

## Research Article

# Methods for Identifying Truck Crash Hotspots

Wenrui Qu <sup>1</sup>, Shaojie Liu <sup>2</sup>, Qun Zhao <sup>2</sup>, and Yi Qi <sup>2</sup>

<sup>1</sup>School of Mathematics and Statistics, Qilu University of Technology (Shandong Academy of Sciences), University Road 3501, Changqing District, Jinan, Shandong, China

<sup>2</sup>Department of Transportation Studies, Texas Southern University, 3100 Cleburne Street, Houston, Texas 77004-9986, USA

Correspondence should be addressed to Yi Qi; [yi.qi@tsu.edu](mailto:yi.qi@tsu.edu)

Received 10 February 2020; Revised 28 September 2020; Accepted 14 October 2020; Published 28 October 2020

Academic Editor: Jing Dong

Copyright © 2020 Wenrui Qu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The goal of this study was to develop a new method for identifying the actual risky spots by using the geographic information system (GIS). For this purpose, in this study, three different methods for detecting hotspots are developed, i.e., (1) the annual average daily traffic (AADT) normalization method, (2) AK crashes (A is the incapacitating crash, and K is the fatal crash) percentage method, and (3) distribution difference method. To evaluate the performances of these three hotspot detection methods along with a baseline method that only considered the frequency of crashes, we applied these three methods to identify the top 20 hotspots for truck crashes in two representative areas in Texas. The results indicated that (1) all three proposed methods produced more reasonable results than the baseline method, and (2) the “distribution difference” method outperformed the other methods.

## 1. Introduction

Due to the size and weight of large trucks, their crashes often result in fatal injuries, property damage, and significant economic losses. According to a study conducted by the National Highway Traffic Safety Administration [1], in 2013, 342,000 large trucks were involved in traffic crashes, and these crashes killed 964 people and injured about 95,000 people. The analysis of historical truck crash data is a reliable, extensively used approach for identifying risk factors and preventing such crashes. However, analyzing crash data and reviewing the police reports for all of the crashes in the entire network is time-consuming and impractical. Detecting the hotspots and analyzing the crashes that have occurred at these locations provide a more effective way to identify the factors that cause crashes and to develop crash-prevention strategies.

Most of the methods that are currently used to analyze crash hotspots have no effective way of considering the impacts of roadway traffic conditions and exposure factors, and very few of them have taken account of the severities of the crashes. As a result, the hotspots that have been identified often are the spots with high traffic volumes or dense

roadway networks instead of especially risky spots. In addition, for some locations that have been identified as hotspots because many crashes have occurred there, most of the crashes have been minor crashes with no injuries. For example, Qi et al. [2] analyzed hotspots for truck crashes in Texas; but because they did not consider the volumes of traffic on different segments of the roadway, 7 of the top 10 hotspots they identified were in congested urban areas. In the same study, because the severities of the crashes that occurred were not considered, most of the top 10 hotspots that were identified for truck crashes were located near locations that generate or attract truck traffic, such as the distribution centers, rest areas, or stopping places for trucks.

The goal of this study was to develop an effective method for detecting truck crash hotspots by using the geographic information system (GIS). For this purpose, in this study, three new methods for detecting hotspots are developed and compared, i.e., (1) the annual average daily traffic (AADT) normalization method that considers both the frequency of crashes and the ADDT on a segment of the roadway; (2) the AK percentage method that considers both crash frequency and percentage of severe level crashes (AK crashes, where A is the incapacitating crash and K is the fatal crash); and (3)

the distribution difference method that is based on the difference between the distribution of AK crashes and the distribution of all types of the crashes. Among these three methods, the AADT normalization method that is also referred to as the crash rate method has been used by many previous studies [3–5]. The other two methods, i.e., AK percentage method and distribution difference method, are proposed by this study and have not been used by other studies in identifying crash hotspots. To evaluate the performances of these three hotspot detection methods along with a baseline method that only considered the frequency of crashes, we applied these three methods to identify the top 20 hotspots for truck crashes in two representative areas in Texas, i.e., the Houston–Galveston area and the Eagle Ford Shale area in South Texas. The results indicated that (1) all three proposed methods produced more reasonable results than the baseline method, and (2) the “distribution difference” method outperformed the other methods. The detected hotspots were evaluated based on the number of spots that were identified to have risky geometric or traffic features as already recognized by the literature and which are referred to in this study as the recognized high-risk spots for trucks. By comparing the numbers of recognized high-risk spots for trucks identified by different methods, recommendations were provided.

