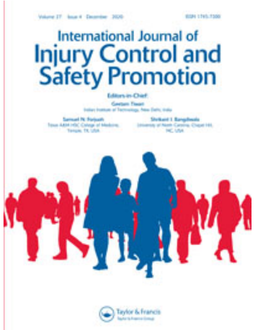

International Journal of Injury Control and Safety Promotion

ISSN: 1745-7300 (Print) 1745-7319 (Online) Journal homepage: https://www.tandfonline.com/loi/nics20

Review of road accident analysis using GIS technique

Monib Shahzad

INTERNATIONAL JOURNAL OF INJURY CONTROL AND SAFETY PROMOTION
2020, VOL. 27, NO. 4, 472–481
https://doi.org/10.1080/17457300.2020.1811732

REVIEW ARTICLE

Review of road accident analysis using GIS technique

Monib Shahzad

Department of Civil Engineering, Pakistan Institute of Engineering & Technology, Multan, Pakistan

ABSTRACT
The number of road crashes is significantly growing worldwide. In the transportation sector, accident outcomes are usually the loss of lives and injuries. To avoid further damages, a tool entitled geographical information system (GIS) could be helpful. GIS has the most demanding tools used to analyze road accidents and road design that can be noteworthy in traffic accident prevention. The purpose of this review is to propose the superlative approach of GIS applicable to accident analysis in different circumstances. The reviewed statistical results of accidents are performed by GIS but the numerical study is not consummate by GIS. Mainly, four essential GIS techniques are introduced and discussed in this review paper to simulate road accidents and suggest some prolific accident analysis tools for road safety.

ARTICLE HISTORY
Received 30 March 2020
Revised 11 August 2020
Accepted 14 August 2020

KEYWORDS
Road safety; transport; geographical information systems; accident analysis

1. Introduction

Road accidents are known to be the main problem for emerging countries, as they offer substantial financial and social dead. According to the World Health Organization (WHO) report of 2018, almost 1.35 million people died globally in road accidents every year. Moreover, the greatest death accidents happen in countries where wages are low (WHO, 2018). Almost 300 thousand people injured and 85,000 died in road accidents in India, annually, tolerates huge human and economic loss (Mohan, 2009). As per the WHO prediction, road crashes will become the third most powerful source of casualties in 2020 (WHO, 2007).

Speed is one of the prime causes of road accidents. This study directed the clustering of crashes before and after speed limit surge occurred in Ankara. There was approximately 10% enhancement in the total number of crashes before and after crossing speed limit, and hotspots traced using geographical information system (GIS). Complete assessment of the study showed that speed limit enlargement was not helpful to improve road safety, hence, threatening the road consumers at different locations (Kibar & Yaman, 2020). Furthermore, Kernel density estimation (KDE) technique has been applied to identify the crashes hotspots pertaining to specific time and season for 3 years using Severity Index (SI). The results disclosed that the ranking of hotspots was more accurate. But KDE alone has a disadvantage that it faces ambiguity regarding the precise location of a road crash (Bíl et al. 2019; Le et al. 2020). Colak et al. (2018) studied the accident blackspots over a 45 km segment of Rize, Turkey using hotspot statistical analysis which is barely employed by other studies. Data collected for 5 years were then analyzed using Getis-Ord tool. Results showed that hotspot statistical analysis provided a more precise hotspot location, hence, this method is favorable when the exact accident location is required.

In the recent past, several studies have been accomplished regarding the road accident analysis execution. Moreover, GISs have the capability to complete the spatial analysis task in a short period. Similarly, scientists used GIS to store geographical data, analyzing the statistical results of crashes, and represent this road crashes data in spatial patterns (Dereli & Erdogan, 2017; Gundogdu, 2010; Kazmi & Zubair, 2014; Loo, 2006; Truong & Somenahalli, 2011). For example: Truong and Somenahalli (2011) identified the hotspots of pedestrians-vehicles crash using contiguous and severity pattern of Getis statistics, and then Moran's *I* was applied to assess the spatial pedestrian-vehicle data. Results revealed that 3 pedestrian-vehicle darkspots were detected at crossing and 10 pedestrian-vehicle darkspots were detected at mid-block position within specific timeframe.

Hashimoto et al. (2016) developed a traffic flow density model using GIS that assures the specific traffic flow to a widespread extent, especially, when the road crash statistics are unobtainable. This model estimated traffic flow relative to other factors including road engineering, population density, and geospatial aspects, etc. This research evaluated the relationship between city characteristics and road crashes. An estimation model of KDE was used in relationship with 16 other models. Results revealed much higher coefficients between actual and estimated numbers. This model can help in locating relative hazardous areas in cities. Tasic et al. (2017) used generalized additive model to evaluate the accumulated statistical crash data of tandem, pedestrian, and vehicles at a macroscopic level in Chicago. The results specified that the crash percentage declines with the

CONTACT Monib Shahzad monibshahzad403@gmail.com
This article was originally published with errors, which have now been corrected in the online version. Please see Correction (http://dx.doi.org/10.1080/17457300.2020.1816260)

rise of the exposure during the spatial examination. This is a way forward towards sustainable infrastructure and useful specially for pedestrians and cyclists for taking further safety measures in transportation.

Mostly scholars correlate the numerical models with GIS to review the traffic accidents hazard, i.e. practically numerical models and GIS techniques have been instigated (Li et al., 2007). The data formation and dissection are usually showing within GIS podium but the Bayesian study is found ineligible in GIS. Moreover, researchers endorsed this process because it helps in the import and export of data into different frameworks makes it user-friendly. This process can be applied autonomously utilizing numerical models which is an added feature (Shankar et al., 1995). Many studies have applied the kernel density method for examining traffic crashes through geospatial perceptions (Blazquez & Celis, 2013; Hashimoto et al., 2016).

KDE is the method commonly used for geographical crash analysis and it also identifies the crashes blackspots (Bailey & Gatrell, 1995). The K-function technique, because it measures point patterns and point-to-point measurements, is considered to be an efficient and compendious spatial method (Yamada & Thill, 2004). Fox et al. (2015) mentioned a Bayesian spatial method to determine pedestrian death accidents, results presented the intrusion policies for high death extents, and showed the dark spot locations where the pedestrian death frequency is tremendous. Cordera et al. (2018) utilized Gravity and Poisson method to assess travel costs and accident occurrence at railway stations. Results found that the gravity model due to spatial effects provides the finest results and offers a faithful trip pattern than the poison method.

A decision-making model was built using the Fuzzy AHP process to recognize road accidents (Nanda & Singh, 2018). Cafiso et al. (2010) functioned the global positioning system and the inspection method to explore rustic roads using road geometry variables. Flahaut et al. (2003) analyzed the brittle-Ord Gi* and Kernel density estimation (K) to identify crash-prone locations. Outcomes revealed that the results of both of these methods were practically matching. Niloofar (2017) analyzed and managed metropolitan road accident data of Iran using the GIS method. Shariat and Shahri (2017) used a semi-local Poisson-Gamma CAR model to analyze road crashes, in correlation with Moran's *I*, however, both models showed the same coefficients with some exceptions.

