Analyzing Road Traffic Incident Hotspots Using Cluster Analysis in Thanh Hoa Province of Vietnam

Ha, L. T.,[1,2] Thao, V. T. P.[1]* and Thao, D. T. P.[1]
[1]Hanoi University of Mining and Geology, Ministry of Education and Training of Vietnam; 18 Duc Thang, Bac Tu Liem, Hanoi, Vietnam, E-mail: vuthiphuongthao@humg.edu.vn,* vtpthao1975@gmail.com
[2]Campus in Ho Chi Minh City, University of Transport and Communications. No.450-451 Le Van Viet Street, Tang Nhon Phu A Ward, Thu Duc City, Ho Chi Minh City, Vietnam
Corresponding Author

DOI: https://doi.org/10.52939/ijg.v20i9.3537

Abstract
This study uses remote sensing images combined with Geographic Information Systems statistical approach - spatial autocorrelation analysis to identify traffic incident hotspot locations. It evaluates the statistical significance of hotspot clusters to support forecasting road traffic incidents applied for Thanh Hoa province of Vietnam with traffic accident data during four years from 2020 to 2023. Based on spatial analysis techniques, including severity index and Getis-Ord Gi statistical analysis with inverse distance weighting interpolation, hotspot clusters are identified and sorted by rank. Applying spatial autocorrelation analysis has important implications in enumerating hotspots in sequence. The analysis results identified 64 hotspots, of which the 32 dangerous incident points are in Thanh Hoa province's transportation system. The results demonstrate the effectiveness of using Getis-Ord Gi* spatial statistical techniques to identify traffic incident hotspot locations and evaluate the statistical significance of hotspot clusters to support forecasting road traffic incidents in Thanh Hoa province.*

Keywords: Autocorrelation, Getis-Ord Gi*, Hotspot Analysis, Inverse Distance Weighting, Severity Index, Traffic Incident

1. Introduction

Road traffic incidents (RTI) are one of the significant concerns and have a severe impact, causing traffic congestion, especially in major cities in Vietnam [1]. The potential effects of RTI, such as injuries, deaths, property damage, disability, and high financial costs on both GDP and road users and involved individuals [2]. According to a report by the World Health Organization, low- and middle-income countries bear the most significant burden of RTI in the world [3]. In Vietnam, road traffic is very complicated because motorbikes are the primary means of transport due to their convenience, ease of use, and low cost.

Due to that convenience, motorbikes are also the type of vehicle that causes many RTIs in Vietnam [4]. Besides motorbikes, many other means of transport, including rudimentary vehicles such as carts, bicycles, cyclos, animal-drawn vehicles, etc, run alongside modern road vehicles such as large and small cars. In addition, many people do not obey traffic regulations when participating in traffic. The traffic system with various rudimentary vehicles and poor awareness of traffic participants is the cause of many traffic accidents in Vietnam [5] and [6] minimize accidents in Vietnam; it is necessary to warn of locations where incidents often occur. That is the location of the hotspot or blackspot of road traffic accidents [7] and [8]. Previously, many methods for identifying hot spots were researched, such as the incident rate method [9], equivalent property damage index [10], Bayesian empirical method (EB) [11], etc. The most effective way to identify hot spots is the EB method [12]. However, the EB method has limitations in requiring training and high skills in statistical analysis.

Over the last two decades, the rapid development and application of GIS technology as a tool to manage, explore, analyze, and visualize spatial data and spatial data analysis methods included in the modules of GIS software have received increasing attention [13]. GIS can easily encode locations and map incidents [14] or create an incident density distribution map [15]. Some studies have focused on two techniques, which are Getis-Ord Gi* statistics [16] and [17] and kernel density estimation (KDE) [18] and [19] using GIS to identify hotspots.

Hotspots are considered as the number of severe injury collisions occurring within a defined length segment (500 m) in 3 years or if the number of deaths is equal to or greater than 10 [20]. In other words, a cluster of crashes with high levels of injuries is called hotspots.

Conversely, if there is a cluster of low-injury crashes, it is called "coldspots ". In this study, hotspots were locations with positive Getis-Ord GI* statistics with significant *z*-scores at a 90% or higher confidence level. Similarly, coldspots are locations with adverse Getis-Ord Gi* statistics and significant z-scores at a confidence level of 90% or higher. In other words, a cluster of crashes with high levels of injuries is called a hotspot. Conversely, a cluster of low-injury crashes is called a coldspot. Black spots in road traffic are also known as dangerous road locations, high-risk locations, accident-prone locations, hot spots, prospective locations, and priority investigation locations [21].