Following a brief review of previous studies on detecting hotspots for crashes, three proposed methods for detecting these hotspots are introduced. After that, descriptions of the study areas are provided. Then, the results of identifying the hotspots by different methods are compared and discussed. Finally, conclusions are presented based on the findings of this study.

## 2. Literature Review

Many studies have been conducted on developing methods for identifying hotspots for crashes in which the geographic information system- (GIS-) based geoprocessing and spatial analysis techniques were used. Among these methods, the point pattern analysis has been the most popular method. In this method, after geocoding of the crash events, the spatial distribution of crash data is analyzed to determine whether an observed distribution of point events results from a random pattern or whether it follows some systematic processes that form a clustered or regular pattern [6]. Some popular methods for point pattern analysis include nearest-neighbor distances, kernel-density estimation (KDE), and  $K$ -function [7].

KDE is one of the most extensively used methods. The goal of standard planar KDE is to develop a continuous surface of density estimates of discrete events, such as road crashes by summing the number of events within a search bandwidth. The KDE method has certain benefits in visualizing the crash density. The density value typically is the highest at the center, and it becomes smaller as the distance from the center increases [8, 9]. Pulugurtha et al. [10] used this method to study the zones in which there were large numbers of crashes involving pedestrians, while Erdogan et al. [11] studied the hotspots associated with highway crashes.

These point pattern analyses mentioned above account for spatial information, but they still treat all sites equally irrespective of their characteristics, i.e., each point is weighted equally. To solve this problem, an advanced point pattern analysis, i.e., the spatial autocorrelation method, was developed to take into account the locations of simultaneously discrete events and their values. The objective of the spatial autocorrelation is to have ranges of numbers to represent two spatial patterns, i.e., cluster and dispersion. The statistical significance of these two spatial patterns can be tested with the “ $z$  score” [12, 13], which is not the case with KDE.

These point pattern analysis methods are used to analyze the point distribution patterns for a 2D planar space, which is certain to raise controversial issues because road crashes often occur on the roads and inside the road networks that are portions of the 2D space. Therefore, road crashes should be considered in a road network space, represented by 1D lines [14]. Several studies have used a road network space to count point events [15–17]. Recently, transportation professionals have developed an ArcGIS-based package for the transport community [18].

Although many previous works have been performed, only a small number of studies have considered the impacts of exposure variables, such as the volume of traffic and the length of the segment of the road. In addition, most of these studies did not take the severity of the crashes into account in their analyses of the hotspots. Note that, in this study, a hotspot for crashes is defined as a location at which the risk of a crash is greater than it is at other locations. Thus, the term “hotspots” refers to locations at which more crashes tend to occur that have a high level of severity. Thus, both the frequency of crashes and their severity must be considered in the analysis of hotspots. This research is intended to ensure that this occurs by proposing three new, GIS-based methods for detecting hotspots for truck crashes and evaluating their performances.

## 3. Methodology

To consider the exposure factors and the weights of severe crashes (AK crashes) in the detection of crash hotspots, three hotspot detection methods were proposed, i.e., the AADT normalization method, the AK percentage method, and the distribution difference method. The basic concepts in these three methods are provided as follows.

**3.1. AADT Normalization Method.** The AADT normalization method considers the exposure factors and the volume of traffic in the detection of hotspots. In general, there are two exposure factors, i.e., the volume of traffic and the length of the segment of the road that is of interest. In this study, initially, all of the links in the roadway network were split into approximately equal, fixed distances. Therefore, the impacts of the length of the segment of the road being studied could be ignored, and only the volume of the traffic was considered as the exposure factor. In this study, the frequency of crashes, i.e., the number of crashes in 5 years,

was calculated for each small segment of the road. Subsequently, a crash rate normalized by AADT was estimated by the following equation:

$$\text{crash rate} = \frac{\text{segment crash frequency}}{\text{segment AADT}}. \quad (1)$$

**3.2. AK Percentage Method and the Distribution Difference Method.** Both the AK percentage method and the distribution difference method are designed to consider the weight of severe crashes (AK crashes) in the detection of crash hotspots. The crash hotspots are defined as the locations at which the risk of a crash is higher than it is at other locations. High risk means that there is a high probability of traffic crashes that meet or exceed a certain level of severity. In this study, we focused on AK crashes, i.e., incapacitating and fatal crashes, because these two types of crashes cause significant economic and social damages.

