Proceedings of 15th IOE Graduate Conference
Peer Reviewed
Year: 2024 Month: May Volume: 15
ISSN: 2350-8914 (Online), 2350-8906 (Print)

Spatial Analysis of Road Traffic Crash Hotspots in Kathmandu Valley, Nepal

Anuradha K.C [a], Rojee Pradhananga [b], Upama Koju [c]

[a, b] *Department of Civil Engineering, Pulchowk Campus, IOE, Tribhuvan University, Nepal*
[c] *Senior Researcher, Forest Action Nepal*
[a] anuradhakc8@gmail.com, [b] rojeepradhananga@gmail.com, [c] upamakoju@gmail.com

Abstract

Kathmandu Valley is one of the rapidly urbanizing cities in Nepal, which registers the highest incidence of road traffic crashes compared to other regions in the country. The current paper aims to demonstrate the applicability of Geographic Information System (GIS) oriented spatial analysis techniques to characterize the spatial patterns of the crash hotspots in the Valley. Within this study, a fusion of spatial autocorrelation (Global Moran's I Index), Getis-Ord Gi* statistic and Kernel Density Estimation (KDE) methods were applied to analyze the three years of crash data (2019-2021) on aggregate basis. A positive value (0.11) of Moran's I Index indicated that the crashes were clustered spatially with high significance. The Getis-Ord Gi* measure identified the highly significant hotspot segments along the Valley roads at 99 % confidence level. The KDE visualization corroborated the hotspot links generated by the Getis-Ord Gi* statistics. The hotpots were further ranked based on weightages relative to severity level. The results showed that the hotspots are concentrated along Ring Road, Tribhuwan Highway, Araniko Highway and feeder roads. The findings of this study will serve useful evidence for traffic safety management agencies to prioritize these specific locations and implement targeted interventions to enhance safety of these road segments.

Keywords

Traffic Crashes, GIS, Spatial Analysis, Hotspots, Moran's I, Getis-Ord Gi*, KDE, Severity

1. Introduction

1.1 Background

Road crashes are first and foremost the preeminent consensus of economies worldwide, driven by the rising number of vehicles and the demand for efficient transportation. The road crashes result in fatalities and injuries affecting 1.19 million and 20-50 million people each year respectively, leading to great socio-economic ordeals at national, community and individual levels. The safety risks are three times higher in low-income countries than the high-income countries where there are 9 deaths per 100,000 people on average [1].

Nepal, like other low-income countries, grapples with traffic safety issues on highways and urban roads with a record of 2883 casualties, 7282 serious injuries, and 25,722 minor injuries in fiscal year 2020/21. Especially, Kathmandu Valley as the primary economic hub in Nepal, is posed with significant crash risks due to its immense population of 2.88 million residents [2] and 0.44 million registered automobiles [3]. According to traffic police records, the Valley witnesses about 10,000 crashes every year causing more than 180 deaths. The elevated crash incidences can be attributed to a number of factors, including an increase in the number of vehicles, insufficient infrastructure capacity, disregarded safety measures, neglected traffic segregation, haphazard roadside parking and weak law enforcement [4]. Although several safety initiatives have been put in place to monitor and reduce the traffic crashes, the problem continues to grow due to insufficient investigation of root causes and a lack of prioritization strategy for the hot spots.

In the context of Kathmandu Valley, the definition of crash hotspots remains elusive due to absence of standardized framework in Nepal. Consequently, local researchers [5, 6] rely upon traditional statistical analysis, site observation and stakeholder interaction techniques to analyze the trends of hotspots just as a general estimate. However, these methods have limited ability to capture all the complex factors contributing to frequencies and severity of crashes at high-risk road locations. The primary constraint lies in the availability of data relevant to crash, road infrastructure and environment, which hinders comprehensive understanding of the factors. Hence, longitudinal studies are desired to identify the patterns of hotspots, their severity and causative factors utilizing more rigorous techniques. The present study endeavors to address this gap by employing GIS based spatial analysis techniques in order to highlight the significant hotspot locations and their patterns across Valley's Road network. Furthermore, the study seeks to rank the identified hotspots based on severity index, and outline probable causes to assist policymakers in acknowledging the potential safety measures.

2. Literature Review

GIS aided spatial data analysis and mapping techniques have been widely applied to combine crash data with spatially referenced variables in order to identify high-risk locations [7, 8, 9, 10]. The spatial analysis typically involves identifying spatial dependency between clusters of locations that share comparable attribute values [11]. Geospatial tools such as Moran's I Index (Global and Local), Getis-Ord Gi* statistics, Kernel Density Estimation (KDE), nearest neighbor distance, and K-mean clustering have been extensively employed in crash hotspots determination [11, 12]. The above methods are sequentially devised to account the shortcomings of previous methods. For example the nearest neighborhood method was replaced by K-function to delineate the clustering extent across wider range of scale. The nearest neighborhood function and K-function merely look at the general tendency of data, so KDE method was devised to localize the clusters. The KDE method involves constructing a smooth density surface within two-dimensional space to estimate aggregate occurrences within search bandwidth. Essentially, the surface value is largest at the point location (the center) and gently decreases to zero along the circle's radius (bandwidth) [13, 14]. The standard KDE method has specific advantages in visualization of location and density of crash points as it has effect of dispersing the risk of crashes. One unavoidable flaw is that it is unable to examine the statistical significance of the crash hotspots [14].