The superiority of the KDE algorithm was compared to other clustering techniques, including k-means, k-medoids, CURE (Clustering Using Representatives), BIRCH (Balanced Iterative Reducing and Clustering Using Hierarchies), STING (Statistical Information Grid), CLIQUE (Clustering in Quest), DBSCAN (Density-Based Spatial Clustering of Applications with noise) and OPTICS (Ordering Points to Identify the Clustering Structure) [22]. The KDE algorithm helps determine the range that covers an incident's danger level. Thus, it visually depicts where problems will likely occur according to spatial relationships.

Furthermore, the KDE analysis process creates a raster layer whose pixels are assigned values corresponding to the intensity that characterizes the entire area for comparison and classification [23]. KDE is, therefore, effective in clearly displaying areas with high incident density [24]. However, the limitation of KDE is that it does not provide statistical significance of hot spots [25] and [26], and there are no criteria to prioritize incident hotspots [2]. To overcome that, some other studies have applied statistical analysis -Getis-Ord Gi* to detect important locations where accidents occur [27] and [28]. Getis–Ord (Gi*) spatial statistics were used to identify hotspots on the highway scale [27]. The method successfully identified accident clusters from more than 47.359 accident incident records from 2016 to 2019. Twenty-five hotspots were identified in this study area, accounting for 26.81% of reported cases, with the most extended hot spot cluster being 31.2 km and the shortest being 300m. There were 16 identified hotspots, equivalent to 64% of the total number of hotspots occurring at locations with intersections, exits, ramps, resorts, or parking stops. The spatial pattern of traffic accident hot spots along the internal and significant road networks was used to present the local hotspot analysis technique (Getis–Ord Gi*) in a GIS environment. Data was collected from 2015 to 2019 for the Irbid Governorate, Jordan.

Analysis of traffic accident incidents at a severe level shows a precise spatial distribution of hotspot locations. Less serious traffic accidents (~95%) occur on the internal road network, at a 99% significance level along the internal and significant road network segments. Both studies define the output of statistical Gi* as z-scores and p-values for each failure location to determine whether the cluster is statistically significant. In addition, the accident severity index (SI) is also considered in most studies. Some authors indicated that it is better to pay attention to the severity of an accident incident because the accident's impact may be related to the injury [29] and [30]. The purpose of this study is to analyze the spatial pattern of road traffic incidents, as well as map incident hotspots in the Thanh Hoa province area. The survey result will assist people in understanding the causes of incidents, handling dangerous locations, and providing reasonable proposals and costs to enhance road safety and minimize the risk of accidents.

2. Materials and Methodology

2.1 Study Area

Thanh Hoa is a large province of the North Central region of Vietnam with geographical coordinates ranging from 20°40′ to 19°18′ North latitude and from 106°04′ to 104°22′ East longitude. There is a convenient transportation system for regional, national, and international travel from Thanh Hoa with the Trans-Vietnam Railway system, Ho Chi Minh Road, many national highways, Nghi Son deep-water seaport, and a river waterway [31]. However, the density of the provincial road network is unevenly distributed among regions, mainly concentrated in urban areas. Furthermore, the scale of the road network and structure is low; the asphalt road system primarily focuses on national and provincial highways, and the asphalt pavement structure is low. That has significantly affected the travel process of traffic participants and caused many severe incidents in road traffic. Figure 1 is the map of Thanh Hoa province.