In recent times, GIS-based research on road crashes and the altitudinal progressive method were performed in the United Kingdom. The research used the K function assessment approach for the identification of clattered hot spot existence. The scientists ultimately found that the unified clattered crashes along these spots could cause accidents to reduce significantly. In the current review paper, several GIS techniques to be implemented on accidental roads for the altitudinal examination have been reviewed. Most vital studies accomplished for traffic analysis using GIS are given in Table 1.

This paper is different from previous studies, as it is a new approach to deal with the methodological problems to identify road crashes. This study is focused on the geospatial mechanism of crashes by making comparisons of four different spatial techniques in terms of strengths and weaknesses (view Table 1) in order to know the practical applications of the study. This review intends to improve the road safety problems across different cities like the application of Getis-Ord and KDE techniques jointly. This study, as a policy development outlook, will help the higher authorities, safety experts, and teachers/students working on GIS to take safety measures against the great number of casualties that occur in the critical places of the world.

The residue sections of this study are organized in the following manner. Section 2 presents a background on the GIS-based techniques or methods. Section 3 and section 4 describe the methodology and geospatial techniques used, while section 5 evaluates the results. Section 6 presents the discussion along with limitations and future research guidelines. The last section presents the conclusion of the study.

2. Background

The target road data are originated in diffusion that makes the analysis part quite complex. A number of academics have successfully examined the road sections and junctions altogether (Erdogan et al., 2008; Huang & Chin, 2010; Ma et al., 2014). But it has been stated that the crashes factors arisen at road sections and the reasons involved on junctions stayed uncoherent, thus, it verifies that roads and crossings could not be mutually investigated (Moore et al., 2011).

2.1. Zone type

The traffic accident record when analyzed guarantees safety measures. Accurate, reliable, and high-quality data is a must to recognize the factors usually involved in road accidents. The data in Table 1 presents that eight varieties of research collected the data from metropolitan, rural, suburban highway, and motorways, whereas eleven studies are examined in metropolitan areas only.

2.2. GIS traffic accident analysis

The methods that are being used by the world scientists over the last twenty years for the GIS analysis of road accidents are shown in Table 1, e.g. observation of accident formation along with geographical dispersion accidents (Budiharto & Saido, 2012; Gundogdu, 2010; Levine et al., 1995 Truong & Somenahalli, 2011), and identification of the areas vulnerable to accidents (Rankavat & Tiwari, 2013). Recently, the trend of using data assessment and GIS together is growing tremendously and this trend of evaluating the traffic accidents is more often employed by the world researchers (Benedek et al., 2016; Erdogan, 2009; Erdogan et al., 2008; Ma et al., 2014; Steenberghen et al., 2004; Tortum & Atalay, 2015; Yalcin & Duzgun, 2015). The spatial clustering of crashes and geographical densities of Shanghai using Moran's Index, KDE and Network KDE (NetKDE) methods has been evaluated by (Chen et al.

Table 1. Comparison of previous researches applying GIS methods for accident analysis.

Writer and publishing period	Zone	Strengths and weaknesses	Methods and summary
Shariat and Shahri (2017)	Mixed	This research technique-AHP can pattern crash data in simple and unique way with more precision. However, when multi-layered data is provided in AHP, it usually fails to provide accurate information.	This study used Fuzzy-AHP technique to examine the most significant factor for road crashes. Implementing spatial autocorrelation model as Moran's Index on 253 traffic analysis zone of Mashhad, Iran indicated more consistency than traditional generalized linear models. Hence, this spatial model can be useful tool for higher authorities to predict their road safety effects.
Niloofar (2017)	Metropolitan	This research analyzed and managed metropolitan road accidents data of Iran in reliable and effective means using GIS. The main weakness of the study is the methodology of collecting accident data. All local authorities have their own way of collecting and maintaining data which are not even linked. This drives to mismatching of data, missing of data and poor results.	This research summarized and analyzed diverse attributes including accident location, road user information, road type, vehicle type, and accident severity using GIS which is quite challenging. The results depicted that the GIS can not only execute the spatial data and highlight black spots but also save the time for data processing.
Yalcin and Duzgun (2015)	Metropolitan	The asset of this research is that it used GIS as a management tool for crash analysis in association with other statistical methods. On the contrary, this study used only Network kernel density method which confines the study's scope.	There are three altitudinal design methods used in this study, like K-function, adjacent neighbor distance, and Kernel density in order to determine the occurrence of clustering and darkspots.
Tortum and Atalay (2015)	Mixed	This research found that the spatial tool turned out to be a unique and useful for identifying problematic or hot accidental areas of Turkey. This study employed conventional regression technique which is, at present, undervalued.	In this study, researchers applied Moran's *I* statistic to clustered the traffic accidents in the provinces of Turkey. The mortality rates in underdeveloped areas were found higher in accordance with road traffic accidents in the provinces of Turkey.
Ma et al. (2014)	Metropolitan	The study showed huge impact on accident hazard impact after comprehensive risk assessment of accidents is done with over detached statistics. This work didn't consider the crashes of different types in regression model.	Scientists applied the statistical analyses method of Quasi-Poisson to find the severity level of crashes. Results suggested that the weight value higher than 0.8 would consider to be a fatal accident.
Effati et al. (2014)	Metropolitan	This study effectively identified the hazardous zones of selected corridors of Iran in cheaper way, even when the existing data is missing and ambiguous. The weaknesses of this study are that the Author was not sure whether the collected crash data was correct and complete or not. Secondly, the number of crashes was not that higher on newly constructed roads, so statistical methods are not convenient on these types of roads.	This study introduces a new Neuro-fuzzy approach for crashes identification of risky areas. In addition, this research method identified some additional zones on the existing hotspots that have not been recognized by conventional statistical methods.
Rankavat and Tiwari (2013)	Mixed	The purpose of this study was to assess the effectiveness of GIS in analyzing intricate road crashes data. GIS solved this problem, and patterned the crashes data. The crash data collected from police, carried out in this research was not accurate, which is one of its weaknesses.	This study identifies the pedestrian zones of Delhi, India that are vulnerable to accidents. Following Kernel density and Moran's *I* methods are being utilized for analyzing statistics from 2006 to 2009. Results revealed that 8503 deaths occurred, and approximately 51% of deaths involved pedestrians at intersections, with 14% cars and 27% buses were responsible for casualties.
Truong and Somenahalli (2011)	Mixed	The research conducted in Adelaide, Australia for identifying pedestrian-vehicle crash which proved to be an accurate and reliable approach. This research analyzes the old pedestrian-crash data, in addition, a multi-criterion ranking approach could bring more improvements.	This research method involves the identification of hotspots of pedestrians-vehicles crash using contiguous and severity pattern of Getis statistics, and then Moran's *I* was applied to assess the spatial pedestrian-vehicle data. Results revealed that 3 pedestrian- vehicle dark spots were detected at crossing and 10 pedestrian-vehicle dark spots were detected at mid-block position.
Steenberghen et al. (2004)	Mixed	The local Moran Index performed well in clustering rural roads of Belgium. While, KDE approach is only applicable for urban areas.	This research work utilized autocorrelation approach by Moran' *I* and KDE for identifying clustered areas of Belgium. Both methods performed well in their own locality.

