CrossMark

Identification of crash hotspots using kernel density estimation and kriging methods: a comparison

Lalita Thakali · Tae J. Kwon · Liping Fu

Abstract This paper presents a study aimed at comparing the outcome of two geostatistical-based approaches, namely kernel density estimation (KDE) and kriging, for identifying crash hotspots in a road network. Aiming at locating high-risk locations for potential intervention, hotspot identification is an integral component of any comprehensive road safety management programs. A case study was conducted with historical crash data collected between 2003 and 2007 in the Hennepin County of Minnesota, U.S. The two methods were evaluated on the basis of a prediction accuracy index (PAI) and a comparison in hotspot ranking. It was found that, based on the PAI measure, the kriging method outperformed the KDE method in its ability to detect hotspots, for all four tested groups of crash data with different times of day. Furthermore, the lists of hotspots identified by the two methods were found to be moderately different, indicating the importance of selecting the right geostatistical method for hotspot identification. Notwithstanding the fact that the comparison study presented herein is limited to one case study, the findings have shown the promising perspective of the kriging technique for road safety analysis.

L. Thakali (✉) · T. J. Kwon · L. Fu
Department of Civil & Environmental Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada
e-mail: lthakali@uwaterloo.ca

T. J. Kwon
e-mail: tkwon@uwaterloo.ca

L. Fu
e-mail: lfu@uwaterloo.ca

L. Fu
School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 610031, China

1 Introduction

Identification of hotspots is a systematic process of detecting road sections that suffer from an unacceptable high risk of crashes. It is a low-cost strategy in road safety management where a small group of road network locations is selected from a large population for further diagnosis of specific problems, selection of cost-effective countermeasures, and prioritization of treatment sites. These identified sites are often known by various terms in literature, such as hazardous locations, hotspots, black spots, priority investment locations, collision-prone locations, or dangerous sites.

There are various approaches that are aimed at identifying hotspots. One of the well-known approaches is using statistical crash models. This approach focuses on relating crashes as a function of potential variables such as road characteristics, traffic level, and weather factors using historical records [1–5] and subsequently uses these models to identify relatively high-risk sections. The other alternative approach is a geostatistical technique. This technique differs from the previous approach by considering the effects of unmeasured confounding variables through the concept of spatial autocorrelation between the crashes event over a geographical space. The focus of this study is to identify crash hotspots using the latter approach. Here, two distinctive geostatistical methods are evaluated and compared: one is the most widely used kernel density estimation (KDE) method and the other is the kriging method. The paper is arranged as follows. The following section provides a review of the existing literature on

Springer

hotspot identification. It is followed by a description of the study methodology with a brief background of the KDE and kriging methods. Then, detailed findings and comparisons are presented in the Results and Discussions section. Lastly, conclusions are made in the last section.

2 Literature review

Hotspots, which are defined as relatively high-risk locations, are commonly identified on the basis of some specific selection criteria. Many different methodologies and criteria have been developed for improving the accuracy of the hotspot identification process, thus the cost-effectiveness of a safety improvement program [6]. One of the most commonly used selection criteria is defined by the expected collision frequencies at the sites of interest. This particular criterion emphasizes on maximizing the system-wide benefits of safety intervention targeted to the hotspots, whereas another commonly implemented criterion is considering the expected collision rate (i.e., expected collision frequency normalized by traffic exposure) which emphasizes on individual road user's equity perspective [7].

The expected collision frequency at a site is commonly estimated using a collision model-based approach, in which collision frequency is statistically modeled as a function of some relevant features such as road characteristics, traffic exposure, and weather factors [1–5]. Roads are normally divided into homogenous sections of equal length and intersections as spatial analysis units. Various count models, with negative binomial (NB) being the most popular, are used to estimate the expected number of crashes over the road network in a study area, and the estimates are subsequently compared with a pre-specified threshold value for determining if a site belongs to a hotspot. Note that the NB models are normally used in empirical Bayes (EB) framework to better capture the local experience of safety levels [1, 6]. One of the most critical parts of this modeling approach is the assumption of a probability distribution for crash count and the functional specification of the model parameters. If these components are incorrectly specified, applying such count models could lead to incorrect hotspots. In addition, this approach is data intensive and requires significant effort in collecting and processing the related data and calibrating the corresponding models [8].

The expected crash frequency could also be estimated using a geostatistical technique by considering the effects of unmeasured confounding variables through the concept of spatial autocorrelation between the crash events over a geographical space [9–13]. KDE is an example which has been used in road safety to study the spatial pattern of crash and identify the hotspots [8–12]. Similarly, there are other geostatisical methods such as clustering methods that

evaluate relative risk based on their degree of association with its surroundings. Examples of these methods used in road safety studies are K-mean clustering [14, 15], nearest neighborhood hierarchical (NNH) clustering [16–18], Moran's I Index, and Getis-Ord Gi statistics [19–21].

Anderson et al. [11] applied KDE method in the City of Afyonkarahisar, Turkey. In this study, the authors were able to detect highly crash risk sections which were highly concentrated in road intersections. Similarly, Keskin et al. [13] and Blazquez and Celis [12] used KDE and Moran's Index method to observe temporal variation of hotspots across the road network. Khan et al. [21] used Getis-Ord Gi statistics to explore the spatial pattern of weather-related crashes, specifically crashes related to rain, fog, and snow conditions. A special pattern was revealed for each category of weather conditions which further suggested the need of prioritizing the treatments based on different weather conditions and locations. Pulgurtha et al. [9], Pulugurtha and Vanapalli [22], and Ha and Thill [23] employed KDE method to investigate the spatial variation of pedestrian crashes and hazardous bus stops. These studies have shown the potential to effectively and economically address pedestrian and passenger safety issues. Another study by Levine [17] and Kundakci and Tuydes-Yaman [18] used NNH clustering method to detect crash hotspots across the road network.

A noticeable difference in aforementioned geostatistical methods is how spatial correlations are considered. For example, in the KDE method, a symmetrical kernel function, which is a function of bandwidth, is placed on each crash point generating a smooth intensity surface. Then, for a given point of interest, the crash intensity is a summation of the entire overlapping surface due to the crashes. In contrary, in the clustering technique such as NNH, a threshold value, which determines the extent of clustering in the neighborhood, is pre-specified. If the distance between crash data point pairs is smaller than the threshold value, then these crashes are grouped into the same cluster. Additional criteria such as minimum number of points to be in a cluster can also be specified. This variation in allocating different weights to the crashes occurring in its neighborhood (e.g., KDE method) or simply grouping crashes into certain clusters clearly indicates that these techniques are likely to have different results in terms of size, shape, and location of hotspots. One of the attractive parts of the KDE method as compared to other variants of clustering methods is that it takes into consideration of spatial autocorrelation of crashes (see Sect. 4.1 for more detailed explanations). Moreover, this method is simple and easy to implement. This could be one of the reasons that KDE method is being widely used in road safety.

In the past efforts on geostatistical based methods, another popular technique called kriging has been rarely

⚫ Springer

J. Mod. Transport. (2015) 23(2):93–106

explored in road safety analysis. As one of the most advanced interpolation methods, kriging has been utilized widely across many different fields of studies in necessity of spatial prediction. With little prior information, this technique is able to provide a best linear unbiased estimator (BLUE) for variables that have tendency to vary over space [24, 25].