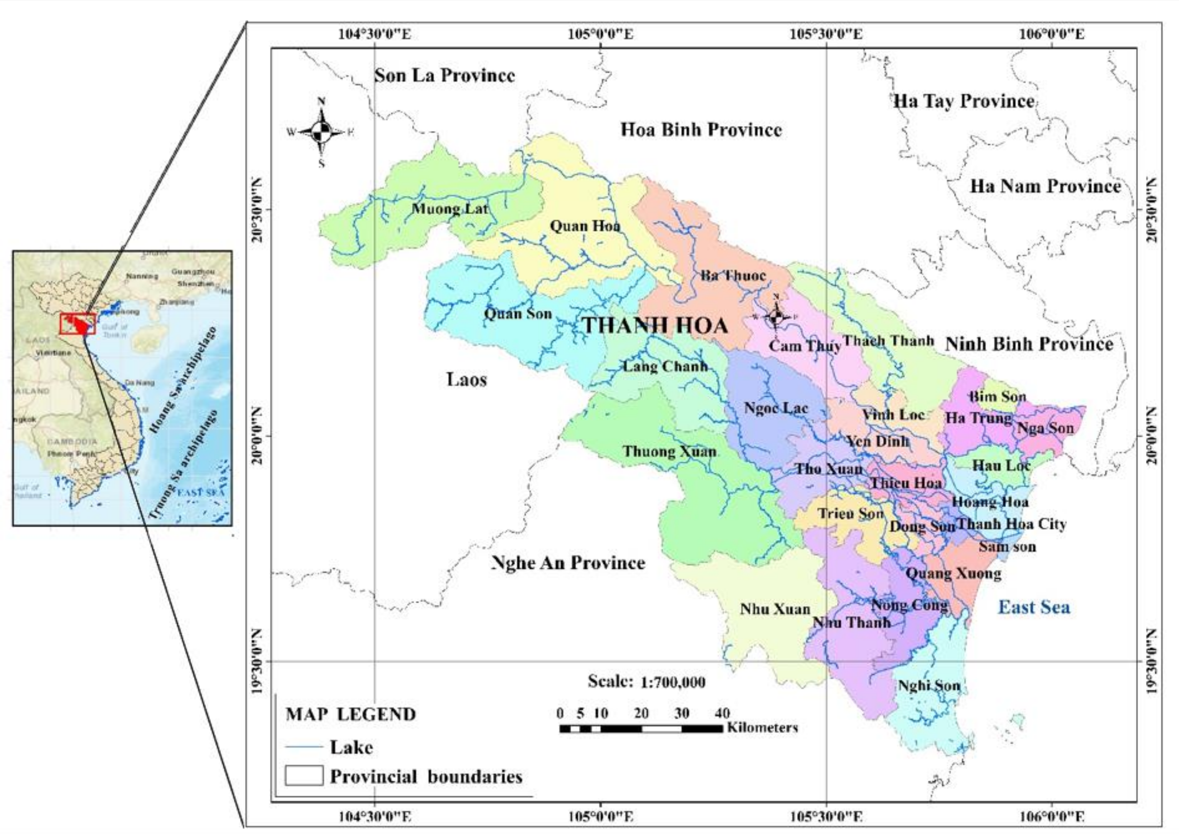

Figure 1: Thanh Hoa province of Vietnam

Table 1: Satellite imagery data

Satellite	Processing level	Path/row	Acquisition Date
Landsat-8	L1GT	127/046	18/04/2022
		127/046	06/09/2022
		127/047	17/08/2023

2.2. Data

In this study, two different datasets were utilized, that is:

1) Map of Thanh Hoa road network and administrative boundaries digitized from Google Earth in shapefile form. Road attribute data includes length, width, road type, speed limit, etc. Collected data from reports provided by Thanh Hoa Department of Natural Resources and Environment. In addition, surface cover is classified and verified from 3 Landsat-8 satellite images with information in Table 1.

2) The Thanh Hoa Transportation Police Department provided an Accident incident data set from 2020 to 2023 in Thanh Hoa province. This dataset includes essential information such as the date, time, and location of the accident, the type of accident and vehicle involved, the number of deaths, and the number of injuries. These two datasets were combined and processed using ArcGIS 10.8 software. Figure 2 illustrates the distribution of road traffic incident locations in Thanh Hoa province over four years from 2020 to 2023.

Data was collected to build a database of road traffic incidents in Thanh Hoa province for four consecutive years, from 2020 to 2023, and a geographic database including road network, administrative boundaries, and surface cover. Analysis results were obtained by running the hotspot analysis tool in ArcGIS software. The study proposes a research framework for analyzing incident hotspots using the Getis-Ord Gi* spatial statistical technique, as shown in Figure 3. It can be divided into three parts: 1) Collected data, 2) Analysed data, and 3) Results.