2018). Results found that 24,147 people were influenced from 2010 to 2012 and 84% victims were male gender. Additionally, 76.2% of vehicles involved in crashes were bikes (60.1%) and buses (16.1%), thus, this research provided help to authorities in pinpointing vulnerable sites.

Kuo et al. (2012) not only compared the polar KDE method with Getis-Ord and NetKDE but also identified hotspots with valuable guidelines regarding methods selection. Additionally, statistical information has a principal role in finding improved results. KDE method has the main weakness that it computes the density of neighborhood crashes without considering the statistical information (Anderson, 2009). Erdogan et al. (2015) revealed that employing divergent hot spots identification methods produces different results.

Flahaut et al. (2003) employed KDE and Moran's *I* method in the vicinity of Belgium for detecting accident hot spots. The outcomes of 510 accidents that occurred from 1992 to 1996 suggest that both methods produced identical results. Zhang (2010) recently found that the local Getis-Ord index identifies the crashes dark spots whenever statistical information is provided while point density model is operable for locating accident hotspots even when statistical data is not provided. Consequently, local Getis-Ord index executed finer results than point density when it comes to risky road spots identification. Pulugurtha et al. (2007) used crash score, KDE, and sum of rank method and indicated 29 pedestrian crash areas in Los Angelos (a metropolitan city of the United States). Furthermore, the results were coherent with slight to no alteration when crash score and sum of rank methods were applied for rankings of pedestrian crash areas.

Truong and Somenahalli (2011) identified the hotspots of pedestrians-vehicles crash using contiguous and severity patterns of Getis statistics, and then Moran's *I* was applied to assess the spatial pedestrian-vehicle data. Results revealed that 3 pedestrian-vehicle dark spots were detected at crossing and 10 pedestrian-vehicle darkspots were detected at the mid-block position. However, a larger proportion of pedestrian-vehicle accidents were found to be at Main North road's crossings. Tortum and Atalay (2015) scrutinized specifically those variables that can make a great impact on the total accident figure. This study applied Moran's *I* statistic to clustered the traffic accidents in the provinces of Turkey. The mortality rates in underdeveloped areas of Turkey's provinces were found higher in accordance with road traffic accidents.

Getis-Ord and Moran's *I* statistics are the two approaches frequently used in the literature. These spatial techniques are useful in analyzing the spatial accident phenomena and to identify the access location. For instance, the occurrence of some sort of perilous accidents in certain locations. These types of GIS spatial techniques are well elucidated in the subsequent section.

Occasionally, factor analysis is also implemented. This analysis tends to exploit important information from small to large data set under restricted conditions.

In contrast, Shaggy sense has already been utilized by scholars. Effati et al. (2014) applied the Neuro-fuzzy approach shown in Table 1 on the highway for marking risky areas. Then, a comparison was made between the risky and actual black spots attained from the numerical method. This method identified some additional zones on the existing hotspots that have not been recognized by conventional statistical methods.

2.3. Statistical accident assessment

The accident numbers like pedestrian clash with cars, motorbikes, buses, collision peak time, road geography, and further mixtures can be identified (Rankavat & Tiwari, 2013). Accidents' identification is possible by using a robust partition. Robust partition is used to spot the clatter data by establishing links in the presence of comparative locations and when regional data are unknown (Yuan et al., 2005). Furthermore, Repeatability and Kernel Density analysis method was utilized by (Erdogan et al., 2008). The collected accident constraints including accident place, vehicle collision time, weather conditions, driver's profile, accident austerity, and vehicular type were used. It is found that repeatability method of analysis measured more hotspots as compared to KDE method. For example, Repeatability showed 12 hotspots at Ayonkarshisar station, while KDE showed 9 hotspots at the same station. This is the way forward towards highway safety measures by GIS. Although, some intellectuals intermix the road traffic statistics and spatial attributes of a road in their imminent researches.

2.4. Objectives of the study

This review article has the following main objectives. Firstly, it proposes a methodology/technique to analyze road crashes. Secondly, it provides a way forward to prevent and identify road crashes using GIS techniques. GIS has a feature of displaying statistical data in the geographic form and identify road crashes hotspots. So, applying GIS will be valuable in achieving above stated goals. In addition, four different GIS techniques have been discussed to study the geospatial pattern of crashes. To be specific, Getis-Ord andKDE techniques jointly will extricate new variables, define populous crash neighborhood, and identify the maximum hotspots in petite time, as these features are no being offered in existing studies. Thus, this study will contribute towards the improvement of road safety problems across different cities of the world.

3. Methods

3.1. Search strategy

The literature was searched pertinent to road accident analysis. The related papers were searched in the database of Scopus (https://www.sciencedirect.com/) and Springer's database (https://www.springernature.com/gp). The search keywords were carefully chosen like 'Road Accident', 'Accident

Analysis' and 'GIS' to classify pertinent journals. These keywords were hunted in the journals presented in Table 2. The papers marked with keywords while searching two databases were ($n = 7658$) in numbers as shown in Figure 1. Papers after excluding duplicates led to 124 papers which were shortlisted for initial scrutiny. Based upon titles of the papers, 65 papers were included in the first round for review.

In the second stage, the abstracts of the selected papers were reviewed based upon topics such as geographical precinct, transportation, road analysis, and road safety. All off topics papers were excluded at this stage. The published works were carefully chosen based on the subsequent standards: (1) the paper had employed GIS technique, (2) the applied techniques had some global impact towards spatial

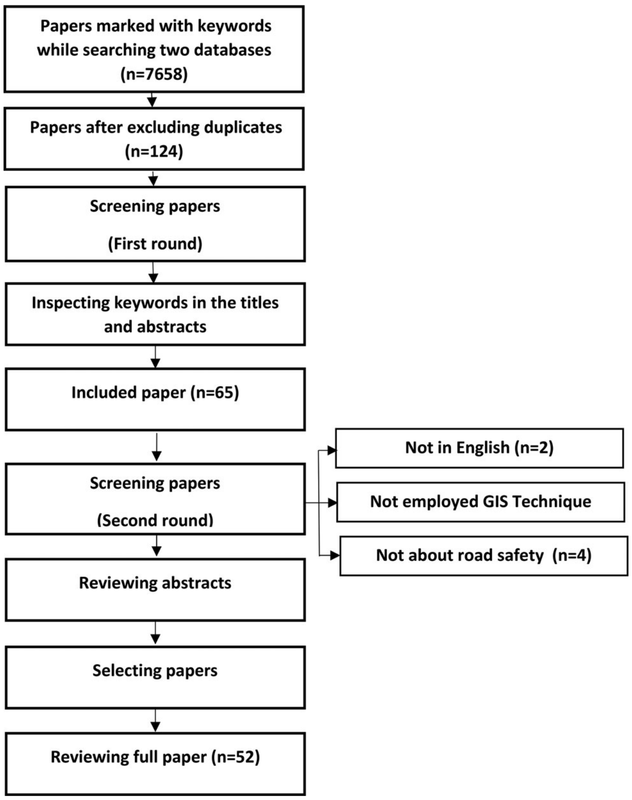

Figure 1. Flowchart of the study.