3 Study area and data description

The region of interest for this study is Hennepin County, Minnesota, which encloses the City of Minneapolis, the 14th largest metropolitan area in the United States Census Bureau [26]. The county has a dense road network with high crash potentials, making it an ideal location for the intended study. The study is based on historical crash data from 2003 to 2007 occurring in major highways, as depicted in Fig. 1. These crash data were originally collected by Minnesota Police Department, and maintained and archived by the Department of Transportation of Minnesota (Mn/DOT). The crashes in the dataset were already geocoded and included some important information such as severity of crashes (i.e., fatality, injury (three different categories) and property damage only), and weather conditions at the time of crashes. Figure 2 shows that more than two-third of crashes occurred in clear weather conditions. Similarly, more than two-third of crashes were property damage only.

The time-of-day that each crash occurred is also known from the dataset, which gives an opportunity to explore temporal trend of hotspots patterns across the highway network. Figure 3 shows total number of crashes that occurred within 5 years of time period based on different times of day. Relatively, the morning crashes are concentrated from a time period of 7–9 AM and evening crashes from 4 to 6 PM. Therefore, we categorized crashes as morning peak hours (7–9 AM), evening peak hours (4–6 PM), and rest as off peak hours. A total of 38,748 crashes were recorded for 5 years, where 5,331 crashes occurred in morning peak, 7,712 in evening peak, and 25,705 in off peak hour. From Fig. 4, it is clear that the peak hours have higher rate of crashes (i.e., per year per hour) than the off peak, which could be due to higher traffic exposures.

4 Methodology

Figure 5 presents an outline of methodology for the intended study. As mentioned previously, two geostatistical methods, KDE and kriging methods, were used to estimate crash intensity over the whole region. A brief description of each method is presented in Sects. 4.1 and 4.2. Following the crash estimation, a buffer of 400 m on either side of the
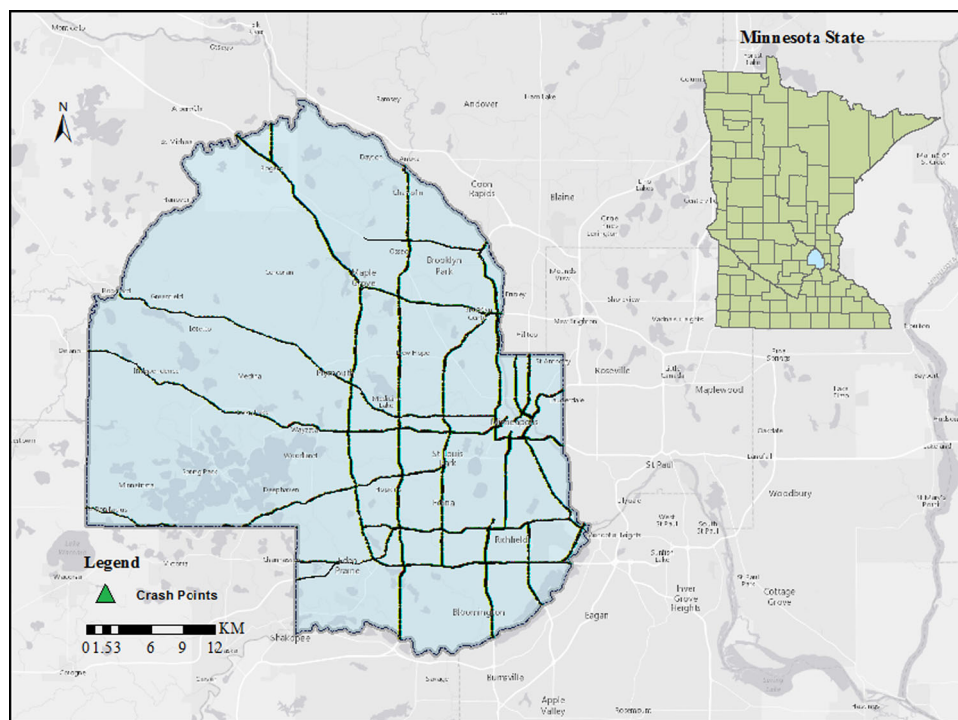

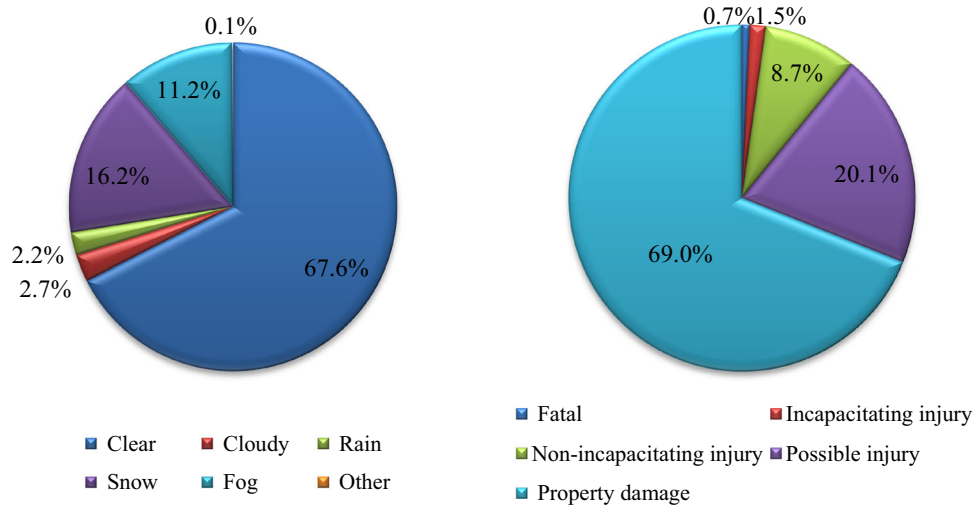**Fig. 1** Study area—Hennepin County, Minnesota

Fig. 2 Percentage of crash occurrences based on weather conditions and crash severities
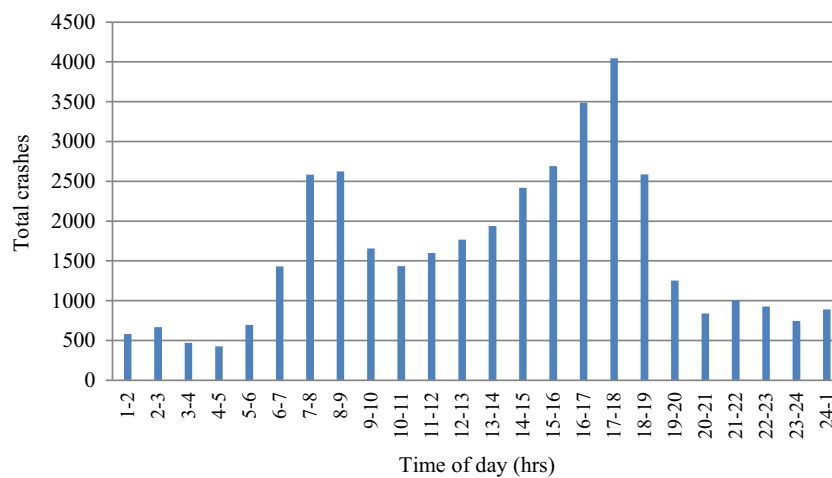