In the analysis of the risk associated with traffic crashes, both the frequency and severity of the crashes must be considered. Otherwise, the real crash risk cannot be identified. For example, Figure 1 shows the crash distribution under different lighting conditions. The left  $y$ -axis is for all the symbols in this figure except the orange curve, while the right  $y$ -axis is particularly for the orange curve. In Figure 1, the AK crash (total crash) distribution is the percentage of the AK crash (total crash) across different lighting conditions, while the AK percentage is only for a particular lighting condition, and it is the ratio of the number of AK crash occurred under a particular lighting condition vs. the total number of crashes occurred under this lighting condition. From Figure 1, it can be seen that, if only the frequency of crashes is considered when analyzing the risk of a truck crash under different lighting conditions, the daylight condition is the riskiest because most of the crashes occurred during that condition. Even if the frequency of AK crashes is the only consideration, it still was found that most of the AK crashes occurred during the daylight condition. As is well known, the daylight condition is not the riskiest condition, and more crashes occurred during this condition simply because there was

much more traffic. Also, it was found that the AK crash distribution at the “daylight” condition was much less than the total crash distribution. However, when the total crash distribution was compared with the AK crash distribution, it was found that the AK crash distribution for the “dark and not lighted” condition was much higher than the total crash distribution. This result indicated that even more crashes occurred during the daylight condition, but most of these were not severe crashes. However, only a few crashes occurred during the “dark and not lighted” condition, but most of them were severe AK crashes. Therefore, the severity of crashes must be considered in identifying the real risky conditions. Note that the curve of “AK crash distribution %–total crash distribution %” in Figure 1 indicates that the “dark and not lighted” condition has the highest value, and it is followed by the “dark and lighted” condition. The “daylight” condition has the lowest value. Thus, the crash risk conditions can be identified quite well by this curve. Similar results also have been obtained when analyzing other crashes risk factors, such as roadway alignment conditions and the condition of the surface of the roadway [2].

According to these findings, the AK percentage method and the distribution difference method were proposed. In the AK percentage method, the major selection criterion is the AK crash percentage, which is defined by equation (2) as follows:

$$\text{AK crash percentage} = \frac{\text{segment AK crash frequency}}{\text{segment total crash frequency}}. \quad (2)$$

In the distribution difference method, the major selection criterion is the difference between the AK crash distribution percentage and total crash distribution percentage, which is defined by equation (3) as follows:

$$\begin{aligned} \text{distribution difference} &= \text{AK crash distribution percentage} \\ &\quad - \text{total crash distribution percentage}, \end{aligned} \quad (3)$$

where

$$\text{AK crash distribution percentage} = \frac{\text{segment AK crash frequency}}{\text{AK crash frequency in the study area}}, \quad (4)$$

$$\text{total crash distribution percentage} = \frac{\text{segment total crash frequency}}{\text{total crash frequency in the study area}}. \quad (5)$$

**3.3. Procedure for Implementing the Proposed Detection Methods for Crash Hotspots.** The proposed methods for detecting crash hotspots can be implemented by using two GIS platforms, i.e., QGIS and ArcGIS. QGIS is used for the purpose of splitting the existing roadway links into segments that have approximately equal distances (500 meters in this study). The rest of the data processing is

performed in ArcGIS, a geographic information system for working with maps and geographical information. The entire procedure can be divided into the following four steps:

Step 1: Input roadway network layers into QGIS to split the road links into segments

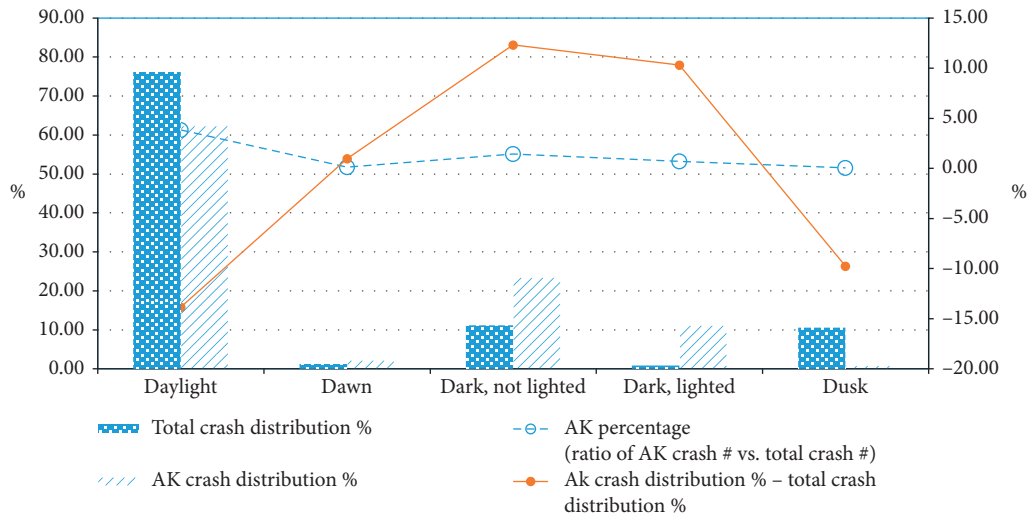


FIGURE 1: Truck crash risk analysis for different lighting conditions. Figure 1 is based on the truck crash data collected from the CRIS system in Texas for the period from 2011 through 2015.