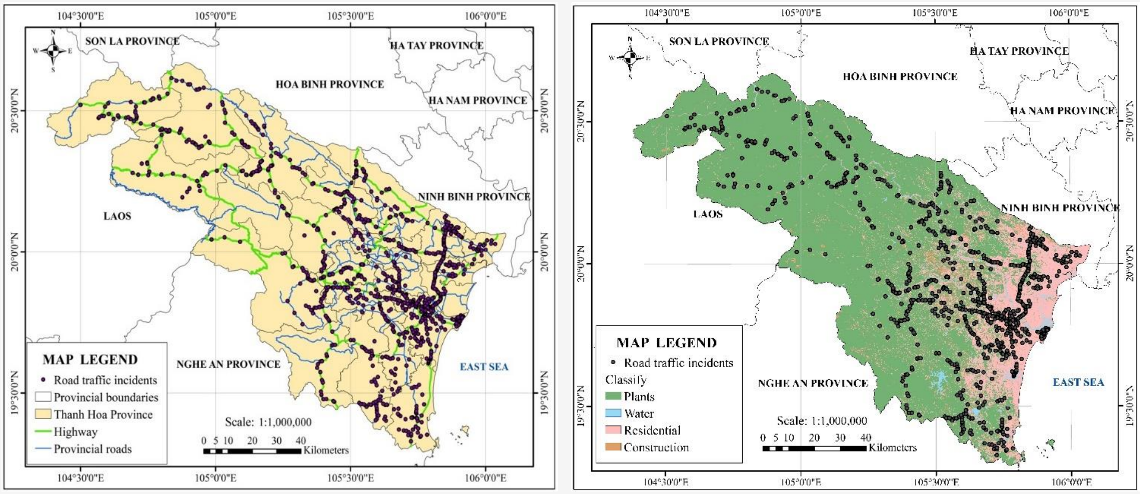

Figure 2: Road traffic incident distribution in Thanh Hoa province from 2020 to 2023

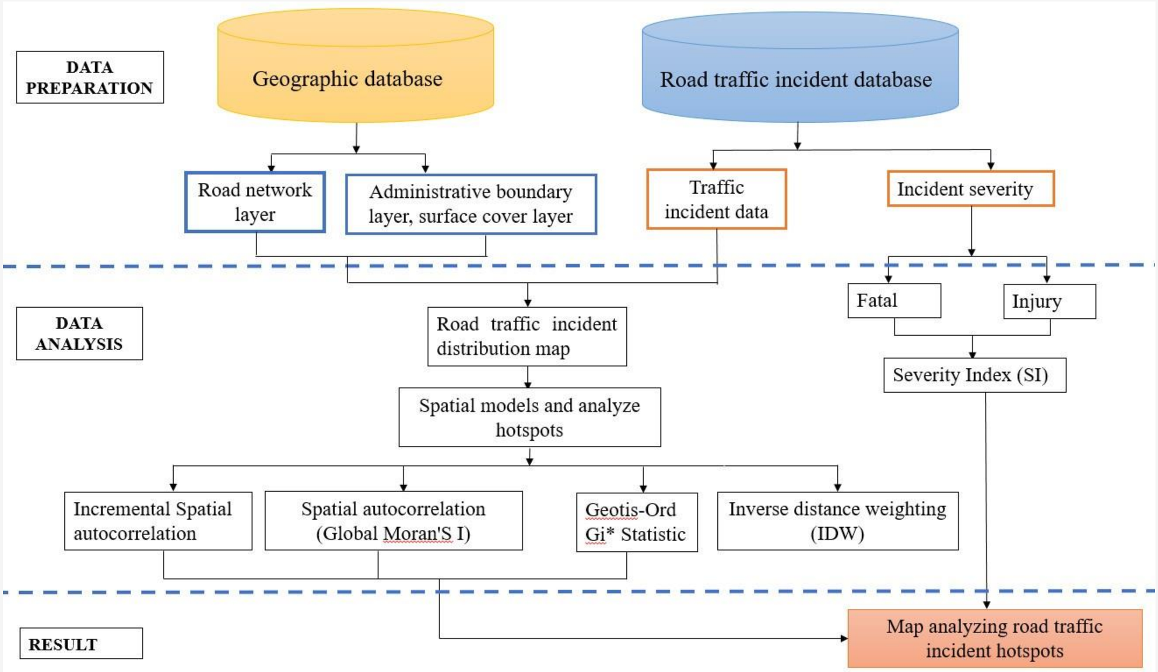

Figure 3: Road traffic incident hotspot analysis process

2.3.2 Data analysis

Moran's Index Statistics of Spatial Autocorrelation:

Statistically significant analysis of hotspots is a measure of spatial autocorrelation by observing incidents' distribution patterns. The test will decide to stop when there is a random distribution and vice versa. The spatial autocorrelation technique examines the location and properties of the incident severity z values of determined Moran's I ranging from -1 to 1, which are considered statistically significant [32]. When the Moran index value is close to 1, the data contain spatial autocorrelation and cluster patterns; when the Moran index value is close to -1, the data is not continuous and scattered.