Whereas, the spatial autocorrelation viz. Moran's I Index and Getis-Ord Gi* statistics categorize the spatial patterns into three kinds; namely clustering, random and dispersive, the statistical significance of which can be evaluated using a z-score, a feature not present in KDE. The Global Moran's I offers several advantages, including its higher level of general stability, testability, adaptability of conditional distribution, and applicability in crash analysis[15]. It specifically evaluates the overall spatial arrangement of a variable without providing statistical inference for identifying particular clusters or hotspots. By complementing this global indicator with local indicators of spatial autocorrelation like local Moran's I and Getis-Ord Gi* statistics, the pockets of hotspots become more apparent. A plus point of the Gi* statistics is their ability to neutralize the spatial distribution of data points, thus facilitating the formulation of hypotheses without the risk of data point patterns biasing the outcomes [16].

The KDE, Moran's I Index and Getis-Ord Gi* statistics have been employed by several researchers to highlight the significant crash hotspots at macroscopic (e.g. state and geographic unit) and microscopic (e.g. road sections and intersections) scales [7, 10, 11, 12, 13, 17, 18, 19]. Prasannakumar et al. [7] examined the spatial distribution of crashes in Thiruvananthapuram city by dividing data into spatial parameters of educational and religious places. Wang et al. [10] analyzed the clustering of six collision types and three severity levels of crashes with optimized hotspot analysis method for Liangshan Yi Autonomous Prefecture, China. The results show that motor vehicle crashes dominate the hotspots in core and outer regions. The study further added that unreasonable threshold distances are culminated when only aggregate dataset is considered. This can jeopardize the significance of cluster identification, also known as the boundary effect. Nanzeen et al. [12] applied the KDE, Getis-Ord Gi* and crash severity map to identify the risky locations in the Indian Reservations. The study suggested that other factors such as traffic volume, speed, crash time and weather should be comprehended to accurately delineate the crash severity risks.

Similarly, Alam et al. [17] imposed these techniques in highways of Ohio and highlighted the importance of temporal trends to uncover the true crash patterns. Harirforoush et al. [11] combined the results of KDE representing crash density with Moran's I to identify significant clusters of traffic crashes in the Sherbrooke region across distinct seasons. The results indicate a notable concentration of traffic crash clusters in the city center and along major roadways during the summer and autumn seasons. Troung et al. [13] investigated the trends of pedestrian-vehicle crashes in Adelaide Metropolitan area using crash data spanning 13 years, and found that majority of pedestrian-vehicle crashes occur at intersection and more severe crashes at mid-blocks. Hammas et al. [18] utilized the KDE and Getis-Ord Gi* to map the hotspots in Medina on annual basis, and made deliberations that these increase proportionately with casual variables like high-speeds, road width and population density. Some extra variables were considered by Afolayan et al. [19] while investigating the traffic crash hotspots on Nigerian highway. The study presented the association of hotspot occurrences with traffic exposure and geometric features.

On the part of visualization, Manepalli et al. [20] compared the KDE and Getis-Ord Gi* (d) statistics and concluded that identical hotspots are outlined by both methods. In contrast, the Kuo [21] and Plug et al. [22] observed KDE maps to possess greater visual significance in comparison to those produced using Gi* z-scores.

3. Study Area

The designated study area is the Kathmandu Valley, a bowl-shaped portion of Bagmati watershed that covers an area of 671.32 sq.km in the central hilly region of Nepal. The study area comprises of 4 local levels (Bhaktapur, Changunarayan, Madhyapur Thimi and Suryabinayak) of Bhaktapur district, 11 local levels (Budhanilkantha, Chandragiri, Dakshinkali, Gokarneshwor, Kageshwori Manahora, Kathmandu, Kirtipur, Nagarjun, Shankharapur, Tarakeshwor and Tokha) of Kathmandu district and 3 local levels (Godawari, Lalitpur and Mahalaxmi) of Lalitpur district.

