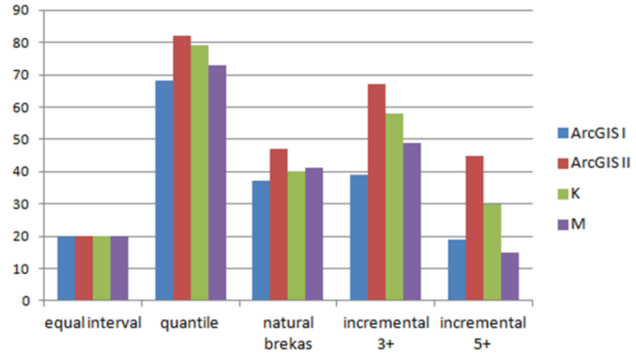

Figure 3. The top class portions of the total data value ranges produced using different classification schemes and bandwidth sizes.

In the case of larger data value ranges (lower hot spot thresholds) lower PAI values were recorded, and in the case of smaller data value ranges (higher hot spot threshold) higher PAI values were recorded. The size of top class data value range affects the size of hot spots, hence influencing HR (larger hot spot size – larger probability to catch more crimes – larger HR) and PAI (larger hot spot size – smaller ratio of events per square unit), so while assessing KDE predictive accuracy this must not be neglected. This is in accordance with Hart&Zandbergen's conclusion that "relatively lower thresholds correspond to relatively higher hit rates across crime types, using a low threshold also appears to correspond to the lowest predictive accuracy when measured in terms of PAI and RRI" [14].

For the largest bandwidths (ArcGIS I and M bandwidths), the highest average PAI was calculated using incremental mean 5+ classification method (the smallest data value range), while for the smallest bandwidths (K and ArcGIS II bandwidths) the highest average PAI was calculated using equal interval classification (the smallest data value range). According to these results, the highest PAI values were produced using classification methods which top (hot) spot classes have small data value range (equal interval and incremental 5+ classification methods) suggesting that the bandwidth size was not able to considerably affect these results (PAIs calculated using equal interval and incremental 5+ were the highest both for small and larger bandwidth sizes).

The highest average RRI value was calculated using equal interval classification method and ArcGIS I and ArcGIS II bandwidth sizes. In the case of equal interval classification and ArcGIS I bandwidth size RRI value for the year 2013 was above 1 indicating increasing concentration from 2008-2010 year. It could be noticed that equal interval classification method applied to ArcGIS II bandwidth size has the highest PAI and RRI value (Table 1).

[7]Quantile classification method typically does not produce intervals that are similar in size. KDE produces skewed distribution, where majority of cells have zero or very close to zero values. Hence lower classes will be filled with similar zero or very close to zero values, while higher classes will encompass the same number of cells with higher, dissimilar values. Hence, top (hotspot) class took larger data value range.

As it can be seen in Figure 2, different classification methods produced hotspots of different sizes, hence affecting HR and PAI values.

If someone (e.g., crime analyst) wants to pursue high HR he should use quantile classification method which generates the largest hotspot sizes across all bandwidths. At the other side, high PAI values may be achieved using equal interval and incremental 5+ classification methods which generated the smallest hotspots. At the same time, rising the hotspot threshold level (hence lowering top class data value range) could ‘inflate’ PAI and ‘blow out’ HR and vice versa. As a result, in terms of predictive accuracy, created hotspot maps may be the example of good hotspot identification work. But is this true from the point of view of someone whom this hotspot maps are intended for - the police decision maker?

Large hotspots (high HR) or too small hot spots (high PAI), from the point of view of police decision maker, who is expected to effectively deploy (often scarce) resources and to do something regarding crime at hotspot, may be of limited operational utility. Saying that hotspot must not be too small nor too large, means that while identifying hotspots, some “accuracy limits” must be set. Otherwise pursuing high HR or PAI may yield hotspots with little operational value. Having in mind that different classification methods produce hot spots of different sizes, this should be taken into account while choosing optimal classification method and threshold value.