At the same time, an absolute zero index value indicates no spatial autocorrelation. Based on the weight matrix, Moran's I is calculated from equation 1 [33]:

$$I = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j} z_i z_j}{S_0 \sum_{i=1}^{n} z_i^2}$$

Equation 1

In which: z_i, z_j are the deviation from the average value; $w_{i,j}$ is the spatial weight between feature *i* and *j*; *n* is the number of features, S_0 is the sum of all spatial weights and is calculated as in equation 2 [34]:

$$S_0 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}$$

Equation 2

The z_i score is calculated from equations 3 to 5:

$$Z_I = \frac{I - E[I]}{\sqrt{V[I]}}$$

Equation 3

With:

$$E[I] = -\frac{1}{n-1}$$

Equation 4

$$V[I] = E[I^2] - E[I]^2$$

Equation 5

Incremental Spatial Autocorrelation (ISA) Tool:
An incremental spatial autocorrelation tool in ArcGIS is used to calculate Moran's I index values and the associated *z*-score, indicating statistical significance at different threshold distances. The distance where the *z*-score is highest is the desired threshold distance and is chosen for cluster mapping using the GetisOrd Gi* function. The ISA tool gives results as a graphical representation between different distance thresholds and the associated *z*-score. The first or highest peak can be selected for optimal clustering among other values.

Getis-Ord Gi Statistics:
Mapping incident clusters in the study area based on hotspot analysis - Getis-ord Gi*. The output of this method includes Gizscores and Gipvalues that indicate whether high-value or low-value objects tend toward clusters [32]. The *z*-score and *p*-value obtained from spatial autocorrelation analysis indicate the incident locations with high values - hotspots and low values -coldspots. However, to be designated as a hot spot, the incident location needs to have a higher value, with surrounding incident characteristics also having high values. The Getis-ord *Gi** statistic was calculated to identify such hotspots according to equation 6:

$$G_i^* = \frac{\sum_{j=1}^{n} w_i x_j - \bar{X}\sum_{j=1}^{n} w_{i,j}}{S\sqrt{\frac{n\sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j}\right)^2}{n-1}}}$$

Equation 6

There, x_j is the attribute value of object *j*, $w_{i,j}$ is the spatial weight between objects *i* and *j*, and *n* is the total number of objects.

With:

$$\bar{X} = \frac{\sum_{j=1}^{n} x_j}{n}$$

Equation 7

$$S = \sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - \bar{X}^2}$$

Equation 8

The correlation between significance level *p*, confidence, and *z*-sore is shown in Table 2. A positive *z*-score value (>2.58) is statistically significant (p =0.01), indicating that the observations are hot spot spatial clusters, and a negative *z*-score (< -2.58) is statistically significant ($p = 0.01$); otherwise, it is the cold spot. Regarding *z*-score > 2.58 and *p*-value < 0.01, the observations are spatial clusters, and Gi* is statistically significant at the 99% confidence level. On the contrary, in the case of *p*-value ≥ 0.01, the observations are random [8] and [32].

Inverse Distance Weighting (IDW):
The calculated hotspot values were interpolated using the inverse distance weighting (IDW) interpolation technique in ArcGIS software. This IDW technique assumes that the influence of the mapping variable decreases with increasing distance from the sampling location. Therefore, this technique gives higher weight to variables near the sampling location and vice versa. The estimated value of the point at a particular location can be calculated using equation 9:

$$z_0 = \frac{\sum_{i=1}^{s} z_i \frac{1}{d_i^k}}{\sum_{i=1}^{s} \frac{1}{d_i^k}}$$

Equation 9

Where z_0 is the estimated value at point *0;* z_i is the measured value at point i; *s* is the number of points used to estimate the unknown value; d_i is the distance from point *i* to point *0;* k is the exponential value that determines the effect of distance.

Table 2: Correlation between confidence level and *z*-score, confidence level and *z*-score *p*-value [8]

Confidence level (%)	z score	p-value
90	<-1.65 or >1.65	<0.10
95	<-1.96 or >1.96	<0.05
99	<-2.58 or >2.58	<0.01

3. Result

3.1 Global Spatial Autocorrelation Analysis -Spatial Autocorrelation Moran's I

According to Figure 4 and the correlation in Table 2, analysis of Global Moran's I show that Moran's I is 0.3840 (> 0.00), which shows that traffic incidents positively correlate with the spatial distribution. A *p*-value of 0.0000 (<0.01) indicates that the probability of random distribution of traffic incidents is less than 1%. In contrast, the spatial clustering state in Figure 4 has a more than 99% probability. Furthermore, the value *z* = 7.6543 is higher than 2.58, which proves that the spatial aggregation trend of accident incidents distributed according to the global Moran is very significant; that is, traffic incidents have a concentrated pattern.

3.2 Incremental Spatial Autocorrelation Analysis

Incremental spatial autocorrelation is applied to analyze spatial autocorrelation using different distance ranges, called bandwidths. Each bandwidth gives a z-score and statistical significance level p, as shown in Table 3. According to Figure 5, the maximum peak value is 10250, with *z*-score of 3.748003 and *p*-value of 0.000178, at which we can find the maximum clustering ability.