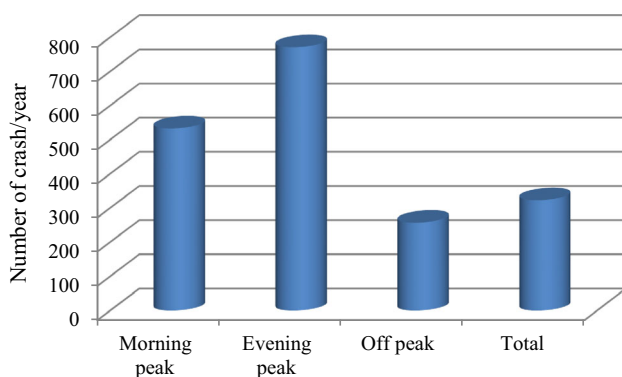Fig. 3 Total numbers of crashes (2003–2007) based on the time-of-day

Fig. 4 Average annual crash based on the time-of-day

highways was used to demarcate the estimated outputs from the two methods. A primary intent of choosing a buffer size of 400 m was to match a predefined spatial analysis unit, which was used to aggregate and produce the resulting outputs (i.e., estimates of crashes) from the two methods considered in this study. In addition, other areas that do not enclose highway segments are less likely (or unlikely) to have any record of crashes; therefore excluding those areas from the analysis was deemed inevitable by considering only the buffered areas. In a real-world application, use of 400 m can be considered reasonable for carrying safety treatments and providing sufficient precision for identifying actual locations of crashes. A smaller cell may be more prone to the problem of producing inaccurate crash statistics, while a larger cell may likely exhibit a loss of detailed information. Most importantly, as one of our objectives is to compare the performance of the two methods, a selection of the equal-sized cell was essential to enforce a fare comparison. A further explanation on selecting a cell size is further discussed in Sect. 4.1.
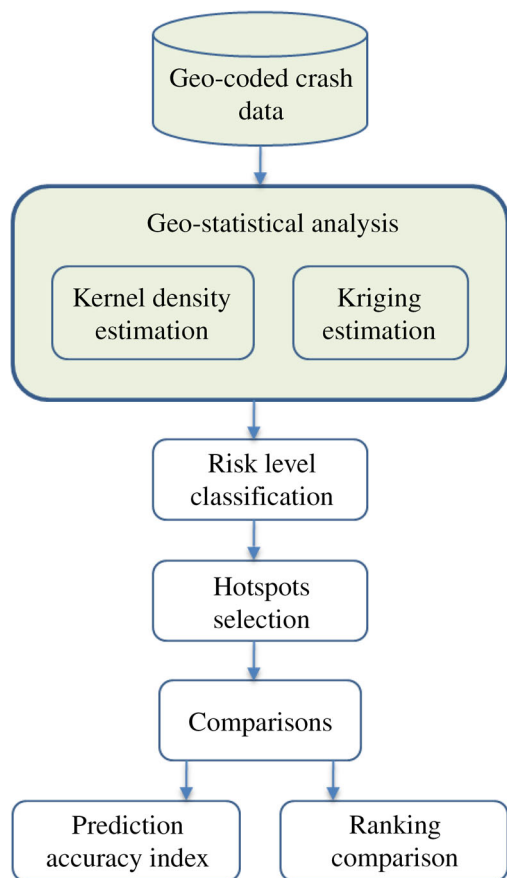
Springer

J. Mod. Transport. (2015) 23(2):93–106

Fig. 5 Flowchart of the comparison study

Estimated results obtained from each method represent a quantitative measure of a risk level reflecting the magnitude of the potential to crash occurrence. Thus, a road section with a larger value indicates that there is a higher chance of crashes than that with a lower value. A risk level along the highways was classified into 10 different levels using a quantile classification method. In this method, the entire set of estimated grid cells (ordered in respect of estimated values) was divided into ten groups with each group having an equal number of cells. Then, the top-most level (i.e., level 10) representing the highest risk highway sections was selected as hotspots. Finally, the selected two estimation methods (i.e., KDE and kriging methods) were compared using prediction accuracy index and ranking process. Details specific to the proposed methods are explained in the following sections.

4.1 Method I: Kernel density

The KDE, a non-parametric approach, is one of the most common used and well-established spatial techniques used to estimate the crash intensity for hotspot identification [9–13]. In this method, a circular search area defined by a
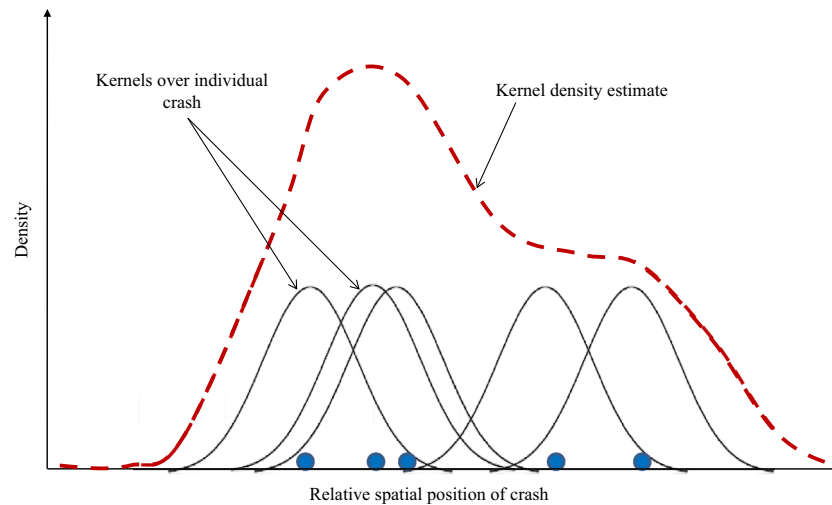
kernel function is placed over each crash (discrete points) resulting in individual smooth and continuous crash density surface (see Eq. (1) and Fig. 6 for 2D visualization). Then, a grid of cells is overlaid over the study area. For a given cell, density is estimated by summing the overlapping density surface resulted from each crash point. This procedure is repeated for all reference grid cells. Note that kernel functions are symmetrical mathematical functions.