Table 2. Journals selected in the reviewed literature.

Journal	Abbreviation	Quartile
Safety Science	SS	Q1
Accident Analysis and Prevention	AAP	Q1
Journal of Transport Geography	JTG	Q1
Traffic Injury Prevention	TIP	Q1
Transport Policy	TP	Q1
Transportation Research Part A: Policy and Practice	TR	Q1
International Journal of Geographical Information Science	IJGIS	Q1
Journal of Public Transportation	JPT	Q2
Journal of Geographical Systems	JGS	Q2
Journal of Safety Research	JSR	Q1
KSCE Journal of Civil Engineering	KSCEJCE	Q2
International Journal of Civil Engineering	IJCE	Q2
Journal of Transportation Engineering	JTE	Q2
Geographical Analysis	GA	Q1
Spatial Information Research	SIR	Q1
Applied Spatial Analysis and Policy	ASAP	Q1
Transactions of the Institute of British Geographers	TIBG	Q1
Environment and Planning B: Planning and Design	EP	Q1
Statistical Methods and Applications	SMA	Q2

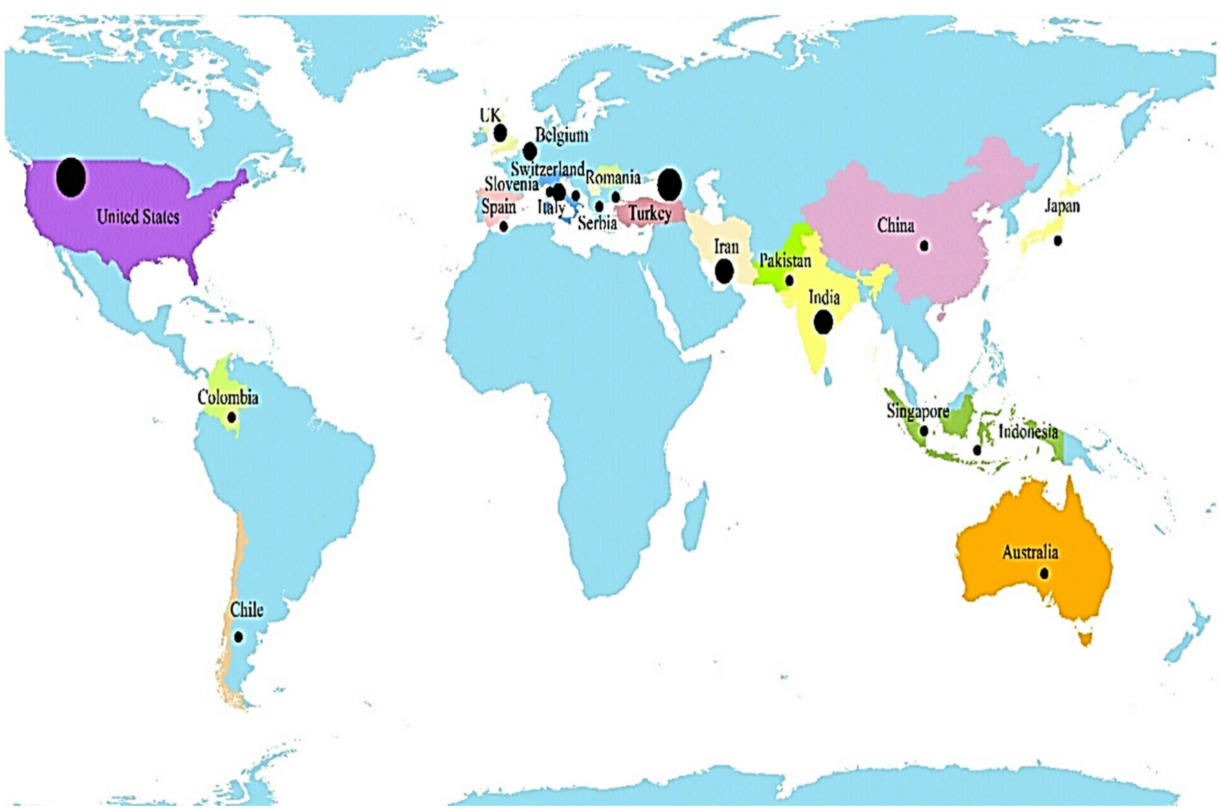

Figure 2. Dispersion of publications by country.

planning, (3) the paper was accessible on internet and (4) the language of paper was English. Next to this screening stage, 52 papers were finally acknowledged for scrutiny in total.

Furthermore, a summary list is constructed based on the papers publishing period, publishing country, numerical method, type of crashes and main findings. Afterward, 52 papers were fully reviewed in total to recognize the employed GIS technique on roads and its perspective on spatial planning. The dispersion of countries where the studies were directed as seen in Figure 2 and are reviewed in this paper.

4. Geospatial assessment techniques

The geographical study usually presents the usage of GIS spatial techniques in identifying road traffic accidents. This spatial technique specifies the crash location geographically and evaluates the map visualization through distribution patterns. The most extensively used tools of GIS reviewed in the study are specified subsequently:

4.1. Kernel density

Kernel density is a typical method used to analyze the spatial data, specifically, in ArcGIS. According to Budiharto and Saido (2012), this method can help professionals to easily measure the accidents that occurred in hazardous extents. Spreading jeopardy means the zone around the clump, where the probability of accidents is much higher. This Kernel density technique is helpful in creating quickest and schuss outcome. Erdogan et al. (2008) determined dark spots on the mainstream roads using a numerical technique (Poisson & Kernel density) and applied management scheme of GIS for crash investigation.

4.2. NetKDE

NetKDE is a mathematical approach, similar to KDE, used to measure density along assigned distances (Produit et al., 2010). It is broadly used by world researchers for spatial flattening (Anselin et al., 2000; Gatrell et al., 1996; Porta et al., 2009). In Ljubljana and Barcelona, scientists studied bicycle application demeanor and commercial events geographical dispersal (Lachance et al., 2011; Produit et al., 2010).

4.3. Getis-Ord

G-Ord has some alluring features that force the users to quantify geographically dispersed variables following Moran's *I*. This technique expands the geographical necessity progression and processed the complicated statistics which is impossible through screening contingency phenomena (Getis & Ord, 1992). This technique is primarily used to indicate the hot-spots of traffic accidents of any type. Getis-Ord measurement shows higher and lower values depending upon the numerical clumping index. Moreover, Getis-Ord statistics heavily assesses the data record of every single hot spot.