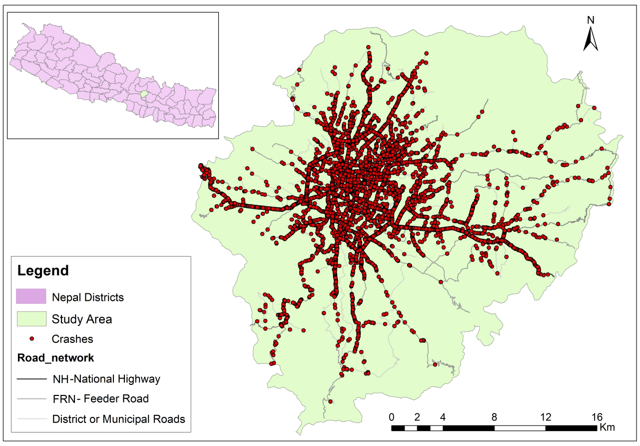

Figure 1: Location Map of Study Area with Road Network and Crash Distribution

As the most populated city in Nepal, the Valley is afflicted by the detrimental effects of haphazard urban expansion. The erratic development has led to an increase in vehicle usage without a corresponding rise in road infrastructures. Most of the highways, feeder and urban roads of the Valley suffer from daily congestion and traffic crashes in relation to heavy traffic flow. Moreover, the Valley roads are reported to have the highest incidences of traffic crashes than elsewhere in the country.

4. Methodology

The present research employed an integration of three spatial analysis methods; Global Moran's I Index, Getis-Ord Gi* statistics, and KDE to evaluate the statistically significant hotspots for aggregate crashes, and assigned the ranking of the identified hotspots based on a severity index. The summary of the research methodology is explained in the sub-sections beneath:

4.1 Data Collection

In the context of this study, the recent traffic crash records from 2019 to 2021 was secured from Nepal Police Headquarters, Naxal in electronic form. The Survey Department provided the road network and administrative dataset in GIS shape file format. The road dataset included attribute information such as ID, link name, start and end of section, road class (municipal/district/feeder/national highway), length and remarks. The overall quality of the road network was ascertained by comparing it with a separate dataset downloaded from OpenStreetMap (OSM) Project using the Geofabrik's server. Additional data on population size were downloaded from census 2021 portal of National Statistics Office.

4.2 Data Preparation and Processing

The crash records, initially in Word format and Nepalese language, was converted into CSV format using a Python program. The CSV file lacked precise coordinates for each crash points. To address this, each record was manually reviewed and distances measured from traffic police station to pinpoint corresponding crash locations in Google Earth. The fields of CSV dataset were formatted to include a unique key (I.D.), date, time, X and Y coordinates, severity level, causes, vehicle type and other variables. The co-ordinates were imported to ArcMap 10.8 to create a layer of point features with attribute information. Road network was added to geo-database and the crash points were adjusted to intersect with nearest edge of road network with snap tool. After removal of outlier points, 23,278 (79.55%) crash records were available for the analysis.

4.3 Spatial Auto correlation

Global Moran's I is one of the earliest markers of spatial auto correlation. It is used to examine whether the crash locations and attributes exhibit any regularity, i.e. whether they are spatially clustered, dispersed or random [7]. It is computed using equation (1):

$$EI = \frac{n}{\sum_{i=1}^{n}\sum_{i=1}^{n} w_{ij}} \times \frac{\sum_{i=1}^{n}\sum_{i=1}^{n} w_{ij}(x_i - X)(x_j - X)}{\sum_{i=1}^{n}(x_i - X)^2} \tag{1}$$

Where, n is total number of cases, w_{ij} is spatial weight between point i and j, x_i is crash count at point i, and X is global mean value [17]. The Z-score to test the null hypothesis of complete randomness of the features is given by equation (2)

$$Z = \frac{(I - E[I])}{\sqrt{V[I]}} \tag{2}$$

Where expected value and variance is represented as:

$$E[I] = \frac{-1}{(n-1)} \tag{3}$$

and

$$V[I] = E[I^2] - (E[I])^2 \tag{4}$$

The "Distance Band from Neighbor Count" geo-processing function generated an average distance to be used as the starting distance, and the increment for the spatial auto correlation analysis. The spatial pattern analysis tool of ArcGIS was run to generate the Moran's I, z-score and p-value using the average distance. The Moran's I value close to -1 denotes dispersed patterns, while the value close to +1 denotes clustered patterns in the analysis regions. The significance of the spatial auto correlation increases with the absolute magnitude of Moran's I [7].

The average distance was also input for the Incremental Spatial Auto correlation function that gave the threshold distance as the analysis scale. The peak distance (distance at which most of the crashes tend to cluster) respective to the maximum z-score and minimum number of neighbors required was selected.