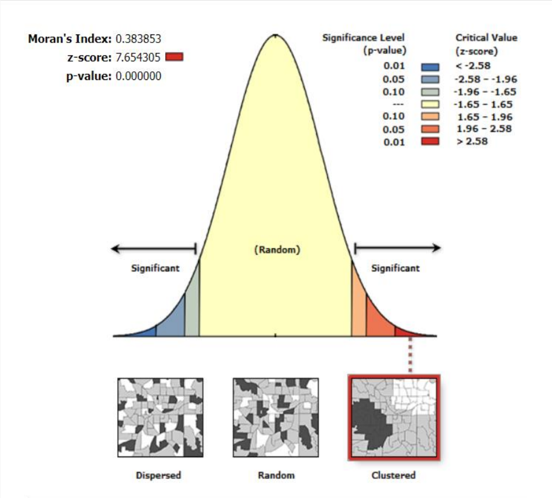

Figure 4: Spatial autocorrelation report

Table 3: Global Moran's I Summary by Distance

Distance	Moran's Index	Expected Index	Variance	z-score	p-value
10000.00	0.018421	-0.000628	0.000026	3.723879	0.000196
10250.00	0.018175	-0.000628	0.000025	3.748003	0.000178
10500.00	0.014999	-0.000628	0.000024	3.177986	0.001483
10750.00	0.016979	-0.000628	0.000023	3.650584	0.000262
11000.00	0.016423	-0.000628	0.000022	3.612572	0.000303
11250.00	0.015534	-0.000628	0.000021	3.543530	0.000395
11500.00	0.016051	-0.000628	0.000020	3.725932	0.000195
11750.00	0.014330	-0.000628	0.000019	3.396120	0.000683
12000.00	0.013956	-0.000628	0.000019	3.368443	0.000756
12250.00	0.011782	-0.000628	0.000018	2.917117	0.003533

First Peak (Distance; Value): 10250.00; 3.748003
Max Peak (Distance; Value): 10250.00; 3.748003
Distance measured in Meters

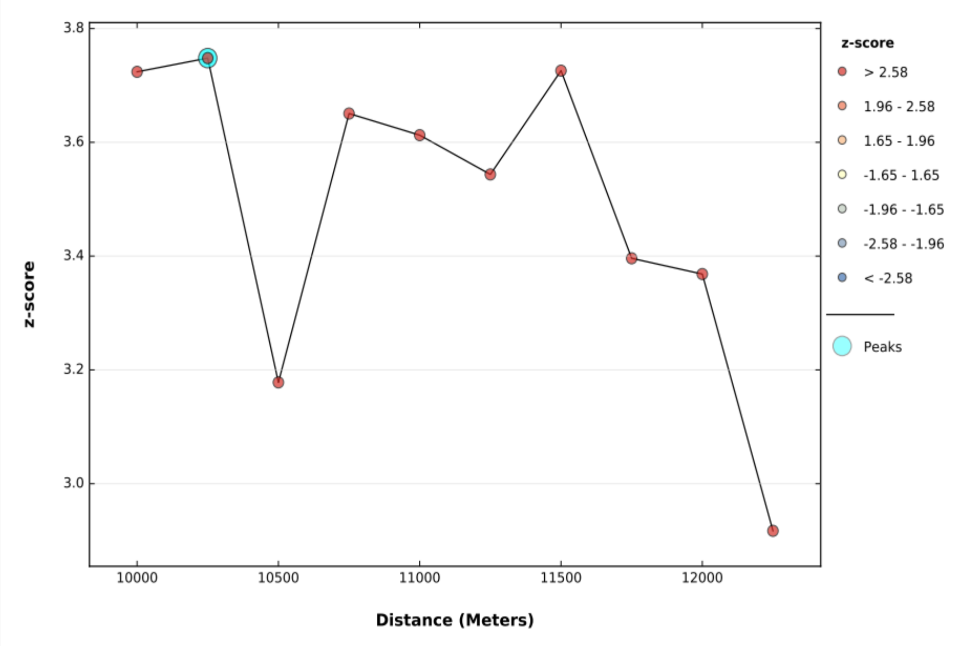

Figure 5: Spatial autocorrelation over distance

3.3 Cluster Analysis and IDW

Analysis of hotspots and coldspots of road traffic incidents in Thanh Hoa province shows that areas with high spatial clustering (red dot clusters) are concentrated on national highways and provincial roads in these areas such as Thanh Hoa city, Bim Son town, Ha Trung district, Hau Loc district, Hoang Hoa district, Sam Son city, Quang Xuong district, Dong Son district. Areas with low spatial clustering (coldspots consisting of light blue dots), such as some locations in Nhu Thanh District, Nong Cong District, Nghi Son Town, Quan Hoa District, and Thach Thanh District. Some remaining districts have no spatial clustering (representing a random distribution of incidents). Figure 6 illustrates an Inverse Distance Weighting incident hotspot Map, while Figure 7 is the distribution of hotspots and coldspots of traffic incidents in Thanh Hoa province.