Grid cell values of the kernel density output are not normally distributed, since the most cell values will have no values or very small values. At the other side a small number of cells will have high values comparing to the rest of the study region. While using KDE attention should be paid to the fact whether chosen data classification method considers data distribution. Because of the skewed distribution, breaking cells into groups that have equal range of values (equal interval classification) or equal number of features (quantile classification) may not be appropriate, especially if the top class is used for hotspot threshold. Unlike equal interval and quantile classification methods which do not consider how the data are distributed, natural breaks classification takes it into consideration.

If we take into account that some classification methods produce too large (quantile classification) or too small (equal interval and incremental mean 5+ classification methods) hotspots, which may not satisfy operational requirements of the hotspot policing, and the fact that equal interval and quantile classification methods do not consider the nature of (skewed) distribution, hence nor statistical requirements, natural breaks and incremental mean approach (above 3 multiples of the grid cells’ mean), may be used as better solution for data classification. Taking into account the size of produced hotspots and applied predictive metrics (calculated HR, PAI and RRI), incremental mean 3+ and natural breaks classification methods may be the most preferable solution (having in mind accuracy vs. precision trade-off). Incremental mean 3+ produced fairly larger hotspots which are not followed by intensive decrease in PAI value. If take into account RRI values, it can be noticed that among natural breaks and incremental mean 3+ classification methods, the highest average RRIs were associated with incremental mean 3+ (except for the K bandwidth).

Finally, while choosing appropriate classification method two additional things should be paid attention to:

1. The ability of map user with limited mapping and/or statistical knowledge to understand the map and the way data is classified (e.g., quantile map can be misleading if the map legend is ignored).
2. The ability to compare different maps if the same classification method is applied (e.g., maps that show different time periods).

Taking into account the fact that incremental mean approach is simpler to understand by the average police officer (crime map user) and the fact that natural breaks classification generate unique classification scheme which undermine map comparison, these criteria favors incremental mean approach.

If we agree that the choice of appropriate classification method must satisfy operational and statistical requirements discussed above, and the necessity that produced map has to be easily understandable for the intended map user (e.g., police officer) and comparable to other maps which depicts the same phenomena (e.g., widely used in hot spot temporal considerations, hot spot strategy evaluations etc.), and if we take into account HR, PAI and RRI values calculated on the hotspots produced using four classification methods applied in this paper, this research suggest using incremental mean 3+ classification method

6. Conclusions

Although there is no agreement that KDE technique is the best in the sense of its capability to accurately identify high crime areas, the KDE technique is currently widely used in police practice, and because of that there is a strong need for (additional) research that could provide guidance on the choice of its parameters in order to maximize its predictive accuracy. The ability of KDE method to predict future crimes depends on several factors, including the parameter settings. Among different parameters, the influence of data classification methods remained under researched.

The “wrong” classification method may lead toward false interpretation of crime pattern which could influence, as it is stated by Bowers et al., police officer(s) “not be somewhere/do something at a place where a crime happens, or be somewhere/do something at a place where a crime does not happen” [1]. Applying one of the default classification methods offered by GIS software becomes common in police practice. In absence of empirical or scientific guidance, efforts made in order to “customize” default classification method to the particular data set were usually guided by map maker prior experience and/or trial and error experimentation. Factors like the nature of mapped data, whom the map is intended for, along with operational benefits of the produced output (hotspot) and statistical justification of the applied classification method, were less likely to be considered.

This research suggest using incremental mean 3+ classification method because it is robust enough, it has acceptable accuracy vs. precision trade-off and produces hotspots that are more in accordance with operational and statistical requirements.