4.4 Getis-Ord Gi* Statistics

The Gi*statistics analyzes each feature in the dataset within the context of neighboring features [16]. It classifies a feature as statistically significant hot spot when both the feature and its neighbors have high value [7, 12, 13]. The statistic is derived from equation (5):

$$G_i^*(d) = \frac{\sum_j \left(W_{ij}x_j - \bar{W}_{ij}\bar{x}\right)}{s^*\sqrt{2\left(\frac{nS_{1i}^* - \bar{W}_{ij}}{n-1}\right)}} \tag{5}$$

Where, $S*$ is the sum of squared weights, $s*$ is the standard deviation of the data, W_{ij} is a spatial weight between features i and j, $W * ij$ is the sum of weights and n is equal to total number of features [19], and:

$$s = \sqrt{\frac{\sum_{j=1}^{n} x_j}{n} - \bar{x}^2} \tag{6}$$

$$\bar{x} = \frac{\sum_{j=1}^{n} x_j}{n} \tag{7}$$

To perform this, the crash data was integrated with road network layer using a spatial join tool. A new joint count field was provided that quantified the number of crashes per segment. The Hot Spot Analysis function (Getis-Ord Gi*) from the Spatial Statistics toolbox of Arc Map 10.8 was fed with this field as the input for hot spot computations. Both the fixed distance band method and inverse distance method were applied for the conceptualization of spatial relationships. A threshold distance of 100 meters was found most suitable for the fixed distance strategy, whilst, the inverse distance method, regardless of threshold input or not, showed better performance than the fixed distance method. So, only the results of the inverse distance method are included in this paper.

The $G_i^*(d)$ statistics returned an output feature class containing z-score (Standard Deviations), p-value, and a bin field for confidence level (*Gi_Bin*) for each feature within the input feature class. Features falling within the ±3 bins accurately represent statistical significance with a confidence level of 99%; whereas those within the ±2 bins indicate a confidence level of 95%, and those within the ±1 bins indicate a confidence level of 90%. A high z-score and small p-value of a feature suggest a spatial clustering of high values, whereas a low negative z-score and small p-value indicate a spatial clustering of low values [23].

4.5 Kernel Density Estimation (KDE)

KDE technique was applied to estimate the density of crashes within a given bandwidth (search radius) surrounding each point across study area. It determines the magnitude per unit area from each hot spot feature; in other word the spread of risks around crash cluster. The mathematical form of the intensity at each location is show in equation (8):

$$f(x,y) = \frac{1}{nh^2}\sum_{i=1}^{n} k\left(\frac{d_i}{h}\right) \tag{8}$$

Where, $f(x,y)$ is the density measured at location (x,y), h is the radius of the circle (bandwidth or kernel size), $K()$ is the kernel which is a function of the bandwidth and distance, and d_i is the distance between point (x,y) and ith location [24].

With the help of the KDE calculator feature offered by the spatial analyst tool, kernel density hot spots were detected using the crash point layer and populated field as none. The primary factor to consider in determining the appropriate density level is the bandwidth and cell size. Through hit and trial method, a bandwidth of 250 m and a cell size of 10 m were found suitable and consistent.

4.6 Identification and Ranking of Hot spots

The statistically significant crash hot spots in the Valley were ascertained by overlapping the high-risk maps created by KDE and Getis-Ord Gi* (d) statistics. Further, severity weighted technique was utilized to pinpoint the hot spots and rank the top ten sites. The weights for each severity level were determined in proportionality to associated crash costs as suggested by Highway Safety Manual (HSM). The crash costs were excerpted from Rizal et al. [25], and the weights were calculated by dividing the crash costs corresponding to different levels of severity by the Property Damage Only (PDO) costs. The derived weights are; 353 for death (A), 8 for major injury (B), 2 for minor injury (C) and 1 for PDO crashes (D). To assign severity index (SI) to each segment, the equation (9) was used:

$$SI = 353A + 8B + 2C + D \tag{9}$$

5. Results and Discussions

For the spatial analysis, the definition of hotspot is taken as a collection of adjacent spatial units characterized by a significant occurrence of crashes. The primary spatial unit considered in this study was the original length of road segments, chosen to account for geometric and environmental changes. The analysis was oriented towards finding the spatial trends in both the aggregate crash data.

5.1 Descriptive Analysis

Figure 2 shows a similarity in the yearly trend of total, death, major injury, minor injury and PDO crashes from 2019-2021. The peak crashes for all types occurred in 2019, decreased in 2020 and again hiked in 2021. The observed reduction in crashes is due to the nationwide lockdown and vehicle travel restrictions imposed during the outbreak of COVID-19 pandemic.

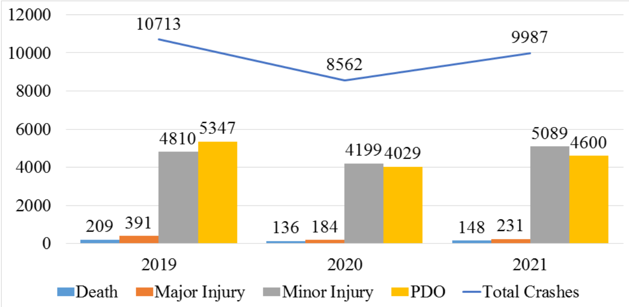

Figure 2: Yearly Distribution of Crashes (2019-2021)