4. Discussion

The IDW interpolation technique is widely implemented to generate spatial hot spot analysis output [18] and [35]. Figure 6 shows continuous smooth surfaces classified into five different classes. Areas with a high risk of incident are shown in red with high values and based on data overlay of analysis results using Getis Ord Gi* statistics to identify incident hotspots and raster surface interpolation from points using inverse distance weighting technique. Figure 6 analysis identified 64 hotspots, of which the 32 most dangerous locations were ranked based on the *z*-score for each location and were statistically significant with a 99% confidence interval. Figure 9 is the map of Thanh Hoa province's road network. The study results show that along National Highway 1A, National Highway 217, and National Highway 10, there are incident points in very high accident-risk areas. Currently, there is one location, location number 345 - an open crossroad connecting to National Highway 1A; 4 locations in the curve area consisting of numbers 422, 393, 419, and 389; 2 black spot locations - numbers 379, 331where serious incidents often occur at intersections. These locations are very dangerous, uncontrolled, and considered a top priority for treatment. The results obtained are consistent with the statistical report of Thanh Hoa Transportation Police Department from 2020 to 2023 and actual observation conditions.

There are four main causes of road traffic incidents in these hotspot areas. The first cause is that there are often no signs at the connecting crossroad location, and the terrain is often lower than the national highway surface, leading to obscured vision. The second cause is that these national highways are being degraded in some sections and are in the process of being upgraded, such as National Highway 1A, National Highway 10, etc, as illustrated in Figure 8. The third cause is frequent illegal behavior in other corridors, such as exposing agricultural products [36].

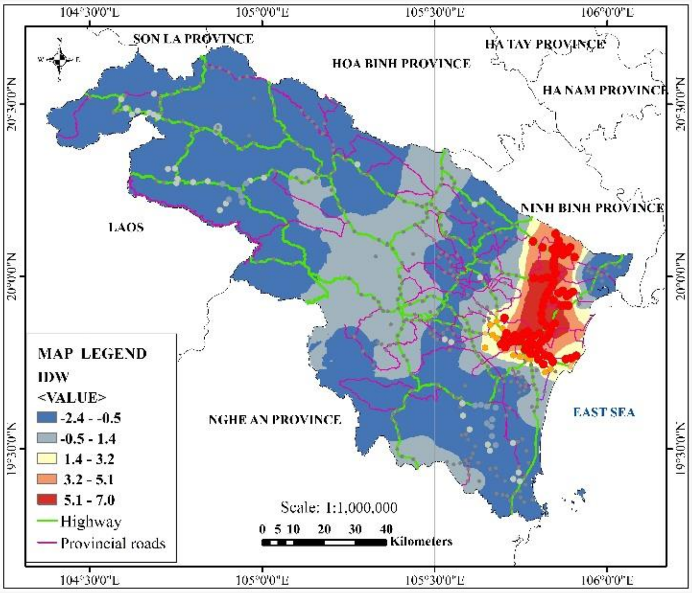

Figure 6. Inverse Distance Weighting incident hotspot

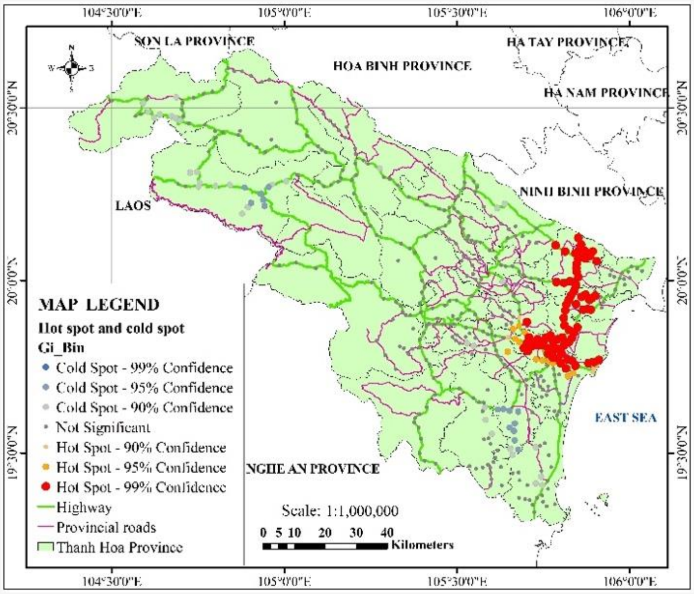

Figure 7: Map Getis Ord Gi* statistical incident hotspot map of Thanh Hoa province