Input the roadway network layer to QGIS, and use the “split” function to split the roadway links into roadway segments. They are the fixed segments. The length of the segments should be determined based on the specific research area and the scope, which can vary for different projects. In this study, the length of the roadway segment was specified to be 500 meters according to Chengye and Ranjitkar [19]. When the split is done, a new roadway network layer is produced, and its length is equal to or less than the given length of the segment. The new layer is imported into ArcGIS in Step 2.

Step 2: Derive the crash frequencies of the small road segments by using ArcGIS

Input historical crash data and the QGIS processed roadway network layer to ArcGIS, and map the crashes to the split roadway segments using the “spatial join”

function. The joining results must be examined carefully to ensure the correct joining between the crash points and the roadway segments based on the roadway information contained in the attribute table of the joined layer.

Step 3: Prescreen the roadway segments based on their crash frequencies

For all of the proposed methods, for detecting crash hotspots, prescreening based on the frequency of crashes is conducted to ensure that the frequencies of crashes on the selected segments are statistically higher than the frequency of crashes on the majority of the roadway segments in the study area. The prescreening threshold is given by the following equation (6):

$$\text{crash frequency} \geq \text{average crash frequency} + 1 \text{ standard deviation of crash frequency.} \quad (6)$$

Prescreening is used to identify the candidate roadway segments that will be analyzed further. In this step, to determine the member of the standard deviation in equation (6), a trial and error method was used. It was found that using one standard deviation can produce the appropriate amount of hotspot candidates for further analysis, while using two standard deviations will result in insufficient hotspot candidates (sometimes even less than 20).

Step 4: Calculate the major selection criteria, and select the riskiest road segments

Different methods for detecting crash hotspots have different major selection criteria. The selection criterion for the AADT normalization method is given by equation (1),

the selection criterion for the AK percentage method is given by equation (2), and the selection criterion for the distribution difference method is given by equation (3). In addition, for the baseline method, the selection criterion is just the frequency of crashes for each segment. By sorting the candidate roadway segments based on the calculated major selection criteria, the segments that have high values of the selection criteria are identified as the crash hotspots.

#### 4. Evaluation and Validation

4.1. Study Areas and Data. To evaluate and validate the proposed crash hotspot methods, two areas in Texas, i.e., the Houston–Galveston area and the Eagle Ford Shale area, were selected to identify the top 20 hotspots of crashes involving



trucks. Figure 2 shows the locations of these two representative areas.

The Houston–Galveston area represents a big metropolitan area with a high population and a high volume of truck traffic. Houston is the fourth largest city in the United States, and the Houston–Galveston area is a major freight traffic hub with the fourth largest port in the nation. Nearly 200 million tons of cargo move through the region annually in addition to commercial traffic generated by the numerous chemical facilities and petrochemical refineries.

The Eagle Ford Shale area represents a suburban or rural area with high truck traffic volume. It covers more than 25 counties in South Texas. Due to the availability of various techniques, such as horizontal drilling and hydraulic fracturing, the Eagle Ford Shale area has become one of the most active drilling areas in the world, resulting in an economic boom in the area. Unfortunately, one of the impacts of this boom has been a dramatic increase in truck traffic and crashes involving trucks.

The data that were used to identify crashes that involved large trucks included historical crash data and roadway network data. Data from the crashes that involved large trucks for the period of 2011–2015 were obtained from TxDOT’s Crash Records Information System (CRIS). The truck crash data that were available included crash severity levels and road characteristics, such as curves, grades, and whether the crashes occurred in rural areas or elsewhere. The GIS roadway network layer with the AADT attribute was downloaded directly from TxDOT’s official website.

**4.2. Evaluation Criteria.** To evaluate the performance of the proposed crash hotspot methods, the following evaluation criteria were used.