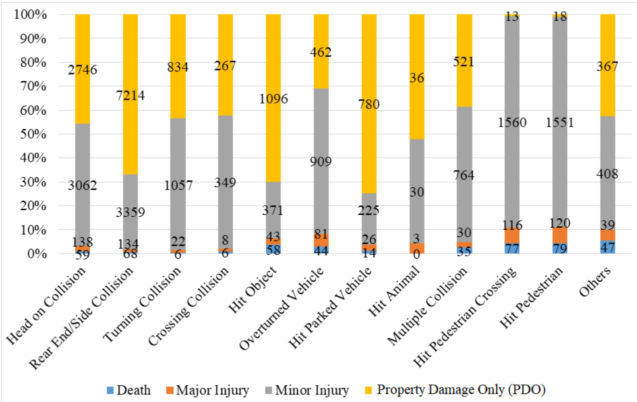

Figure 3: Collision Types and Resulting Severity

Figure 3 reveals that head on, rear end/side, overturned vehicle and hit pedestrian (crossing or walking) types of collision contribute to deaths and major injuries. The impact of collisions is particularly high when the vehicular speed is considerable as illustrated by Figure 4. Driver carelessness and driving under the influence of alcohol are also held accountable for the seriousness of the crash events.

The crash density pattern of the Valley exhibits a significant concentration towards central urban area in excess of 900 crash counts (Figure 5). The core region comprises of Kathmandu Metropolitan City, Lalitpur Metropolitan City, Madhyapur Thimi Municipality and Bhaktapur Municipality. These areas are distinguished by elevated levels of population

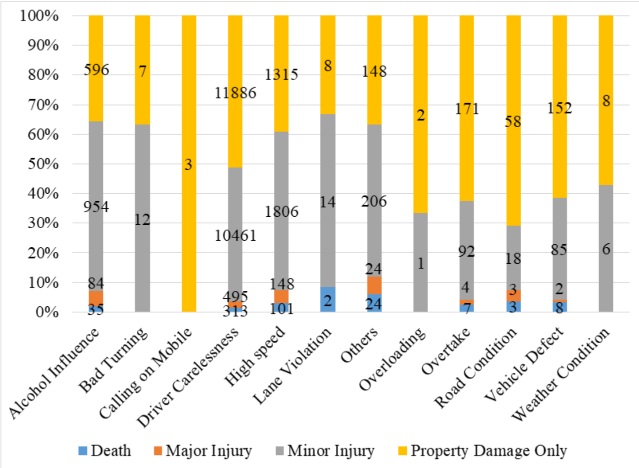

Figure 4: Main Reasons of Crashes and Resulting Severity

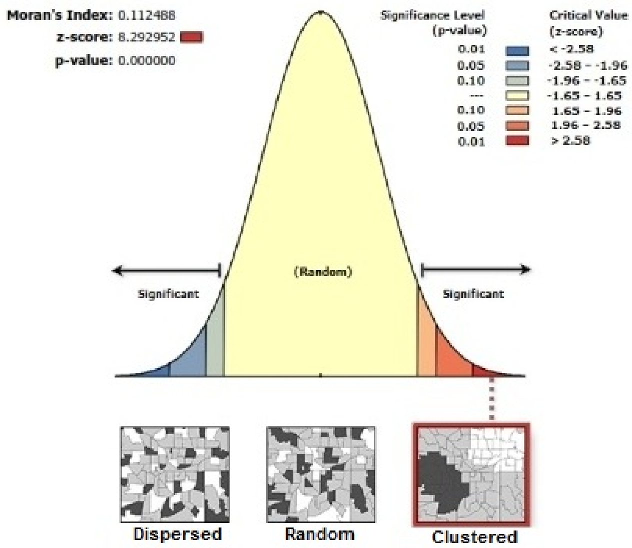

Figure 6: Spatial Auto correlation Results

density (over 5500 heads/sq.km), socio-economic engagement, extensive road infrastructure, and extreme traffic flows.