Researching the influence of different classification methods on predictive accuracy of KDE hotspot maps is not less important than those conducted in order to provide guidance for choosing other KDE parameters. In the future similar researches should be conducted, but in the way that overcome the limitations of this research. They should be conducted using much larger data sets, over longer time periods and comparing results among different crime types. This will bring us closer to what is considered as “ideal” data classification method for particular dataset. Recommendations regarding “ideal” data classification method, along with “ideal” bandwidth and cell size, which should be achieved through already conducted and future researches, should improve the predictive reliability of KDE, and confirm the benefits hotspot identification and crime mapping discipline have in everyday police practice.

Acknowledgment

This work was supported by the Ministry of Science and Technological Development of the Republic of Serbia, under Grant No. 179045.

References

[1] Bowers K., Johnson S., and Pease K., “Prospective Hot-Spotting: The Future of Crime Mapping?” *British Journal of Criminology*, vol. 44, no. 5, pp. 641-658, 2004.

[2] Braga A. and Weisburd D., *Policing Problem Places: Crime Hot Spots and Effective Prevention*, Oxford: University Press, 2010.

[3] Caplan J. and Kennedy L., *Risk Terrain Modeling: Crime Prediction and Risk Reduction*, University of California Press, 2016.

[4] Chainey S., Reid S., and Stuart N., *When is A Hotspot and Hotspot? A Procedure for Creating Statistically Robust Hotspot Maps of Crime*, CRC Press, 2002.

[5] Chainey S. and Ratcliffe J., *GIS and Crime Mapping*, John Wiley and Sons, 2005.

[6] Chainey S., Tompson L., and Uhlig S., “The Utility of Hotspot Mapping For Predicting Spatial Patterns of Crime,” *Security Journal*, vol. 21, pp. 4-28, 2008.

[7] Chainey S., “Examining the Influence of Cell Size and Bandwidth Size on Kernel Density Estimation Crime Hotspot Maps for Predicting Spatial Patterns of Crime,” *Bulletin of the Geographical Society of Liege*, vol. 60, pp. 7-19, 2013.

[8] Eck J., Chainey S., Cameron J., Leitner M., and Wilson R., *Mapping Crime: Understanding Hotspots*, National Institute of Justice, 2005.

[9] Eldrandaly K. and Naguib S., “A Knowledge-Based System for GIS Software Selection,” *The International Arab Journal of Information Technology*, vol. 10, no. 2, pp. 152-159, 2013.

[10] Fan S., The Spatial-Temporal Prediction of Various Crime Types in Houston, TX Based on Hot-Spot Techniques, Thesis, Louisiana State University, 2014.

[11] Ghosh A., Lagenbacher M., Duda J., and Klofas J., “The Geography of Crime in Rochester-Patterns over Time (2005-2011),” Working Paper, Center for Public Safety Initiative, 2012.

[12] Haberman C. and Ratcliffe J., “The Predictive Policing Challenges of Near Repeat Armed Street Robberies,” *Policing: A Journal of Policy and Practice*, vol. 6, no. 2, pp. 151-166, 2012.

[13] Harries K., *Mapping Crime: Principle and Practice*, Office of Justice Programs, 1999.

[14] Hart T. and Zandbergen P., “Effects of Data Quality on Predictive Hotspot Mapping,” Technical Report, Submitted to US Department of Justice, 2012.

[15] Hart T. and Zandbergen P., “Kernel Density Estimation and Hotspot Mapping: Examining the Influence of Interpolation Method, Grid Cell Size, And Bandwidth on Crime Forecasting,” *Policing: An International Journal of Police Strategies and Management*, vol. 37, no. 2, pp. 305-323, 2014.

[16] Johnson S., Bowers K., Birks D., and Pease K., “Predictive Mapping: Accuracy, Units of Analysis and the Environmental Backcloth,” *in Proceedings of Putting Crime in its Place: Units of Analysis in Geographical Criminology*, New York, pp. 171-198, 2009.