Evaluation criteria: the numbers of identified hotspots that are recognized as types of high-risk locations

Based on a previous study conducted by employees of the University of Kentucky [20] and the traffic safety statistics [1], the following specific types of roadway locations were identified as the riskiest locations for truck crashes:

- Grades
- Curves
- Rural roads
- Intersections
- Interchanges

In this study, these types of locations are referred to as recognized high-risk locations for trucks. If an identified hotspot belongs to one of the recognized high-risk locations, it is more likely that this spot presents high risks for trucks. Therefore, the total numbers of identified hotspots that belong to each type of the recognized high-risk location are the measurements for the correction of the detection results. Higher values of this criterion indicate better detection results.

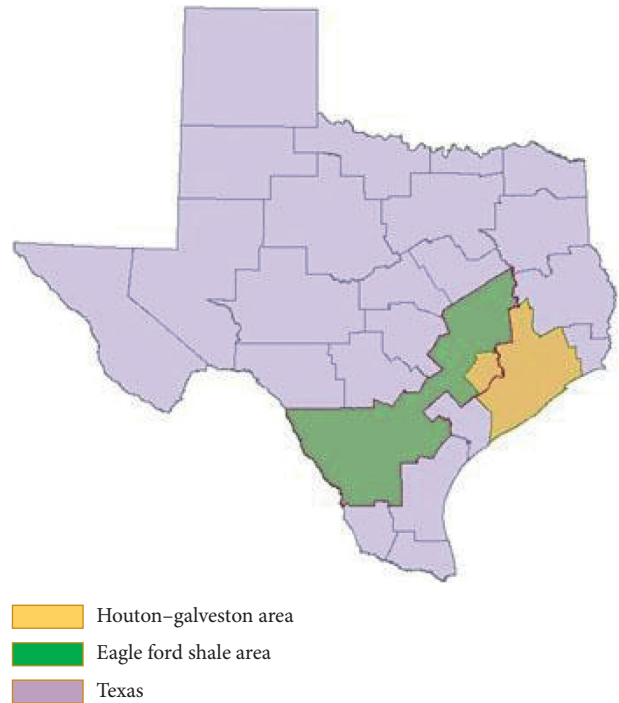


FIGURE 2: The two study areas that were selected.

## 5. Results and Discussion

The proposed three methods for detecting hotspots, along with the baseline method that only considers the crash frequency, were used to identify the top 20 hotspots for truck crashes in the two study areas. As mentioned earlier, one of the problems in the existing hotspot analysis method is that many of the detected hotspots are near a truck distribution center (TDC). These spots usually are not very risky, and the high crash rates more likely were due to the high volume of truck traffic. Therefore, if most of the identified hotspots are close to a TDC, it indicates that the detection result is biased and not reliable. In this study, 1000 meters was selected as the threshold for determining whether or not a hotspot was close to a TDC, and the number of identified hotspots that were close to truck distribution centers was counted. The statistic description of the numbers of identified hotspots close to a TDC is presented in Table 1. According to Table 1, it can be observed that the baseline method, i.e., crash frequency, has the highest number of observations close to the TDC in both studied areas. Additionally, the distribution difference method has the lowest number of observations close to TDC. These results potentially indicate that the distribution difference method can detect the crash hotspots by considering the impacts of traffic exposures.

**5.1. Houston–Galveston Area.** For this study area, the results of the hotspots that were identified by three different methods are presented in the maps in Figure 3. It is apparent that the identified hotspots are different. The top 20 hotspots identified by the three proposed methods are farther apart

TABLE 1: Numbers of identified hotspots close to TDC.

	Close to TDC	
	Eagle Ford Shale	Houston–Galveston
Baseline method: crash frequency only	<b>16</b>	<b>5</b>
AADT normalization method	2	4
AK parentage method	4	2
Distribution difference method (AK%–TC%)	<b>1</b>	<b>2</b>