4.4. Moran's *I* statistic

It is a special technique used to analyze the geographical data of accidents. Moran's *I* technique estimates the meditation rate and also examine the spatial data when it is present in the scattered profile. Erdogan (2009) checked the altitudinal pattern of clumping data by only providing a single stretch of longitudinal connection. Astringency index of binary points can be well-defined as the variance between individual value and the total. Pirdavani et al. (2014) built accident speculation models whose prime purpose was to drive the Moran's *I* conditional and unconditional variables. The outcome showed the significance of accident speculation models while making altitudinal connections.

5. Results

5.1. Findings from reviewed studies

GIS performed numerous high-risk accidents on different spot points. Thus far, hot spots have been examined using alinear analysis approach with the help of accident statistics. But it is not the only standard to analyze the accidents, besides, mortalities, injuries record, and degree of damage have been taken into account for this study. The study used 6 different routes for hotspots identification. From results, hotspots to be found at different courses of the study area will make further developments towards route sections (Gundogdu, 2010).

A study conducted by Tasic et al. (2017), has processed the data using the 2D function of a geographical correlation model with a prime purpose of analyzing accidents. Three variables were taken into account data for data analysis involving tandem, pedestrian, and vehicular data. Results were notable as per the tandem-pedestrian accidents rate was far lesser than the vehicular accidents, providing advanced safety level.

Moran's *I* geographical analysis showed pedestrian-vehicle clatter numbers with substantial clumping patterns for the municipal area. Similarly, the pedestrian and vehicle accidents dark spots are also been speckled by the Getis-Ord statistic. Most of the dark spots were spotted adjacent to the crossings of dominion road and main road. Site assessments indicate

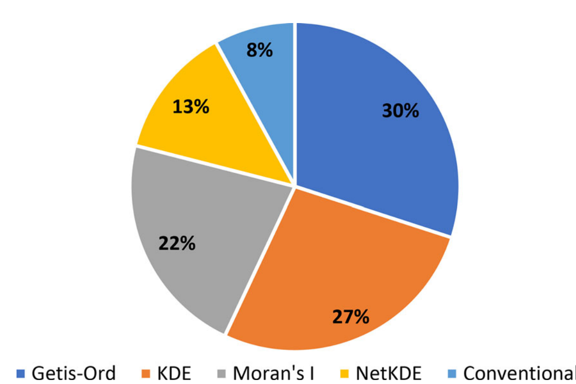

Figure 3. Distribution of techniques adopted in the articles reviewed.

that general road services are not provided for instance, traffic signals are offered at main road and crossing of dominion road, hence, triggering accidents. Getis-Ord, in results, showed considerable competency and constituency in finding the insecure bus stop status and dark spots of vehicle with pedestrian accidents (Truong & Somenahalli, 2011).

Rankavat and Tiwari (2013) analyzed pedestrian accidents arisen in Delhi through GIS database. In road accidents, the deaths that occurred for a period of three years were recorded. In addition, accident data formerly interpreted by making cardinal maps via GIS. Results indicated five pedestrian blackspots positions which were expressively clumped near crossings.

5.2. Data analysis

In the current literature, research papers were broadly reviewed in order to acquire precise results. The GIS techniques used by researchers in their studies are mainly Getis-Ord, Moran's *I*, Kernel density (KD) and NetKDE. However, results displayed in Figure 3 and presented in Table 2 revealed that the adaptation frequency of Getis-Ord and KD techniques stands 30% and 27% respectively, while it is 22% for Moran's *I*, 13% for conventional techniques and this adaptation rate stayed lowest (8%) for Network NetKDE. Similarly, the flowchart in Figure 4, demonstrates the distribution of spatial methods from reviewed papers (52). NetKDE was utilized 4 times, whereas Kernel density and Getis-Ord was employed 14 times and16 times, respectively in reviewed research, which turned out to be 30 in total.

Hence, it shows that these two techniques Getis-Ord and KD technique are more favorite among GIS users, and should be jointly applied. Moreover, data presented in Figure 5 incarnates the publications record from the year 1992 to 2018. The researches carried out in this review were highest for year 2010 and 2015, while no studies were undertaken for year 1993–1994, 1997–1999 and 2001–2002.

In summation, this review study on accident analysis via GIS, has mainly focused on the results of worldwide research published from 20+ years. In compliance with the overhead findings and reviewed studies, we propose a more comprehensive technique for critically analyzing road crashes. The proposed technique is unique because it determines the highly populated crash zones, distinguish crash variables, and will identify the maximum number of dark spots at the same time, therefore, would bring further improvements in the existing geospatial methods.

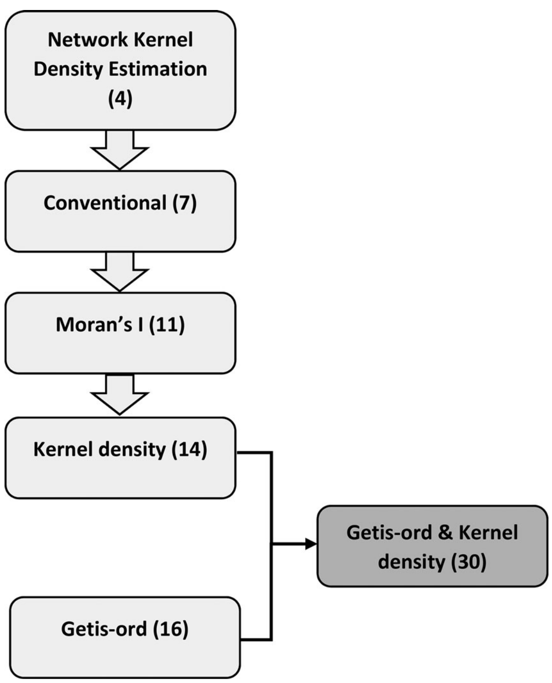

Figure 4. Distribution summary of spatial methods carried out in reviewed studies.

6. Discussion

The investigation of the literature that was executed in the current study has provided some notable findings regarding geographical or spatial applications on-road crashes. It seems that several multidimensional techniques and methods have been implemented, with a more focus on Kernel density and Getis-Ord. This provides the premium tools usually not available for accident statistical examination with more precision. Nevertheless, these methods have shown dearth in making a comparison of two diverse

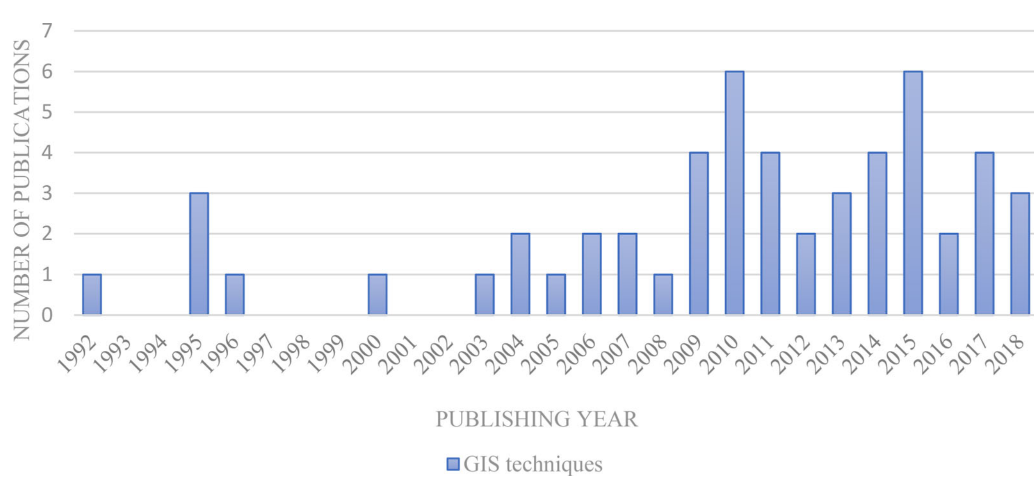

Figure 5. Number of Publications per annum from period 1992 to 2018.