[17] Johnson S., Bernasco W., Bowers K., Elffers H., Ratcliffe J., Rengert G., and Townsley M., “Space-Time Patterns of Risk: A Cross National Assessment of Residential Burglary

Victimization," *Journal of Quantitative Criminology*, vol. 23, no. 3, pp. 201-219, 2007.
[18] Levine N., "The "Hottest" Part of A Hotspot: Comments on "The Utility of Hotspot Mapping for Predicting Spatial Patterns of Crime," *Security Journal*, vol. 21, no. 4, pp. 295-302, 2008.
[19] Levine N., *CrimeStat: A Spatial Statistics Program for the Analysis of Crime Incident Locations (v 4.0)*. Ned Levine and Associates, Houston, TX and the National Institute of Justice, 2013.
[20] Monmonier M., "Lying with Maps," *Statistical Science*, vol. 20, no. 3, pp. 215-222, 2005.
[21] Ratcliffe J. and Rengert G., "Near Repeat Patterns in Philadelphia Shootings," *Security Journal*, vol. pp. 21, 58-76, 2008.
[22] Sherman L., *Hot Spots of Crime and Criminal Careers of Places*, Criminal Justice Press, 1995.
[23] Smith S. and Bruce C., *CrimeStat III: User Workbook*, the National Institute of Justice, 2008.
[24] Tompson L., Partridge H., and Shepherd N., "Hot Routes: Developing a New Technique for the Spatial Analysis of Crime," *Crime Mapping: A Journal of Research and Practice*, vol. 1, no. 1, pp. 77-96, 2009.
[25] Van Patten I., McKeldin-Conor J., and Cox D., "A Microspatial Analysis of Robbery: Prospective Hot Spotting in A Small City," *Crime Mapping: A Journal of Research and Practice*, vol. 1, no. 1, pp. 7-32, 2009.
[26] Walter L., McInnis B., Price C., Smith S., and Hollywood J., *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*, RAND Corporation, 2013.
[27] Williamson D., McLafferty S., Goldsmith, V., Mollenkopf, J., and McGuire P., "A Better Method to Smooth Crime Incident Data," *ESRI Arc User Magazine*, pp. 1-5, 1999.
[28] Zandbergen P., "Geocoding Quality and Implications for Spatial Analysis," *Geography Compass*, vol. 3, no. 2, pp. 647-680, 2009.

Nenad Milić received the B.Sc. degree and M.Sc. degree in 2000. and 2004. Respectively from the Police Academy, Belgrade, Serbia, and Ph.D. degree in 2012. from the Faculty of Law, University of Kragujevac, Serbia. He is Associate Professor at the Academy for Criminalistic and Police Studies in Belgrade, Serbia, Department of Criminalistic. His research interest includes GIS, crime mapping and crime analysis.

Brankica Popović received the B.Sc. degree in electrical engineering in 1990. from the Faculty of Electrical Engineering, University of Pristina, M.Sc. degree in 1997. from the Faculty of Electrical Engineering, University of Belgrade and the Ph.D. degree in 2011 from the Faculty of Electronic Engineering, University of Nis, Serbia. She is Associate Professor at the Academy for Criminalistic and Police Studies in Belgrade, Serbia, Department of Informatics and Computer Sciences. Her research interest include image processing, biometric systems, information security and information systems with emphasize on information technology utilization in law enforcement agencies. She is a member of IEEE.

Saša Mijalković received the B.Sc. degree at Police Academy (2001), M.Sc. degree (2004) and the Ph.D. degree (2006) from the Faculty of Security of the University of Belgrade, Serbia. He is Full Professor at the Academy for Criminalistic and Police Studies in Belgrade, Department of Criminalistic. His research interest includes national security, organized crime, terrorism, intelligence and security affairs and services.

Darko Marinković received the B.Sc. degree in 1999. from the Police Academy, Belgrade, Serbia, M.Sc. degree and Ph.D. degree in 2003. and 2008. respectively from the Faculty of Law, University of Belgrade, Serbia. He is Full Professor at the Academy for Criminalistic and Police Studies in Belgrade, Serbia, Department of Criminalistic. His research interest include criminal investigation, organized crime, drug related crimes, special investigative techniques.