$$f(x,y) = \sum_{i=1}^{n} \frac{1}{n \times 2 \times \pi h^2} \times W_i \times K\left(\frac{d_i}{h}\right), \tag{1}$$
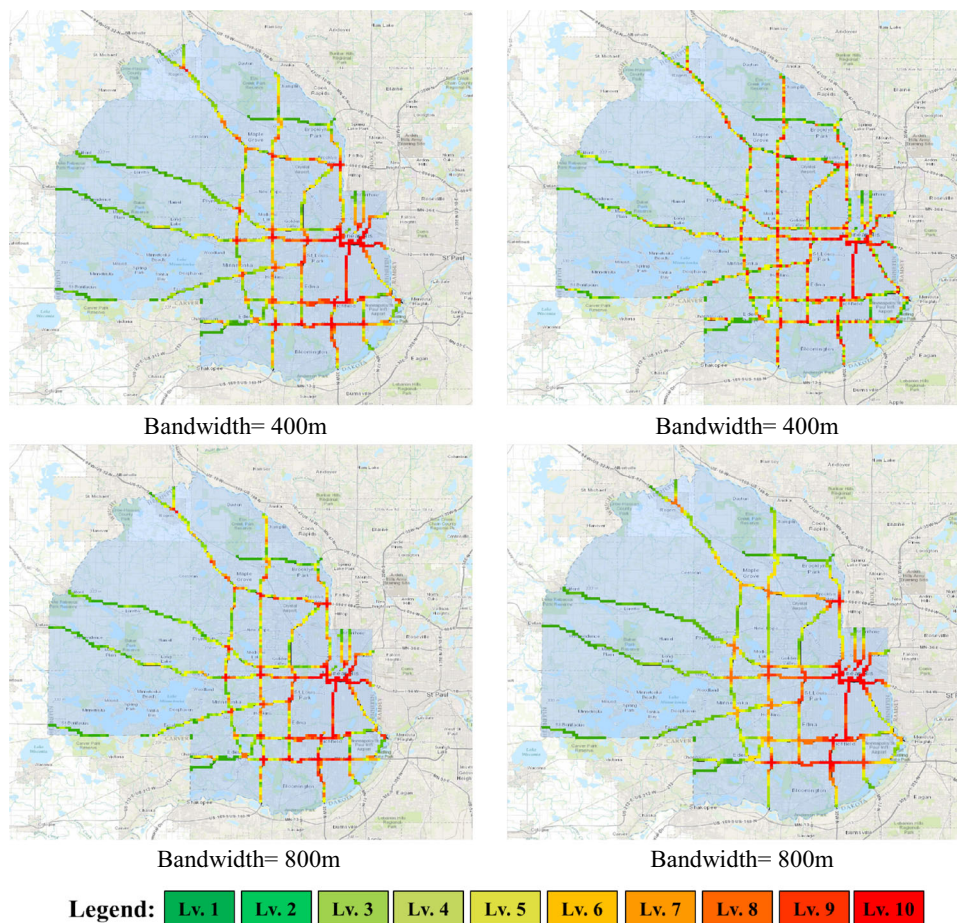
where $f(x,y)$ is the density estimate at the location (x,y); n is the number of observations; h is the bandwidth; K is the kernel function and d_i is the distance between the location (x, y) and the ith observation; and W_i is the intensity of the observation. For the crash count, W_i is unit, whereas this may vary when we consider different weights for different severities of crashes.

There is a wide choice of kernel functions such as normal, uniform, quartic, epanichnikov, and triagular. Among them, the most popular is normal [16, 17] used in CrimeStat and quartic functions [20] used in ArcGIS. According to Silverman [27], the selection of kernel function is less sensitive to the outcomes. In our study, we initially considered normal and quartic kernels using bandwidth of 400 and 800 m to estimate the density for all crash cases. As shown in Fig. 7, a general pattern of density estimation represented by the color-coded map appears to be very much similar. For example, if we look at the highest risk zones in red, they appear to be very similar. With this supportive information, we choose to consider normal kernel for the rest of the KDE estimations. Note that CrimeStat and Arcmap were used as the GIS platform for the analysis.

The two main parameters which affect the KDE are bandwidth and cell size. The output of KDE is presented in a raster format consisting of a grid of cells. Intuitively, the size of cell has to be reasonable to represent crash cluster occurring in reality. The selection of size is also a trade-off between the computation time, sample size, and the information to retain. Having larger grid cell size saves the processing time; however, the information is likely to be averaged in a larger area, thus resulting in loss of information. Meanwhile, too small grid cell size increases the computation time. Also, a lower level of granularity may not be necessary from the aspect of designing a safety program. Considering safety treatments for a reasonable length of section and keeping some space for the potential inaccuracy in geocoding of the crash location, 400 m of grid cell was used. Sizes could vary from one study to another (e.g., Anderson [11] used 100 m; Erdogan et al. [10] used 500 m; Blazquez and Celis [12] used 100 m).

J. Mod. Transport. (2015) 23(2):93–106

Springer

Fig. 6 Kernel density estimation method

Bandwidth= 400m

Bandwidth= 400m

Bandwidth= 800m

Bandwidth= 800m

Legend: Lv. 1 | Lv. 2 | Lv. 3 | Lv. 4 | Lv. 5 | Lv. 6 | Lv. 7 | Lv. 8 | Lv. 9 | Lv. 10

Fig. 7 KDE estimation using two different kernels (*left* Quartic kernel and *right* Gaussian kernel function)

Another important parameter in the KDE method is the selection of bandwidth which determines the extent of search area. Depending on the type of kernel estimate used, this interval has a slightly different meaning. For the normal kernel function, the bandwidth is the standard deviation of the normal distribution. For the uniform, quartic, triangular kernels, the bandwidth is the radius of the search area to be interpolated. The choice of bandwidth is quite

🍥 Springer

J. Mod. Transport. (2015) 23(2):93–106

subjective [11, 28, 29]. Typically, a narrower bandwidth interval will lead to a finer mesh density estimate with all the peaks and valleys detected, whereas a larger bandwidth interval will lead to a smoother distribution and, therefore, detect less variability between areas. While smaller bandwidths show greater differentiation among areas, one has to keep in mind the statistical precision of the estimate. Brimicombe [30] suggested that the bandwidth to be 6, 9, or 12 times the median of nearest neighbor distance.[1] In general, it is a good idea to experiment with different fixed intervals to see which results make the most sense [28]. Previous researchers have used values ranging from 20 to 1,000 m (e.g., Xie and Ya [22] used 20, 100, 250, and 500 m; Ha and Thil [23] used 400 and 800 m; Erdogan et al. [10] used 500 m; Keskin et al. [13] used 200 m; Blazquez and Celis [13] used 1,000 m). In our study, we considered two different bandwidth values, i.e., 400 and 800 m (equal and double the cell size). It is reasonable to consider that the correlation of crashes within a short length of 200 m on either side exists (i.e., 200 m on either sides of road section means total section length of 400 m). Moreover, 800 m was used to study the sensitivity bandwidth in the hotspots pattern.

4.2 Method II: Kriging

Kriging is a generic term coined by geostatisticians for a family of generalized least squares regression algorithms in recognition of the pioneering work of a mining engineer, Danie Krige [31]. The main idea behind kriging is that the predicted outputs are weighted average of sample data, and the weights are determined in such a way that they are unique to each predicted point and a function of the separation distance (lag) between the observed location and the location to be predicted. In other words, kriging provides estimates at unknown locations based on a set of available observations by characterizing and quantifying spatial variability of the area of interest. Let x and x_i be location vectors for estimation point and a set of observations at known locations, respectively, with $i = 1,… n$. In this study, x indicates a single point/location where a number of crashes likely to occur is estimated using nearby observations, x_i.

Based on n number of available crash frequencies, we are interested in estimating a number of crashes at any given location, denoted by $\hat{Z}(x)$. The expression of a general kriging model is as follows [32]:

$$\hat{Z}(x) = m(x) + \sum_{i=1}^{n} \lambda_i [Z(x_i) - m(x_i)], \qquad (2)$$

where $m(x)$ and $m(x_i)$ are expected values of the random variables $Z(x)$ and $Z(x_i)$, and λ_i is a kriging weight assigned to datum $Z(x_i)$ for estimation of a crash frequency at any location x. The random field, $Z(x)$, can be decomposed into two components namely residual component $R(x)$ and a trend component $m(x)$, and expressed as $Z(x) = R(x) + m(x)$. Each of three main variants of kriging namely simple kriging (SK), ordinary kriging (OK), and universal kriging (UK) can be distinguished according to the model considered for the trend component, $m(x)$.

The most widely used kriging approach, OK, assumes the constant mean over each local neighboring area, whereas SK assumes a constant mean over the entire study area, a characteristic that often limits the wide application. UK is a hybrid method which is based on point observations and regression of the target variable on spatially exhaustive auxiliary information [25]. In our analysis, OK was used as it is relatively simple yet powerful and less data intensive.