Table 3. Results comparison from reviewed studies for accident analysis.

Author	Year	Country of study	Results/key findings
Hashimoto	2016	Japan	This research evaluated relations between city characteristics and road crashes. An estimation model of KDE was used in relationship with 16 other models. Results revealed much higher coefficients between actual and estimated number. This model can help in locating relative hazardous areas in cities.
Blazquez	2013	Chile	This study indicated 7 most critical zones of child-pedestrian crashes using Kernel density and Moran's *I* methods. Results confirm that the factors; careless crossing of pedestrian, pedestrian's impudence, pedestrians and drivers violates crosswalks, contributes 81.31% of crashes that occurred during a span of 8 years (2000–2008).
Truong and Somenahalli	2011	Australia	This study involves the identification of hotspots of pedestrians-vehicles using contiguous and severity pattern of Getis statistics and then Moran's *I* specified a quite significant clumping pattern of pedestrian-vehicle accident data. From results, both Moran's and Getis techniques proved to be competent and consistent in locating vehicle-pedestrian accident's black spots
Produit et al.	2010	Spain	The outcomes of the research showed the potential of NetKDE by presenting the economic schedules along road mesh. Using ArcGIS, Scientists successfully predicted the density of social activities in complex restraint environment.
Erdogan et al.	2008	Turkey	This study intended to apply GIS as a management tool and to locate crashes hotspots. Furthermore, Repeatability and Kernel Density analysis method was utilized. It is found that repeatability method of analysis measured more hotspots as compared to KDE method. For example, Repeatability showed 12 hotspots at Ayonkarshisar station, while KDE showed 9 hotspots at the same station. This is the way forward towards highway safety measures by GIS.
Pulugurtha et al.	2007	USA	This research used crash score, KDE and sum of rank method and indicated 29 pedestrian crash areas in Los Angeles city of the United States. The results were coherent with slight to no alteration when crash score and sum of rank methods were applied for rankings of pedestrian crash areas.

outcomes. Furthermore, the outmoded model like generalized linear model is not worthy of the given findings. These conventional models look forthright in result evaluation and interpolation. But the suggested methods performed well in interpolation mechanism of road crashes.

Moreover, the position of the hotspot is a further result of geographical or spatial research. Scientists found that the hotspots are not always coherent but exhibited considerable distinction regarding the vehicular type, and several new features applied during analysis tremendously alter the cardinal plan. Certainly, the methods being implemented on specific research will disturb the investigation result. Scientists must be cautious enough to alter unnoticed factors into noticed ones, to secure accurate and significant hotspot plan.

The GIS technique captures the researcher's attention owing to its advance data-analysis abilities. In the first, it provides a precise and careful data collection. Likewise, this technique also performs prior and subsequent-processing of results specified in spatial analysis (Dereli & Erdogan, 2017; Gundogdu, 2010).

Frequently, the geographical study is mostly performed by rejuvenating nations (like the United States and China), although not often seen in developing nations. Multi-power tools of GIS have been broadly analyzed; whose findings can be seen in Table 3. Each GIS tool demarcated its demerits and demerits. Numerous researchers used parameters such as vehicle speed, travel distance, and road size, and placed these parameters in assessment model in shape of variables. The model portrayed useful output when the data was not readily available.

The geographical analysis of road accidents indicates upsurge mortality proportion in less developed areas as compared to well-established areas of the city (Tortum & Atalay, 2015).

Travel distance and travel speed are the two essential factors but ignored in the past. Furthermore, some geometrical aspects of the road including road width, road slope, and road arc are not widely used, however, data inaccessibility for different study precincts may be one of the reasons for this reduction. Adding statistical data, for example, road accident data, geographic data, and population density might find scientists' interest rather than road geometrical data, hence, securing less contemplation.

Furthermore, data accessibility and reliable data within restricted time have always been a barrier while carrying out the geographical examination. However, the geographical examination has successfully analyzed various problems that a researcher can face. In addition, there is a need to dig out the variables that have not been analyzed before due to data inaccessibility such as road characteristics or profile, etc.

6.1. Limitations and future research directions

Regarding limitations, this study detected the limited techniques functioned in recent literature. In literature, 52 studies were comprehensively reviewed between 1992 and 2018.

The database search was limited to journal papers that were available in English and are peer-reviewed.

Moreover, GPS is a well-equipped device to find out exact crash locations (Bíl et al., 2013). Nonetheless, other distinguished models such as deep learning approach might be more constructive in future for a certain variety of road accident variables within explicit nature of framework. As the Support Vector Algorithms model in deep learning approach is effectively used by (Effati et al., 2015) in inspecting accident severity index, however, turned out to be a replacement of conventional regression techniques. The machine learning approach is rarely used by researchers in their studies. The concept of machine learning because of its advance programming ability can provide fruitful results with more precision. The future research direction of this study is to focus more on machine learning in road safety researches.

7. Conclusions

The aim of this review article was to study the GIS techniques valuable to ensure road safety. The study focuses on the geospatial mechanism of crashes by making comparisons of different techniques in order to know the practical applications of the study. Furthermore, 52 papers have been reviewed in relevance to road accidents to attain study aims. Most of the primary studies applied Moran's *I* and Getis-Ord to acquire precise results, but no study applied Getis-Ord and KDE methods jointly. From results, the adaptation frequency of Getis-Ord and KDE were appeared to be 30% and 27%, respectively, while it is 22% for Moran's *I*, 13% for conventional techniques and this adaptation rate stayed lowest (8%) for NetKDE. In conclusion, this study highly recommends applying Getis-Ord and KDE techniques in combination, because this proposed mechanism will not only determine the highly populated crash zones, distinguish new crash variables but will also identify the maximum number of hotspots in faster and easier mode. This new study will allow GIS professionals and government agencies of different countries of the world to practically implement this method in their domestic areas.

Disclosure statement

No potential conflict of interest was reported by the author.

ORCID

Monib Shahzad http://orcid.org/0000-0003-3190-0709

References

Anderson, T. (2009). Kernel density estimation and K-means clustering to profile road accident hotspots. *Accident; Analysis and Prevention*, *41* (3), 359–364. https://doi.org/10.1016/j.aap.2008.12.014

Anselin, L., Cohen, J., Cook, D., Gorr, W., & Tita, G. (2000). Spatial analyses of crime. *Criminal Justice*, *4*, 213–262.