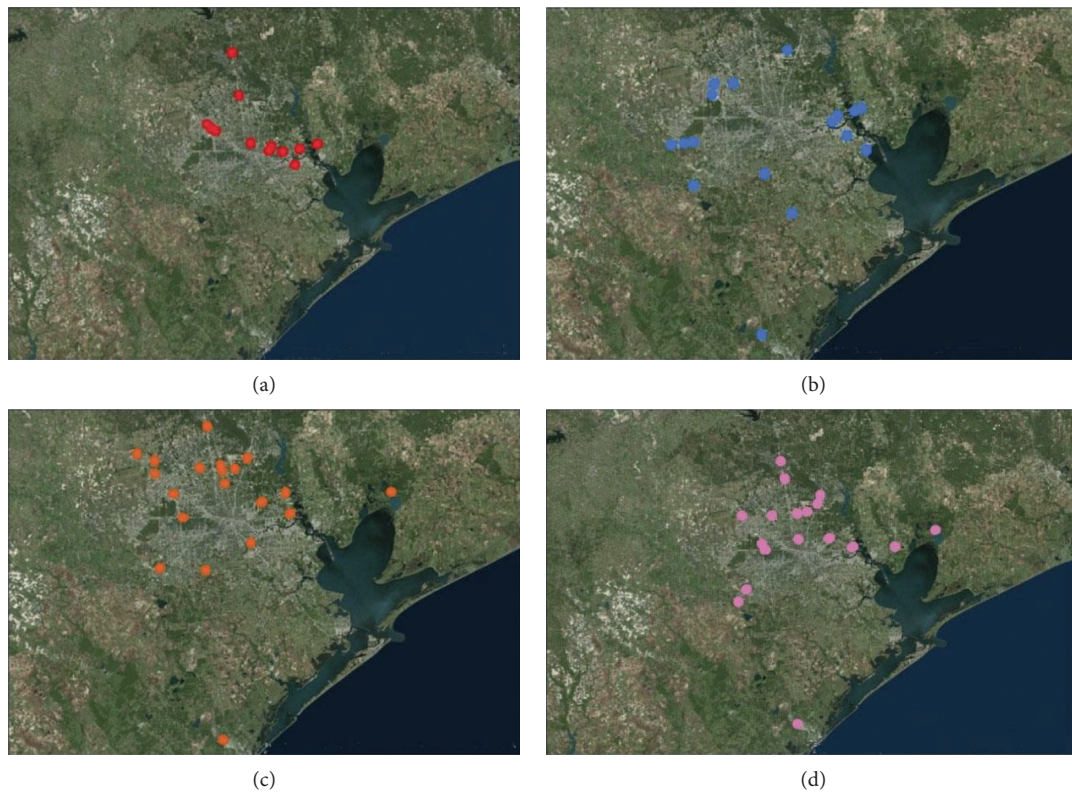


FIGURE 3: Hotspots identified in the Houston–Galveston area by four different methods for detecting crash hotspots. (a) Hotspots identified by the baseline method. (b) Hotspots identified by the AADT normalization method. (c) Hotspots identified by the AK percentage method. (d) Hotspots identified by the distribution difference method.

than the 20 hotspots that were identified by the baseline method that only considers the frequency of crashes.

By closely examining the detected results, the number of identified hotspots that belong to each type of recognized high-risk location was derived for each method. The results are presented in Table 2. If an identified hotspot belongs to one of the recognized high-risk locations, it is more likely that this spot presents high risks for trucks. Therefore, the more detected hotspots belonging to the identified high-risk location types, the better the detection result.

In Table 2, by comparing the detection results of different hotspot detection methods, the best results were identified and marked in green, and the worst ones were also identified and marked in red. It can be seen that, overall, the distribution difference method has the best performance because it identified more hotspots than the other methods that belong to the recognized high-risk location types, including

at curves, at locations where there were grades, in rural areas, and at interchanges. However, the distribution difference method detected fewer hotspots at intersections, but the AK parentage method detected more hotspots at intersections. The baseline method identified the fewest hotspots at curves and in rural areas.

**5.2. Eagle Ford Shale Areas.** For this study area, the results of the identified hotspots by four different methods are presented in the maps in Figure 4. The differences in the spatial patterns are even more conspicuous in this area. The hotspots identified by the three proposed methods are not near each other, whereas the hotspots identified by the baseline method were clustered at one location at which the volume of truck traffic was relatively high.

The results of the four methods in the Eagle Ford Shale area are presented in Table 3.

TABLE 2: Geographic features of the hotspots in the Houston–Galveston area.

	Number of identified hotspots at the curve	Number of identified hotspots at a location with grade	Evaluation criteria		
			Number of identified hotspots at intersection	Number of identified hotspots in the rural area	Number of identified hotspots at the interchange
Baseline method: crash frequency only	8	12	8	10	12
AADT normalization method	10	12	8	17	10
AK parentage method	8	6	12	12	7
Distribution difference method (AK%–TC%)	11	14	7	15	12