As mentioned previously, a fundamental assumption for geostatistical methods is the existence of spatial autocorrelation. The investigation of autocorrelation is essential in most geostatistical analyses that are done by modeling the spatial dissimilarities (semivariogram) based on the available sampling of the attribute of interest [24]. A most commonly adopted tool for capturing the spatial data that exhibit weak stationarity is semivariogram given in Eq. (3).

$$\hat{\gamma}(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} [Z(x_i + h) - Z(x_i)]^2, \qquad (3)$$

where $\hat{\gamma}(h)$ is the sample semivariogram, $z(x_i)$ is a crash frequency measurement taken at location x_i, and $n(h)$ is the number of pairs of available crash frequency observations separated by the lag distance h. Typically, a mathematical model is utilized to describe the sample semivariances, and a few examples of those are exponential, Gaussian, and spherical models.

Esri's ArcGIS 10.2 comes equipped with a geostatistical analyst package that offers a user-friendly kriging interpolation tool. OK is utilized to obtain the interpolated surface for all data with different temporal units at an aggregation level of 400 m^2. The extensive amount of heuristic trial and error is carried out to ensure that the semivariogram model and the parameters selected produce unbiased results (i.e., the mean prediction error should be close to 0, while the RMSE should be close to 1). There are several parameters that need to be carefully determined when constructing a semivariogram model and some are *sill*, *range*, and *nugget*. *Sill* represents the level of the plateau (if it exists), while *range* represents the distance

---

[1] Nearest neighbor distance is the distance from each point of event to its nearest neighbor.

J. Mod. Transport. (2015) 23(2):93–106

🕿 Springer

where the semivariogram reaches the *sill*, also commonly interpreted as degree of spatial correlation. *Nugget* represents that there is to account micro-scale variation and measurement errors or any spatial variability that exists at a distance smaller than the shortest distance of two measurements. For more information on how to build a good semivariogram model, readers are advised to refer to a comprehensive work done by Olea [33].

4.3 Hotspots selection criteria

After estimating the number of crashes over the grid cells, the outputs are presented in a color coded map. In both the methods, the estimation takes place over the entire study region. Those areas without the road network are likely to have mostly zero values except regions which are very close to the network. Meanwhile, those areas outside of networks are not of an interest. Such areas are discarded by extracting the results lying only within a buffer of 400 m on each side of highways. This particular value was selected as to make sure that all the grid cells of 400 by 400 m in the vicinity of highways are included. Note that only the major roads under the jurisdiction of Minnesota Department of Transportation in Hennepin County have been considered in this study.

The next step is selecting a set of high-risk zones (i.e., the hotspots). There is no universal rule or threshold values to benchmark for what should be the hotspots. It is an arbitrary selection of a cutting off value that screens relatively higher risk areas over the given study area. An example of this value could be considering an overall average of the estimated output [6]. When the estimated value for a given location is higher than this threshold value, then it is considered as hotspot. However, in a real world, this could be decided based on budget availability. Another alternative method, which is used in this study, is using quantile method where we classify the estimated values in different classes. For this, we pick a certain number of classes to be created, and then the data are distributed equally between the classes resulting into equal numbers of grid cells in each class [9, 34]. Note that in both the methods we use the same grid cell size (i.e., 400 by 400 m^2), thus controlling the numbers of total cells for the comparison. Each class represents the order of severity based on crash risk level. We label these categories as risk level 1, risk level 2, and so on. "Hotspot" is then determined by the top thematic risk level, i.e., level 10. As we are comparing the performance of KDE and kriging method with different output units, this approach makes a fair comparison from the perspective of consistency in hotspot coverage.

4.4 Comparisons

Prediction accuracy index (PAI) developed by Chainey et al. [34] was used as a performance measure to compare the performance of the two proposed methods (see Eq. (4)). This was initially developed in crime mapping context [34, 35, 37] and has been used in road safety as well [18]. Here, we have made a slight modification in the denominator using length of road segment instead of using its area. This is reasonable as highway network is better represented by a linear 1-D feature rather than a 2-D feature. Meanwhile, we also calculated the PAI in terms of area:

$$\text{PAI} = \frac{\frac{n}{N} \times 100}{\frac{m}{M} \times 100}, \tag{4}$$

where n is the number of crash in hotspots, N is the total number of crashes, m is the length of highway section in hotspots or area covered, and M is the total length of highway section or total area covered.

As seen in Eq. (4), PAI is a ratio of percentage of crashes occurring within the identified hotspots (say A) to the percentage of area covered by it (say B). Intuitively, the higher the PAI value, the better the performance. Note that one of the reasons for normalizing "A" by "B" is that the higher "A" value may not always necessarily indicate better ability in identifying risk zones without the normalization. For example, say we identify whole region as hotspots ("B" is 100), which means "A" is 100, then PAI would have been 100. However, with the normalization by "B," PAI index becomes 1, which is reasonable. Moreover, PAI index measures an ability to locate high number of potential crashes in a small area. A convincing example could be, say we have 25 % of crash occurring in hotspots which represents 50 % of total area, similarly 25 % crash in 80 % of area, then PAI values will be 0.5 and 0.31, respectively. Scientifically, we would choose the method that results the first case (i.e., PAI 0.5) as road agencies can allocate resources effectively by mobilizing them on a smaller area in treating high crash potentials.

In addition, another measure, which aims to compare the physical locations of hotspots delineated by KDE and kriging, was used to see if their outcomes are similar or different, and investigate if one approach can be used for another. For this comparison, only those hotspot locations commonly identified by both the methods were extracted, and their matching rate was computed with respect to the total area of hotspots.