Bailey, T. C., & Gatrell, A. C. (1995). *Interactive spatial data analysis*. Longman Scientific & Technical.

Benedek, J., Ciobanu, S. M., & Man, T. C. (2016). Hotspots and social background of urban traffic crashes: A case study in Cluj-Napoca (Romania). *Accident; Analysis and Prevention*, *87*, 117–126. https://doi.org/10.1016/j.aap.2015.11.026

Bíl, M., Andrašik, R., & Janoška, Z. (2013). Identification of hazardous road locations of traffic accidents by means of kernel density estimation and cluster significance evaluation. *Accident; Analysis and Prevention*, *55*, 265–273.

Bíl, M., Andrášik, R., & Sedoník, J. (2019). A detailed spatiotemporal analysis of traffic crash hotspots. *Applied Geography*, *107*, 82–90. https://doi.org/10.1016/j.apgeog.2019.04.008

Blazquez, C. A., & Celis, M. S. (2013). A spatial and temporal analysis of child pedestrian crashes in Santiago, Chile. *Accident; Analysis and Prevention*, *50*, 304–311. https://doi.org/10.1016/j.aap.2012.05.001

Budiharto, U., & Saido, A. P. (2012). Traffic accident blackspot identification and ambulance fastest route mobilization process for the city of Surakarta. *Jurnal Transportasi*, *12*(3), 237–248. https://doi.org/10.26593/jt.v12i3.508.%25p

Cafiso, S., Di Graziano, A., Di Silvestro, G., La Cava, G., & Persaud, B. (2010). Development of comprehensive accident models for two-lane rural highways using exposure, geometry, consistency and context variables. *Accident; Analysis and Prevention*, *42*(4), 1072–1079. https://doi.org/10.1016/j.aap.2009.12.015

Chen, X., Huang, L., Dai, D., Zhu, M., & Jin, K. (2018). Hotspots of road traffic crashes in a redeveloping area of Shanghai. *International Journal of Injury Control and Safety Promotion*, *25*(3), 293–302. https://doi.org/10.1080/17457300.2018.1431938

Colak, H. E., Memisoglu, T., Erbas, Y. S., & Bediroglu, S. (2018). Hot spot analysis based on network spatial weights to determine spatial statistics of traffic accidents in Rize. *Arabian Journal of Geosciences*, *11*(7), 151. https://doi.org/10.1007/s12517-018-3492-8

Cordera, R., Sañudo, R., dell'Olio, L., & Ibeas, Á. (2018). Trip dstribution model for regional railway servies considering spatial effects between stations. *Transport Policy*, *67*, 77–84. https://doi.org/10.1016/j.tranpol.2018.01.016

Dereli, M. A., & Erdogan, S. (2017). A new model for determining the traffic accident black spots using GIS-aided spatial statistical methods. *Transportation Research Part A: Policy and Practice*, *103*, 106–117. https://doi.org/10.1016/j.tra.2017.05.031

Effati, M., Rajabi, M., Samadzadegan, F., & Shabani, S. (2014). A geospatial neuro fuzzy approach for identification of hazardous zones in regional transportation corridors. *International Journal of Civil Engineering*, *12*(3), 289–303.

Effati, M., Thill, J. C., & Shabani, S. (2015). Geospatial and machine learning techniques for wicked social science problems: Analysis of crash severity on a regional highway corridor. *Journal of Geographical Systems*, *17*(2), 107–135. https://doi.org/10.1007/s10109-015-0210-x

Erdogan, S. (2009). Explorative spatial analysis of traffic accident statistics and road mortality among the provinces of Turkey. *Journal of Safety Research*, *40*(5), 341–351. https://doi.org/10.1016/j.jsr.2009.07.006

Erdogan, S., Ilci, V., Soysal, O. M., & Kormaz, A. (2015). A model suggestion for the determination of the traffic accident hot spots on the Turkish highway road network: A pilot study. *Boletim de Ciências Geodésicas*, *21* (1), 169–188. https://doi.org/10.1590/S1982-21702015000100011

Erdogan, S., Yilmaz, I., Baybura, T., & Gullu, M. (2008). Geographical information systems aided traffic accident analysis system case study: City of Afyonkarahisar. *Accident; Analysis and Prevention*, *40*(1), 174–181. https://doi.org/10.1016/j.aap.2007.05.004

Flahaut, B., Mouchart, M., San Martin, E., & Thomas, I. (2003). The local spatial autocorrelation and the kernel method for identifying black zones: A comparative approach. *Accident; Analysis and Prevention*, *35* (6), 991–1004. https://doi.org/10.1016/S0001-4575(02)00107-0

Fox, L., Serre, M. L., Lippmann, S. J., Rodríguez, D. A., Bangdiwala, S. I., Gutiérrez, M. I., Escobar, G., & Villaveces, A. (2015). Spatiotemporal approaches to analyzing pedestrian fatalities: The case of Cali, Colombia. *Traffic Injury Prevention*, *16*(6), 571–577. https://doi.org/10.1080/15389588.2014.976336

Gatrell, A. C., Bailey, T. C., Diggle, P. J., & Rowlingson, B. S. (1996). Spatial point pattern analysis and its application in geographical epidemiology. *Transactions of the Institute of British Geographers*, *21*(1), 256–274. https://doi.org/10.2307/622936

Getis, A., & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis*, *24*(3), 189–206. https://doi.org/10.1111/j.1538-4632.1992.tb00261.x

Gundogdu, I. B. (2010). Applying linear analysis methods to GIS-supported procedures for preventing traffic accidents: Case study of Konya. *Safety Science*, *48*(6), 763–769. https://doi.org/10.1016/j.ssci.2010.02.016

Gundogdu, I. B. (2011). A new approach for GIS-supported mapping of traffic accidents. *Proceedings of the Institution of Civil Engineers - Transport*, *164*(2), 87–96. https://doi.org/10.1680/tran.2011.164.2.87

Hashimoto, S., Yoshiki, S., Saeki, R., Mimura, Y., Ando, R., & Nanba, S. (2016). Development and application of traffic accident density estimation models using kernel density estimation. *Journal of Traffic and Transportation Engineering (English Edition)*, *3*(3), 262–270. https://doi.org/10.1016/j.jtte.2016.01.005

Huang, H., & Chin, H. (2010). Modeling road traffic crashes with zero-inflation and site-specific random effects. *Statistical Methods & Applications*, *19*(3), 445–462. https://doi.org/10.1007/s10260-010-0136-x

Kibar, F. T., & Yaman, H. T. (2020). GIS-based evaluation of the speed limit increase on urban arterial traffic safety in Ankara. *Arabian Journal of Geosciences*, *13*, 476. https://doi.org/10.1007/s12517-020-054865

Kazmi, J. H., & Zubair, S. (2014). Estimation of vehicle damage cost involved in road traffic accidents in Karachi, Pakistan: A geospatial perspective. *Procedia Engineering*, *77*, 70–78. https://doi.org/10.1016/j.proeng.2014.07.008

Kuo, P. F., Zeng, X., Lord, D. (2012). *Guidelines for choosing hot-spot analysis tools based on data characteristics, network restrictions, and time distributions*. In Proceedings of the Ninety-first Annual Meeting of the Transportation Research Board, pp. 22–26.