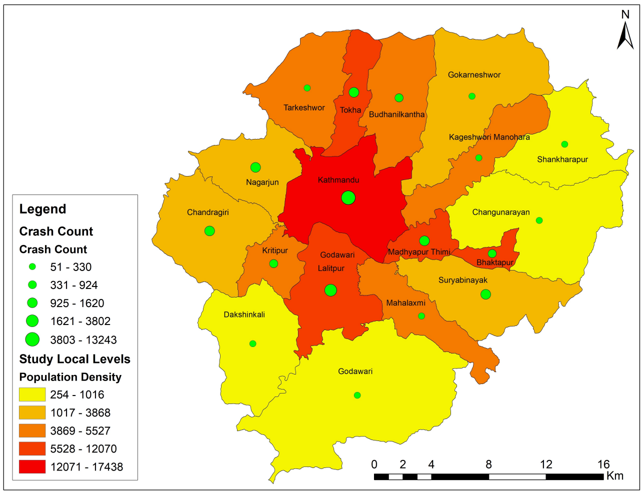

Figure 5: Population Density vs. Crash Frequency

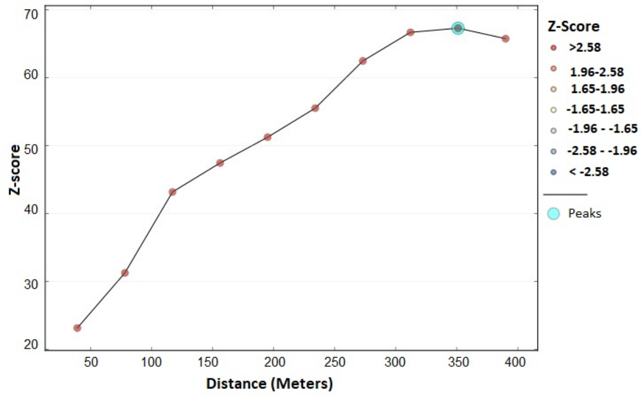

Figure 7: Incremental Spatial Auto correlation Results

5.2 Spatial Autocorrelation Analysis

The spatial autocorrelation analysis of aggregate dataset indicates a significant clustering of high-value crashes as seen by Moran's Index of 0.11, a z-score of 8.29, and a p-value of 0.00 at a 99% confidence interval (Figure 6). This provides a sufficient basis for the rejection of the null hypothesis, concluding that the dataset manifests a discernible clustering tendency rather than a purely random dispersion.

The result of incremental spatial autocorrelation for total crashes varies on a scale of 50 m to 400 m (Figure 7) with one peak at 351 m (z-score of 67.27). The peak distance led to overestimation of hotspot features; to address this problem, a reduced threshold distance was applied in the hotspot analysis.

5.3 Getis-Ord Gi* Analysis

The Getis-Ord Gi* statistics was computed to identify road sections that are significantly more hazardous than adjacent features. The GiZ-score ranged from -0.60 to 19.76 and GiP value ranged from 0 to 0.999976. The positive GiZ-score and small GiP value indicates statistically significant clustering of high values, while the negative GiZ-score and high GiP values means intense clustering of low values. The statistics categorized 302 segments as hotspots in +3 bins (red color), 89 segments as hotspots in +2 bins (orange color) and 95 segments as hotspots in +1 bins (90% yellow color) and rest of the insignificant segments in 0 bin (grey color). Figure 8 suggests that the spatial hotspots established with 99% confidence ($p < 0.01$) cover a substantial portion of the spatial units, whereas the prevalence of orange and yellow segments are less conspicuous.

The highly significant hotspots at 99% confidence are located as discontinuous clusters along Ring Road, Araniko Highway, Tribhuwan Highway, Tripureshwor – Ring Road, Thapathali – Ekantakuna, Samakhusi – Tokha – Gurje, Balkhu – Chovar, Satdobato – Pulchowki and Satdobato – Dhapakhel road sections (Figure 8). These hotspots are more persistent where there is presence of intersections, curves, bridge approaches, access roads, median barriers, bus stops, roadside objects,

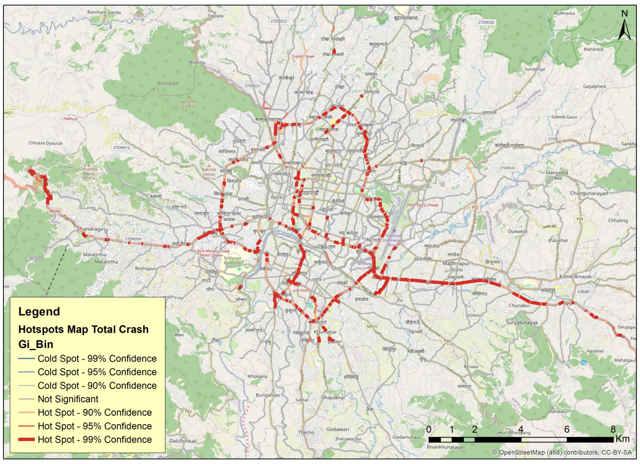

Figure 8: Getis-Ord Gi* Hotspot Map

pedestrian crossings, fuel stations and other features. These clusters are closely linked with traffic volume, vehicle speed, population density, residential developments and commercial/institutional activities (hospitals, schools, shopping complex, markets etc.). Additionally, the slope extremity and sharp bends may contribute to the higher RTC frequency observed in the eastern Araniko Highway and western Tribhuwan Highway.

5.4 KDE Mapping

In Figure 9, the results of the estimated kernel density are displayed in color gradient symbolizing the intensity of hotspots, ranging from 776.67 points/sq.km to 5,436.69 points/sq.km. White color denotes absence of hot spots, yellow indicates mild concentration, and finally red color represents significant concentrations of hotspots. Based on the visual interpretation of Figure 9, the prevalence of hotspots is higher on the Ring Road compared to other highways. On the Ring Road, hot spots are more visible at intersections, namely Koteshwor, Tinkune, Chabahil, Sukedhara, Narayan Gopal Chowk, Samkhusi, New Buspark, Kalanki, Balkhu, Nakhu Dobato, Ekantakuna, Satdobato and Gwarko areas.