5 Results and discussions

Two different geospatial techniques, the KDE and kriging methods, were employed for the hotspot analysis. Figure 8
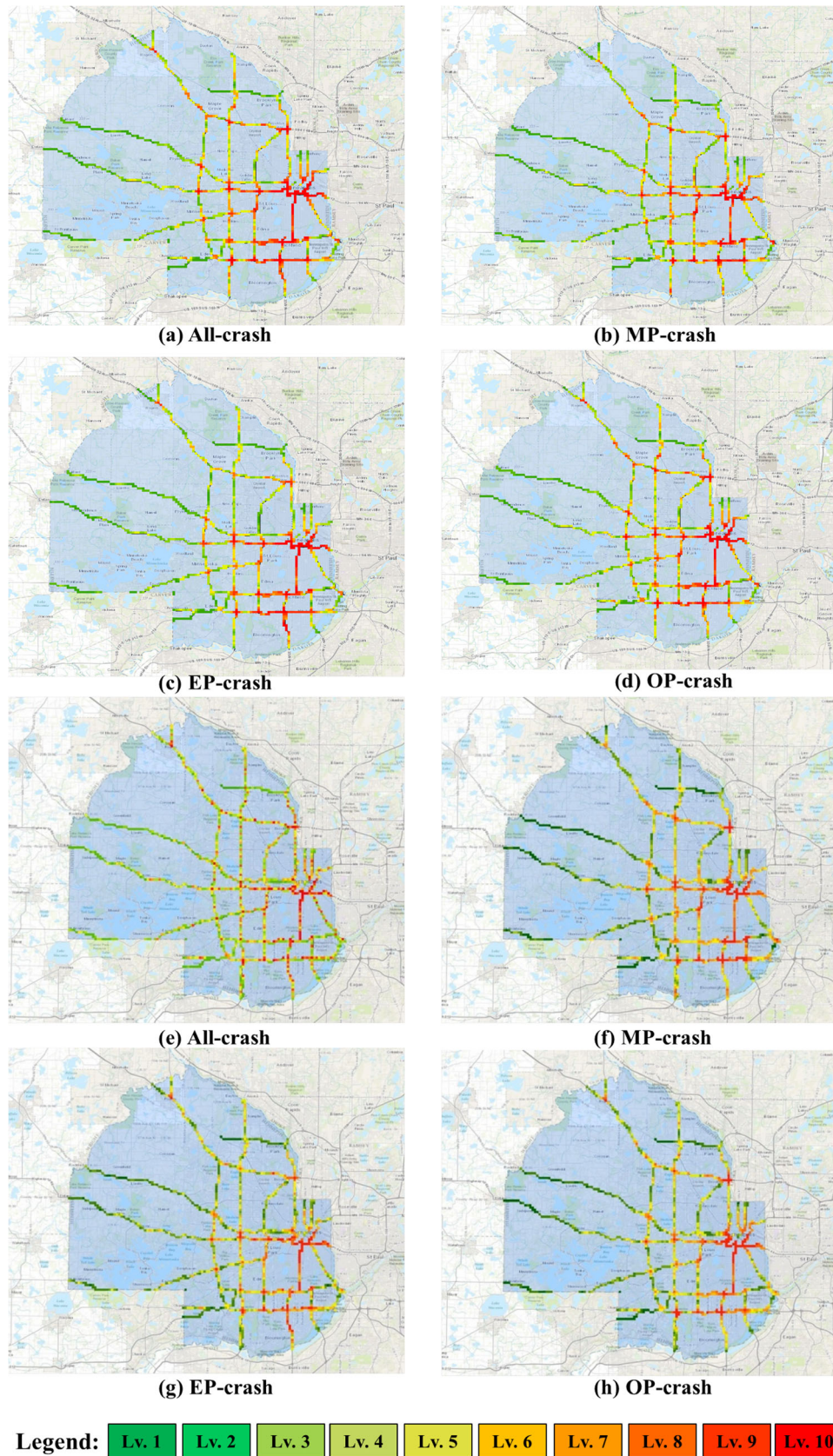
Fig. 8 KDE estimation results for bandwidth 400 m: (**a**)–(**d**); kriging estimation results: (**e**)–(**f**) (where *MP* morning peak, *EP* evening peak, *OP* off-peak)

illustrates the estimated results on a temporal basis using the two methods. For KDE method, two different bandwidths (i.e., 400 and 800 m) were considered to evaluate its sensitivity in the estimation process but only the results for 400 m are presented here due to the limited space. One of the reasons for using 400 m as a minimum spatial unit was to sufficiently cover the size of the predefined grid cell. In other words, a selection of bandwidth less than the size of the grid cell would be worthless for KDE method because its output would eventually be averaged over each grid. Note that in both methods grid cell size of 400 by 400 m was used. The results presented are categorized into ten different levels arranged in an increasing order of risk level, each of which represents 10 % coverage of the total buffered area. This classification in a color-coded map provides a clear visualization of where the crash-prone areas are, for example, increase in the degree of redness indicates higher risk sections. Such crash risk mappings could be a value to engineers and planners in road agencies in making a good decision for planning road safety budget. In general, both the methods showed that the high crash-prone zones are concentrated in the vicinity of Minneapolis city. This is intuitive as the higher level of traffic interactions generates more safety problems. As the highways are extending outward from the core urban areas, the risk level is decreasing. As seen in the figures, this macro-level of visualizing safety risks demonstrates little difference between the methods and temporal units. Further close investigation and comparisons are made by selecting a set of hotspots.

Criteria in selecting a set of hotspots may vary across studies, as there is no any universal rule of selection. Whatever the method is, the main controlling idea is to select a set of sections with higher safety risk. In our study, we used a simple quantile method in which the top risk level from the previously mentioned classes of risk levels was selected. We could also use a threshold cutoff value approach where estimated values in each location are compared with a critical value, and the locations exceeding the critical value are screened as hotspots. In such cases, cutoff values could be determined based on statistic of estimated crash over the area, such as mean and standard deviation.

Figure 9 presents selected hotspots (i.e., represented by red rectangles) locations using KDE and kriging methods. As observed, the spatial locations of hotspots identified by these two different methods are not identical, and it is important to make a decision of which method performance is better based on their performance evaluation. One of the common approaches of this is done by comparing the actual values against estimated results. However, unlike in classical statistical modeling method (e.g., using NB Models) which commonly uses this approach, it is not straightforward with geostatistical techniques. For example, in KDE method, we estimate the density of crashes. As a result, it is

not convenient to evaluate its performance by comparing output results (density) against its corresponding actual (count) values. Most importantly, as we are comparing two methods, a common measure is needed. This was addressed by adopting a performance measure, i.e., PAI, which was initially proposed by Chainey et al. [34].

Table 1 presents a comparison between these two methods in terms of PAI using both the length of highway section and buffer area coverage. A slight variation in hotspot location is observed among different times of day, suggesting that the hotspots may vary by the times of day. Comparatively, most of the hotspots are located around intersections and interchanges in both the methods; however, the hotspots from the kriging method are a little spread out. In all the cases, kriging method has higher PAI values compared to the KDE method. As explained in Sect. 4.4, higher PAI value indicates better ability of a method to locate high potential crash in a small area, which practically helps road agencies to efficiently mobilize limited resource. With this evidence of PAI values, kriging method is better performing compared to KDE. Meanwhile, this method of identifying hotspots could be concluded better in future by comparing with the statistical modeling approach. A similar hotspot pattern was observed with both the bandwidths in KDE method as presented in Table 1 (only output map for 400 m is reported). The results obtained from both the bandwidths were comparable, showing less sensitivity in the selected values as shown in Fig. 10.

In addition, another bandwidth (i.e., 800 m) was tested to check the sensitivity of different bandwidths but only a small difference was observed in their performance. For example, as shown in Table 1, the PAI (length) value for all crash case for 400 m bandwidth, the PAI index was found to be 2.75, while a comparative value of 2.72 was found for 800 m bandwidth. The magnitude of their difference is very marginal (i.e., 1.09 %). A similar trend was also observed when compared with other times of day as presented in Fig. 10. From this, it can be asserted that the PAI index seems to be higher when a smaller bandwidth is implemented irrespective of time-of-day. This also shows that a further analysis testing the sensitivity of different bandwidth sizes may not be necessary for comparison with the kriging method.

Note that the total areas of the selected hotspots are not exactly the same (see Table 1) but this small discrepancy is caused during GIS data processing. However, this minor issue pertaining to an insignificant discrepancy in the total area does not affect the resulting PAI indexes as they were calculated using the normalized numerical figures (refer to Eq. (4)).