Lachance, B. N., Produit, T., Tominc, B., Matej, N., & Golicnik Marusic, B. (2011). Network based kernel density estimation for cycling facilities optimal location applied to Ljubljana. *ICCSA, Part II, LNCS*, *6783*, 136–150.

Le, K. G., Liu, P., & Lin, L. T. (2020). Determining the road traffic accident hotspots using GIS-based temporal-spatial statistical analytic techniques in Hanoi, Vietnam. *Geo-Spatial Information Science*, *23*(2), 153–164. https://doi.org/10.1080/10095020.2019.1683437

Levine, N., Kim, K. E., & Nitz, L. H. (1995). Spatial analysis of Honolulu motor vehicle crashes: I. Spatial patterns. *Accident Analysis & Prevention*, *27*(5), 663–674. https://doi.org/10.1016/0001-4575(95)00017-T

Li, L., Zhu, L., & Sui, D. Z. (2007). A GIS-based Bayesian approach for analyzing spatial–temporal patterns of intra-city motor vehicle crashes. *Journal of Transport Geography*, *15*(4), 274–285. https://doi.org/10.1016/j.jtrangeo.2006.08.005

Loo, B. P. (2006). Validating crash locations for quantitative spatial analysis: A GIS-based approach. *Accident; Analysis and Prevention*, *38*(5), 879–886. https://doi.org/10.1016/j.aap.2006.02.012

Ma, L., Yan, X., & Qiao, W. (2014). A quasi-Poisson approach on modeling accident hazard index for urban road segments. *Discrete Dynamics in Nature and Society, New York, NY*, *2014*, 1–8. https://doi.org/10.1155/2014/489052

Moore, D. N., Schneider, W. H., Savolainen, P. T., & Farzaneh, M. (2011). Mixed logit analysis of bicyclist injury severity resulting from motor vehicle crashes at intersection and non-intersection locations. *Accident; Analysis and Prevention*, *43*(3), 621–630. https://doi.org/10.1016/j.aap.2010.09.015

Mohan, D. (2009). Road accidents in India. *IATSS Research*, *33*(1), 75–79. https://doi.org/10.1016/S0386-1112(14)60239-9

Nanda, S., & Singh, S. (2018). Evaluation of factors responsible for road accidents in India by fuzzy AHP. In *Networking communication and data knowledge engineering*. Lecture Notes on Data Engineering and Communiction Technology (pp. 179–188). Springer. https://doi.org/10.1007/978-981-10-4585-1_15

Niloofar, H. M. A. (2017). Introducing GIS as legitimate instrument to deal with road accident data: A case study of Iran, Tehran. *Spatial Information Research*, *25*(1), 151–159. https://doi.org/10.1007/s41324-017-0083-9

Pirdavani, A., Bellemans, T., Brijs, T., & Wets, G. (2014). Application of geographically weighted regression technique in spatial analysis of fatal and injury crashes. *Journal of Transportation Engineering*, *140*(8), 04014032. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000680

Porta, S., Latora, V., Wang, F., Strano, E., Cardillo, A., Iacoviello, V., Messora, R., & Scellato, S. (2009). Street centrality and densities of retails and services in Bologna. *Environment and Planning B: Planning and Design*, *36*(3), 450–465. https://doi.org/10.1068/b34098

Produit, T., Lachance-Bernard, N., Strano, E., Porta, S., & Joost, S. (2010). A network based Kernel density estimator applied to Barcelona economic activities. *Lecture Notes in Computer Science*, *6016*, 32–45. https://doi.org/10.1007/978-3-642-12156-2_3

Pulugurtha, S. S., Krishnakumar, V. K., & Nambisan, S. S. (2007). New methods to identify and rank high pedestrian crash zones: An illustration. *Accident; Analysis and Prevention*, *39* (4), 800–811. https://doi.org/10.1016/j.aap.2006.12.001

Rankavat, S., & Tiwari, G. (2013). Pedestrian accident analysis in Delhi using GIS. *Journal of the Eastern Asia Society for Transportation Studies*, *10*, 1446–1457. https://doi.org/10.11175/easts.10.1446

Shariat, M. A., & Shahri, M. (2017). Crash prediction modeling using a spatial semi-local model: A case study of Mashhad. *Applied Spatial Analysis and Policy*, *10*(4), 565–584. https://doi.org/10.1007/s12061-016-9199-x

Shankar, V., Mannering, F., & Barfield, W. (1995). Effect of roadway geometrics and environmental factors on rural freeway accident frequencies. *Accident Analysis & Prevention*, *27*(3), 371–389. https://doi.org/10.1016/0001-4575(94)00078-Z

Steenberghen, T., Dufays, T., Thomas, I., & Flahaut, B. (2004). Intra-urban location and clustering of road accidents using GIS: A Belgian example. *International Journal of Geographical Information Science*, *18*(2), 169–181. https://doi.org/10.1080/13658810310001629619

Tasic, I., Elvik, R., & Brewer, S. (2017). Exploring the safety in numbers effect for vulnerable road users on a macroscopic scale. *Accident; Analysis and Prevention*, *109*, 36–46. https://doi.org/10.1016/j.aap.2017.07.029

Tortum, A., Atalay, A. (2015). Spatial analysis of road mortality rates in Turkey. Proceedings of the Institution of Civil Engineers-Transport, *168*(6), 532–542. https://doi.org/10.1680/jtran.14.00029

Truong, L. T., & Somenahalli, S. V. (2011). Using GIS to identify pedestrian-vehicle crash hot spots and unsafe bus stops. *Journal of Public Transportation*, *14*(1), 99–114. https://doi.org/10.5038/2375-0901.14.1.6

World Health Organization (WHO). (2007). *Road traffic crashes leading cause of deaths among young people*. WHO Press.

World Health Organization (WHO). (2018). *Global status report on road safety*. WHO Press.

Yalcin, G., & Duzgun, H. S. (2015). Spatial analysis of two-wheeled vehicles traffic crashes: Osmaniye in Turkey. *KSCE Journal of Civil Engineering*, *19*(7), 2225–2232. https://doi.org/10.1007/s12205-015-0661-0

Yamada, I., & Thill, J. C. (2004). Comparison of planar and network K-functions in traffic accident analysis. *Journal of Transport Geography*, *12*(2), 149–158. https://doi.org/10.1016/j.jtrangeo.2003.10.006

Yuan, J., Zhang, B., & Lin, F. (2005). Graph partition model for robust temporal data segmentation. *Knowledge Discovery. Data Mining: 9th Pacific-Asia Conference*, *3518*, 758–763.

Zhang, Y. (2010). *Hotspot analysis of highway accident spatial pattern based on network spatial weights*. A&M University.