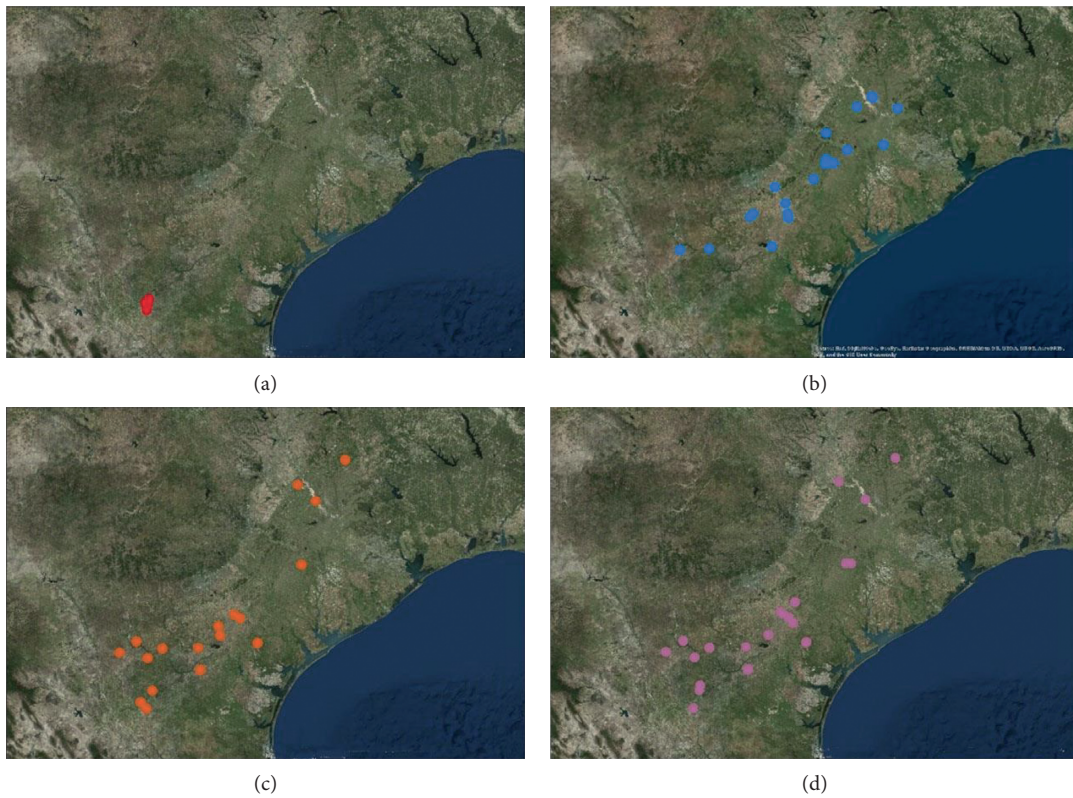


FIGURE 4: Hotspots in the Eagle Ford Shale area as identified by four different methods for detecting crash hotspots. (a) Hotspots identified by the baseline method. (b) Hotspots identified by the AADT normalization method. (c) Hotspots identified by the AK percentage method. (d) Hotspots identified by the distribution difference method.

TABLE 3: Geographic features of the hotspots in the Eagle Ford Shale area.

	Number of identified hotspots at the curve	Number of identified hotspots at a location with grade	Evaluation criteria		
			Number of identified hotspots at intersection	Number of identified hotspots in the rural area	Number of identified hotspots at the interchange
Baseline method: crash frequency only	6	7	19	5	1
AADT normalization method	7	4	17	19	0
AK parentage method	6	7	19	18	1
Distribution difference method (AK%–TC%)	7	13	10	20	3



The results in Table 3 indicate that the distribution difference method still outperformed the other three detection methods, since it identified more hotspots that belong to the recognized high-risk location types, including at curves, at locations where there were grades, in rural areas, and at interchanges. However, it also was found that the distribution difference method detected less hotspots at intersections. Similar to the results in area 1, the AK percentage method can detect more hotspots that are at intersections. In addition, the baseline method identified the fewest hotspots at curves, in rural areas, and at interchanges.

## 6. Conclusions

In this study, we proposed three hotspot detection methods that can consider the impacts of exposure variables or the severity levels of crashes, i.e., the AADT normalization method, the AK percentage method, and the distribution difference method. To evaluate their performances, these three hotspot detection methods, along with a baseline method that only considered the frequency of crashes, were used to identify the top 20 truck crash hotspots in two representative areas in Texas. Based on the detection results, the following key findings were obtained:

- (1) If only the crash frequency is considered in the process of identifying crash hotspots, the identified hotspots are likely to cluster in one area where there is a high volume of traffic
- (2) Overall, the distribution difference method outperformed the baseline method, the AADT normalization method, and the AK percentage method because it was able to detect more spots associated with locations that are recognized as risky for trucks and to detect fewer spots that are near a TDC
- (3) The AK percentage method is recommended for detecting hotspots at intersections

This research provides useful ideas on the detection of crash hotspots and a new type of criteria for evaluating the performance of crash hotspot detection methods. One limitation of this study is that the proposed methods for detecting crash hotspots were evaluated based only on their detection results in two representative areas. To further validate and refine the proposed methods, more locations with different traffic and roadway network conditions should be selected as study areas in the future.