As outlined previously, the hotspots identified by the two methods are compared by matching their physical

Ⓢ Springer

J. Mod. Transport. (2015) 23(2):93–106
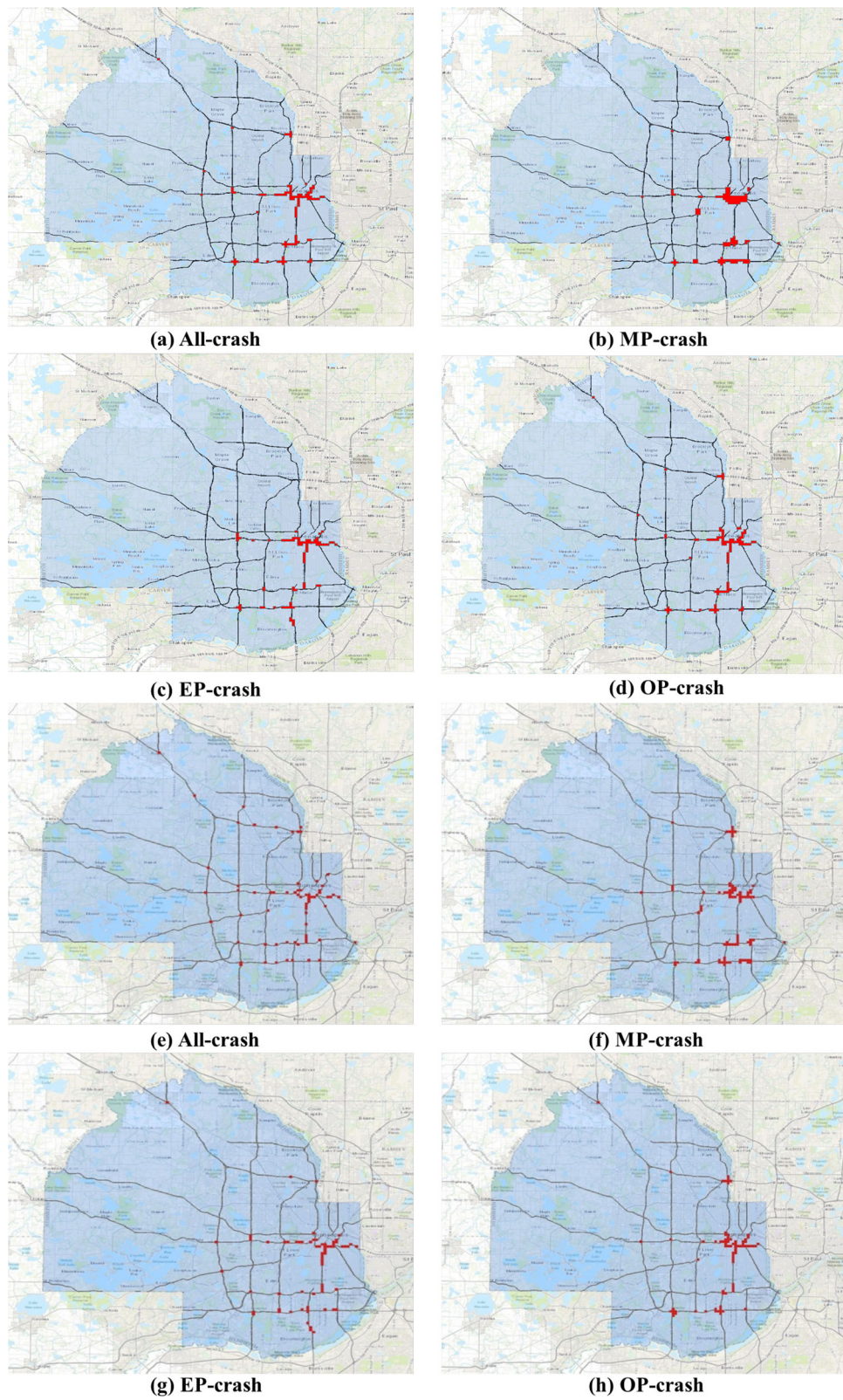
Fig. 9 Hotspots by KDE method for bandwidth 400 m: (**a**)–(**d**); kriging method: (**e**)–(**f**) (where *MP* morning peak, *EP* evening peak, *OP* off-peak)

Table 1 Performance comparisons of KDE and kriging methods

Time-of-day	No. of crashes in hotspot	Total crashes	Length of highway (km, two way)	Area coverage (km²)	PAI (length)	PAI (area)
KDE method (Bandwidth 400 m)						
All crash	14,239	38,748	107.04	16.32	2.75	3.69
MP crash	1,611	5,331	88.17	13.76	2.74	3.60
EP crash	2,700	7,712	82.79	13.28	3.38	4.32
OP crash	9,376	25,705	102.27	15.68	2.85	3.81
Total (T) & Avg. (A)	–	–	799.67 (T)	163.96 (T)	2.93 (A)	3.86 (A)
KDE method (Bandwidth 800 m)						
All crash	12,504	38,748	94.70	15.36	2.72	3.44
MP crash	1,175	5,331	82.89	11.84	2.13	3.05
EP crash	2,252	7,712	75.04	12.00	3.11	3.99
OP crash	8,146	25,705	89.26	14.24	2.84	3.65
Total (T) & Avg. (A)	–	–	799.67 (T)	163.96 (T)	2.7 (A)	3.53 (A)
Kriging method						
All crash	14,839	38,748	100.62	14.88	3.04	4.22
MP crash	1,907	5,331	106.61	15.36	2.68	3.82
EP crash	3,558	7,712	108.71	16.00	3.39	4.73
OP crash	10,886	25,705	107.17	16.64	3.16	4.17
Total (T) & Avg. (A)	–	–	799.67 (T)	163.96 (T)	3.07 (A)	4.23 (A)

The area coverage is not exactly 10 % due to loss of some cell sections during GIS processing
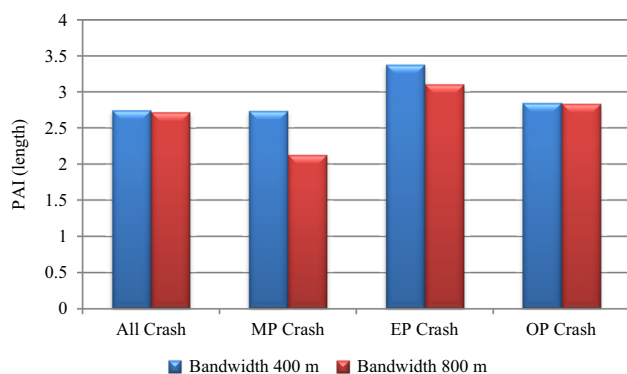
Fig. 10 Sensitivity of bandwidth in KDE method

locations for all four temporal groups. The intent of this non-performance-oriented comparison is to see if there exists a high (or low) match between the outcomes of the two methods, and investigate the feasibility of using one approach over another. The findings showed that the matching rates between the outcomes of kriging and KDE with 400 m bandwidth were found to be 52 %, 75 %, 66 %, and 71 % for All, MP, EP, and OP crash groups, respectively. This clearly suggests that there exist significant discrepancies between the two methods in identifying common hotspots. Similarly, matching rates using KDE with bandwidth 800 m were 45 %, 77 %, 62 %, and 69 % for All, MP, EP, and OP crash groups, respectively. Note that the comparison outcomes using two different bandwidths in KDE method were found comparable.

The above findings can be interpreted as follows: first, the average matching rate of 65 % indicates that the outcomes of the two test methods experience significant difference, suggesting that one method may not be used as a replacement of another. Moreover, as the PAI measures indicate that kriging method has better performance compared to KDE, we may conclude that kriging method, which is less explored in road safety, could be one of the potential methods in hotspots analysis. However, an open research question raises about which methods would produce more accurate results, should the reliability and credibility of the PAI index be questioned. Therefore, a further investigation is of high necessity to assert.