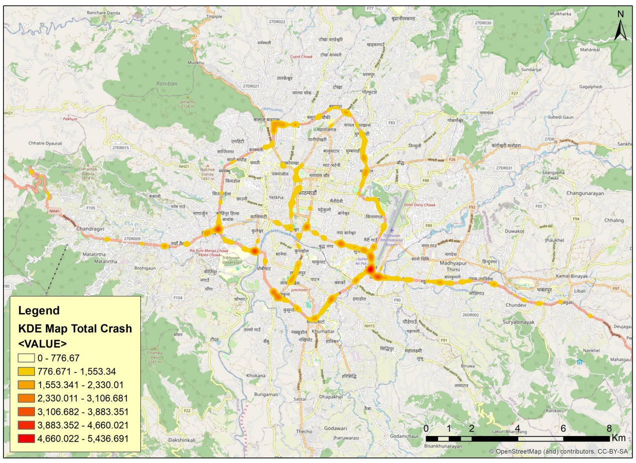

Figure 9: KDE Map for Total Crashes

By overlaying the maps created from Getis-Ord Gi* and KDE methods, it is evident that similar hotspots are identified by both the approaches. Though KDE effectively detected spatial hotspots, it failed to provide insights into the statistical significance of high or low crash frequencies at various locations. On the other hand, the Getis-Ord Gi* analysis revealed some additional hotspots along northern and southern roads of the Valley, but overgeneralized the results by marking entire road segments as hotspots irrespective of the length of segment that actually has a high concentration of crashes. Nevertheless, this approach precisely delineated the significant hotspot segments for subsequent ranking purpose.

5.5 Hotspots Ranking

The 287 common hotspots at 99% confidence level which are characterized by both the Getis-Ord Gi* and KDE methods were subjected to ranking process using the Equation (9). The ranking results are shown in Table 1. The top ten hotspots exhibit high number of deaths and major injury crashes in comparison to other hotspots identified at 99% confidence. As a result, these specific locations were assigned with relatively high weights of severity score, despites the fact that other locations had high number of crash incidences. It is critical for transportation agencies to prioritize the prevention of fatalities and severe injuries at these identified hotspots, due to their profound effects on people, families, and economic fabric.

Table 1: Top Ten Significant Hotspots

S.N.	Location	Crash Count	SI	Rank
1	Balkhu (TU)	81	2017	1
2	Dhungeadda	191	1862	2
3	Kaushaltar Chowk	123	1751	3
4	Chardobato Chowk	82	1723	4
5	Chundevi Chowk	74	1705	5
6	New Buspark	79	1634	6
7	Srijananagar Chowk	48	1561	7
8	Thasikhel Chowk	84	1331	8
9	Chabahil Chowk	29	1305	9
10	Nagdhunga	85	1203	10

6. Conclusions

This study introduced a novel hotspot analysis technique based on GIS spatial tools to facilitate the effective visualization of high crash risk areas and possible causes in the Kathmandu Valley. The study employed three methods; Global Moran's I, Getis-Ord Gi* and KDE to explore the geographical patterns of the highly significant hotspots. Based on the findings, the following conclusions were drawn:

- The positive Moran's I statistic, high z-score and low p value indicate that the patterns of hotspots are clustered in nature.
- The Getis-Ord Gi* and KDE techniques highlight similar locations of hotspots, and hence, their collective potential can be utilized to quickly visualize the statistically significant hotspots.
- The concentration of crash hotspots is significant along Ring Road, Araniko Highway, Tribhuwan Highway, Tripureshwor-

Ring Road, Thapathali-Ekantakuna Road and other main urban roads.

- Hotspots with the higher incidence of crashes do not necessarily correlate with higher severity score, and hence, they do not occupy higher position in ranking.
- Top ten hotspot sections as identified by the weighted severity index underscores the need for immediate mitigative action due to high incidences of deaths and major injuries associated with these sites.
- The ranked locations will provide a a strategic pathway to the native transportation management authorities to focus their investigative endeavors on the risk determinants at these locations, and ensuring optimum allocation of their constrained financial resources to reduce both the frequency of crashes and seriousness of impacts.

Future Research

The present paper is limited to the study of spatial patterns of the crashes and their severity, and neglects the correlation with other factors like land use, vehicle ownership, road geometry (such as length, sight distance and gradient), weather, time of crash etc. The limitation warrants future research to be focused around integration of these spatial and temporal variables in order to accurately assess the inter dependency of hotspot occurrences on such contributing factors.

Acknowledgments

The authors extend their sincere appreciation to Pulchowk Engineering College, Lalitpur for offering a conducive platform to conduct this research, and providing continuous guidance throughout the thesis work. The authors are also grateful to the Traffic Directorate, Naxal for their assistance in supplying the crash data.

References

[1] World Health Organization. Global status report on road safety 2023, 2023.

[2] Central Bureau of Statistics - Nepal. National population and housing census 2021 results, 2021.