## Data Availability

The data used to support this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This research was supported in part by the United States Department of Transportation (USDOT) (grant #69A3551747133) and the U.S. Department of Homeland Security (DHS) (grant #2014-ST-062-000057-02).

## References

- [1] National Highway Traffic Safety Administration, *Traffic Safety Facts: Rural/Urban Comparison*, US Department of Transportation, Washington, DC, USA, 2014.
- [2] Y. Qi, Q. Zhao, P. Liu et al., *Development of Systemic Large Truck Safety Analyses*, Texas Department of Transportation (TxDOT), Washington, TX, USA, 2017.
- [3] A. P. Tarko and M. Kanodia, "Hazard elimination program-manual on improving safety of Indiana road intersections and sections," 2004.
- [4] R. Elvik, "Comparative analysis of techniques for identifying locations of hazardous roads," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2083, no. 1, pp. 72–75, 2008.
- [5] A. Montella, "A comparative analysis of hotspot identification methods," *Accident Analysis & Prevention*, vol. 42, no. 2, pp. 571–581, 2010.
- [6] T. Bailey and T. Gatrell, *Interactive Spatial Data Analysis*, Longman Scientific & Technical, Harlow Essex, UK, 1995.
- [7] I. Yamada and J.-C. Thill, "Comparison of planar and network  $K$ -functions in traffic accident analysis," *Journal of Transport Geography*, vol. 12, no. 2, pp. 149–158, 2003.
- [8] S. S. Vemulapalli, "GIS-based spatial and temporal analysis of aging-involved crashes in Florida," Doctoral dissertation, The Florida State University, Florida, FL, USA, 2015.
- [9] L. T. Truong and S. V. Somenahalli, "Using GIS to identify pedestrian-vehicle crash hotspots and unsafe bus stops," *Journal of Public Transportation*, vol. 14, no. 1, p. 6, 2011.
- [10] S. Pulugurtha and V. Vanapalli, "Hazardous bus stops identification: an illustration using GIS," *Journal of Public Transportation*, vol. 11, no. 2, pp. 65–83, 2008.
- [11] S. Erdogan, I. Yilmaz, T. Baybura, and M. Gullu, "Geographical information systems aided traffic accident analysis system case study: city of afyonkarahisar," *Accident Analysis & Prevention*, vol. 41, 2008.
- [12] D. W. S. Wong and J. Lee, *Statistical Analysis of Geographic Information with ArcView GIS and ArcGIS*, Wiley, Hoboken, MA, USA, 2005.
- [13] A. Getis and J. K. Ord, "The analysis of spatial association by use of distance statistics," *Geographical Analysis*, vol. 24, no. 3, pp. 189–207, 1992.
- [14] Z. Xie and J. Yan, "Detecting traffic accident clusters with network kernel density estimation and local spatial statistics: an integrated approach," *Journal of Transport Geography*, vol. 31, pp. 64–71, 2013.
- [15] J. Young and P. Y. Park, "Hotzone identification with GIS-based post-network screening analysis," *Journal of Transport Geography*, vol. 34, pp. 106–120, 2014.
- [16] K. Sugihara, T. Satoh, and A. Okabe, "Simple and unbiased kernel function for network analysis," in *Proceedings of the 2010 International Symposium on Communications and Information Technologies (ISCIT)*, pp. 827–832, IEEE, New York, NY, USA, 2010.



- [17] A. Okabe, K.-i. Okunuki, and S. Shiode, "SANET: a toolbox for spatial analysis on a network," *Geographical Analysis*, vol. 38, no. 1, pp. 57–66, 2006.
- [18] A. Okabe, T. Satoh, and K. Sugihara, "A kernel density estimation method for networks, its computational method and a GIS-based tool," *International Journal of Geographical Information Science*, vol. 23, no. 1, pp. 7–32, 2009.
- [19] P. Chengye and P. Ranjitkar, "Modeling motorway accidents using negative binomial regression," *Journal of the Eastern Asia Society for Transportation Studies*, vol. 10, pp. 1946–1963, 2013.
- [20] J. G. Pigman and R. Kenneth, "Agent heavy truck involvement in traffic accidents and related countermeasures," *University of Kentucky*, vol. 10, 1999.