6 Conclusions and recommendations

This paper describes a comparative analysis on two geo-statistical-based approaches for estimating the expected collision frequency of individual road sections and identifying crash hotspots in a highway network. In contrast to the widely adopted safety model-based approach, a geo-statistical-based hotspot identification method is less data intensive and easier to implement, as it does not require extensive information about the underlying road network such as road geometry and traffic volume. The two geo-statistical methods considered in this analysis are called KDE and kriging. The KDE approach has been applied in a

few prior road safety studies while kriging, one of the least explored methods in road safety studies, was introduced in this research as a promising alternative because of its advantages in handling spatially autocorrelated datasets and success in other applications.

The two methods were compared in a case study for identifying crash hotspots in the road network of the Hennepin County, Minnesota. Five years of historical crash data, which were aggregated by different times of day, were used for geostatistically inferring the spatial distribution of the expected crash frequency using the two methods. The estimated crash frequencies were then used for subsequent hotspot identification, and the identification results were then compared using two criteria, namely PAI and percentage difference in hotspots identified. It was found that, according to the PAI criterion, kriging is superior in its ability to pinpoint the hotspots than that of the KDE. A comparison on hotspot ranking indicates that the two methods have resulted in moderately different lists of hotspots.

Regardless of the credibility of the evaluation criteria, it is worthwhile to note that kriging, which has seldom been used for road safety analysis, was shown to be a promising technique. The findings suggest that a further investigation is required to achieve more definite conclusions.

This research can be further extended to several directions to overcome a few limitations of the study conducted herein. First, a further investigation is needed to address the issue of how to incorporate the severity of individual crash data in hotspot identification. Second, instead of using the PAI measure, the performance of kriging can be benchmarked with the outcomes from the conventional crash model-based approach. Third, if we have historical geocoded data for potential crash influencing factors such as traffic exposure and weather conditions, we could apply universal kriging method and identify if these factors make significant contribution to hotspots. Such weather-related crash studies could be valuable for road agencies, especially in cold countries, for planning proactive winter road maintenance.

References

1. Hauer E (1992) Empirical Bayes approach to the estimation of "unsafety": the multivariate regression method. Accid Anal Prev 24(5):457–477

2. Saccomanno FF, Grossi R, Greco D, Mehmood A (2001) Identifying black spots along highway SS107 in southern Italy using two models. J Transp Eng 127(6):515–522

3. Greibe P (2003) Accident prediction models for urban roads. Accid Anal Prev 35(2):273–285

4. Miranda-Mareno LF (2006) Statistical models and methods for identifying hazardous locations for safety improvements. PhD thesis report, University of Waterloo

5. Cheng W, Washington S (2008) New criteria for evaluating methods of identifying hot spots. Transp Res Rec 2083(1):76–85

6. Highway Safety Manual (HSM) (2010) American association of state highway and transportation officials (AASHTO). Washington

7. Tarko, A.P., Kanodia, M., 2003. Hazard elimination program—manual on improving safety of Indiana Road Intersections and Sections

8. Kuo PF, Lord D, Walden TD (2012) Using geographical information systems to effectively organize police patrol routes by grouping hot spots of crash and crime data. Zachry Department of Civil Engineering, Texas A&M University

9. Pulugurtha SS, Krishnakumar KV, Nambisan SS (2007) New methods to identify and rank high pedestrian crash zones: an illustration. Accid Anal Prev 39(4):800–811

10. Erdogan S, Yilmaz I, Baybura T, Gullu M (2008) Geographical information systems aided traffic accident analysis system case study: city of Afyonkarahisar. Accid Anal Prev 40(1):174–181

11. Anderson TK (2009) Kernel density estimation and K-means clustering to profile road accident hotspots. Accid Anal Prev 41(3):359–364

12. Blazquez CA, Celis MS (2013) A spatial and temporal analysis of child pedestrian crashes in Santiago, Chile. Accid Anal Prev 50:304–311

13. Keskin F, Yenilmez F, Çolak M, Yavuzer I, Düzgün HS (2011) Analysis of traffic incidents in METU campus. Procedia Soc Behav Sci 19:61–70

14. Kim KE, Yamashita EY (2004) Using K-means clustering algorithm to examine patterns of pedestrian involved crashes in Honolulu, Hawaii CD- Rom TRB. National Research Council, Washington DC

15. Levine N, Kim KE, Lawrence HN (1995) Spatial analysis of Honolulu motor vehicle crashes: 1. spatial patterns. Accid Anal Prev 27(5):663–674

16. Levine N (2006) Crime mapping and the CrimeStat program. Geogr Anal 38(2006):41–56

17. Levine N (2009) A motor vehicle safety planning support system: the Houston experience. In S. Geertman and J. Stillwell (eds) Planning support systems: best practice and new methods. Springer, New York, pp 93–111

18. Kundakci E, Tuydes-Yaman H (2014) Understanding the distribution of traffic accident hotspots in Ankara, Turkey. In: Proceedings of the 93rd annual TRB conference, Washington, 12–16 Jan

19. Prasannakumar V, Vijith H, Charutha R, Geetha N (2011) Spatio-temporal clustering of road accidents: GIS based analysis and assessment. Procedia Soc Behav Sci 21:317–325

20. Troung LT, Somenhalli S (2011) Using GIS to identify pedestrian-vehicle crash hot spots and unsafe bus stops. J Public Transp 14:99–114

21. Khan G, Qin X, Noyce DA (2008) Spatial analysis of weather crash patterns. J Transp Eng 134(5):191–202

22. Pulugurtha S, Vanapalli V (2008) Hazardous bus stops identification: an illustration using GIS. J Public transp 11:65–83

23. Ha HH, Thill JC (2011) Analysis of traffic hazard intensity: a spatial epidemiology case study of urban pedestrians. Comput Environ Urban Syst 35(3):230–240

24. Olea RA (1999) Geostatistics for engineers and earth scientists. Kluwer, Boston

Springer

25. Oliver MA, Webster R (1990) Kriging: a method of interpolation for geographical information systems. Int J Geogr Inf Syst 3:313–332
26. U.S. Census Bureau (2014). Metropolitan and micropolitan statistical areas. United States Census Bureau, Population Division. Retrieved 9 June 2014
27. Silverman BW (1986) Density estimation for statistics and data analysis(Vol. 26). CRC press, Boca Raton
28. Bailey TC, Gatrell AC (1995) Interactive spatial data analysis. Longman, Harlow
29. Levine N (2004). CrimeStat III: a spatial statistics program for the analysis of crime incident locations (Version 3.0), National Institute of Justice, Washington, DC
30. Brimicombe AJ (2004). On being more robust about hotpots. Paper presented at the 7th annual international crime mapping research conference, April 2004
31. Krige DG (1951) A statistical approach to some basic mine valuation problems on the Witwatersrand. J Chem Metall Min Soc S Afr 52(6):119–139
32. Goovaerts P (1997) Geostatistics for natural resources evaluation. Oxford University Press, New York
33. Olea RA (2006) A six-step practical approach to semivariogram modeling. Stoch Environ Res Risk Assess 20:307–318
34. Chainey S, Tompson L, Uhlig S (2008) The utility of hotspot mapping for predicting spatial patterns of crime. Secur J 21(1):4–28
35. Van Pattern J, McKeldin-Coner D Cox (2009) A microspatial analysis of robbery: prospective hot spotting in a small city. Crime Mapp 1(1):7–32
36. Xie Z, Yan J (2008) Kernel density estimation of traffic accidents in a network space. Comput Environ Urban Syst 32(5):396–406
37. Hart TC, Zandbergen PA (2012) Effects of data quality on predictive hotspot mapping. Publication 239861. National Criminal Justice Research Service, Washington, DC