[3] S. Adhikari. City report: Kathmandu metropolitan city, nepal, 2014.

[4] Jyoti Kuikel, Bhagwan Aryal, Tilak Bogati, and Bhagabati Sedain. Road traffic deaths and injuries in kathmandu. *Journal of Health Promotion*, 10(1):73–88, 2022.

[5] Hemant Tiwari. Dependency of road accidents with volume and speed (a case study of major black spot location within kathmandu valley). In *Proceedings of IOE Graduate Conference*, pages 339–343, 2015.

[6] Sali Shakya. *Ranking Road Safety Hazardous Locations in Nepal: A Case Study of Kalanki (Ch. 10+ 600)–Koteshwor (Ch. 20+ 994) Road Section*. PhD thesis, Pulchowk Campus, 2020.

[7] V Prasannakumar, H Vijith, R Charutha, and N Geetha. Spatio-temporal clustering of road accidents: Gis based analysis and assessment. *Procedia-social and behavioral sciences*, 21:317–325, 2011.

[8] Michal Bíl, Richard Andrášik, and Jiří Sedoník. A detailed spatiotemporal analysis of traffic crash hotspots. *Applied geography*, 107:82–90, 2019.

[9] Gholam Ali Shafabakhsh, Afshin Famili, and Mohammad Sadegh Bahadori. Gis-based spatial analysis of urban traffic accidents: Case study in mashhad, iran. *Journal of traffic and transportation engineering (English edition)*, 4(3):290–299, 2017.

[10] Zhuang-Zhuang Wang, Yi-Ning Lu, Zi-Hao Zou, Yu-Han Ma, and Tao Wang. Applying ohsa to detect road accident blackspots. *International journal of environmental research and public health*, 19(24):16970, 2022.

[11] Homayoun Harirforoush and Lynda Bellalite. A new integrated gis-based analysis to detect hotspots: A case study of the city of sherbrooke. *Accident Analysis & Prevention*, 130:62–74, 2019.

[12] Sahima Nazneen, Mahdi Rezapour, and Khaled Ksaibati. Application of geographical information system techniques to determine high crash-prone areas in the fort peck indian reservation. *The Open Transportation Journal*, 14(1), 2020.

[13] Long Tien Truong and Sekhar VC Somenahalli. Using gis to identify pedestrian-vehicle crash hot spots and unsafe bus stops. *Journal of Public Transportation*, 14(1):99–114, 2011.

[14] Tessa K Anderson. Kernel density estimation and k-means clustering to profile road accident hotspots. *Accident Analysis & Prevention*, 41(3):359–364, 2009.

[15] Benoît Flahaut, Michel Mouchart, Ernesto San Martin, and Isabelle Thomas. The local spatial autocorrelation and the kernel method for identifying black zones: A comparative approach. *Accident Analysis & Prevention*, 35(6):991–1004, 2003.

[16] Arthur Getis and J Keith Ord. The analysis of spatial association by use of distance statistics. *Geographical analysis*, 24(3):189–206, 1992.

[17] Md Saiful Alam and Nusrat Jahan Tabassum. Spatial pattern identification and crash severity analysis of road traffic crash hot spots in ohio. *Heliyon*, 9(5), 2023.

[18] Mahmoud Ahmed Mahmoud Hammas. *Spatial analysis of traffic accidents in the city of Medina by using GIS*. PhD thesis, King Abdulaziz University, 2023.

[19] A Afolayan, SM Easa, OS Abiola, FM Alayaki, and O Folorunso. Gis-based spatial analysis of accident hotspots: A nigerian case study. infrastructures 2022, 7, 103, 2022.

[20] UR Manepalli, Ghulam H Bham, and Srinadh Kandada. Evaluation of hotspots identification using kernel density estimation (k) and getis-ord (gi*) on i-630. In *3rd International Conference on Road Safety and Simulation*, volume 21, pages 14–16. National Academy of Sciences Indianapolis Indiana, United States, 2011.

[21] Pei-Fen Kuo, Xiaosi Zeng, and Dominique Lord. Guidelines for choosing hot-spot analysis tools based on data characteristics, network restrictions, and time distributions. In *Proceedings of the 91 Annual Meeting of the Transportation Research Board*, pages 22–26. Citeseer, 2011.

[22] Charlotte Plug, Jianhong Cecilia Xia, and Craig Caulfield. Spatial and temporal visualisation techniques for crash analysis. *Accident Analysis & Prevention*, 43(6):1937–1946, 2011.

[23] Environmental Systems Research Institute (ESRI). How hot spot analysis (getis-ord gi*) works, 2021.

[24] A Stewart Fotheringham, Chris Brunsdon, and Martin Charlton. *Quantitative geography: perspectives on spatial data analysis.* Sage, 2000.

[25] Samikshya Rizal and Hemant Tiwari. Analysis of road traffic crash cost in kathmandu valley. In *2nd International Conference on Integrated Transport for Sustainable Mobility*, 2023.