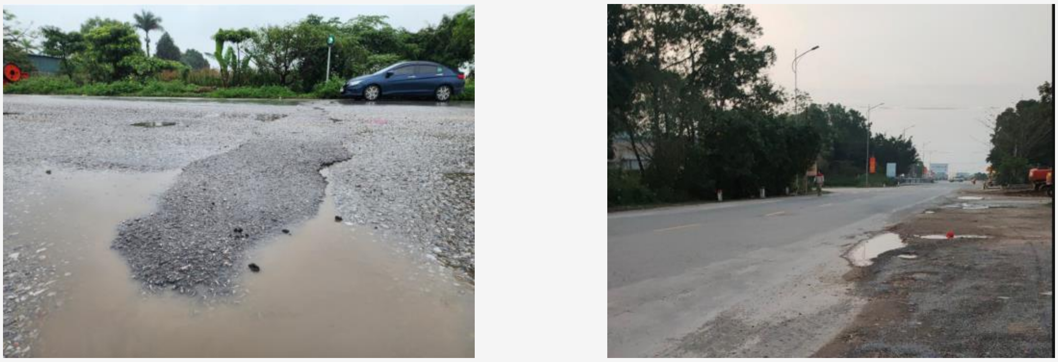

Figure 8: National highways of Thanh Hoa province are being degraded in some sections

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

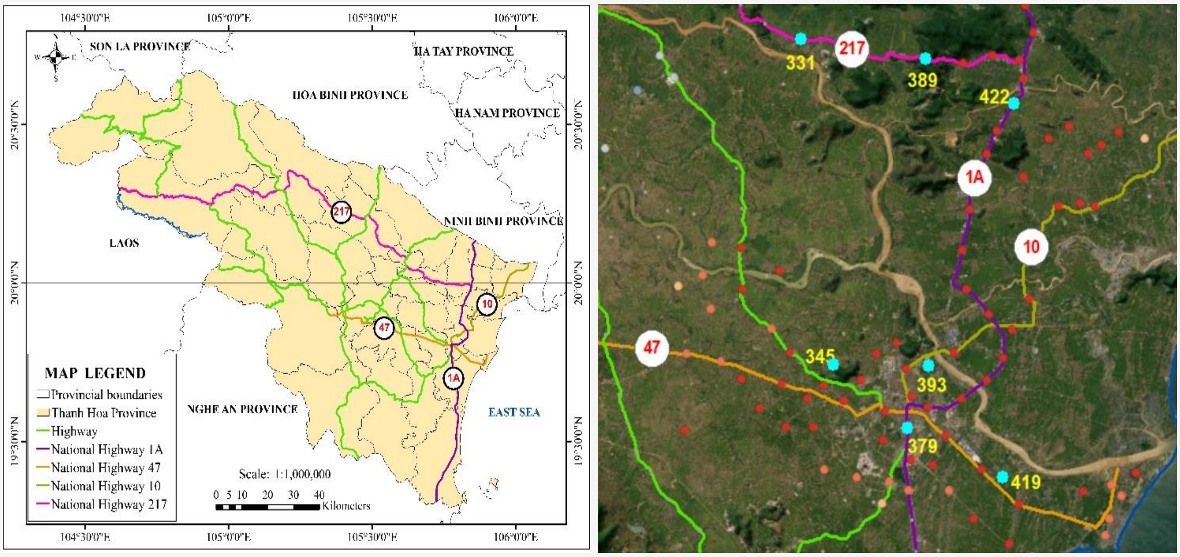

Figure 9: Road network of Thanh Hoa province

And final cause, there are many undulating buildings obscure the view, along with the high density of vehicles participating in traffic, typically in the location of black point 379 (at Km16+570/QL47) at the intersection of Highway 47 with Nguyen Duy Hieu Street in the center of Thanh Hoa City and black point 331 (at Km 23+780/QL 47) at the intersection of Nguyen Mong Tuan Street with National Highway 47.

5. Conclusion

The study proposed a process for analyzing road transportation hotspots using Getis-Ord Gi* spatial statistical techniques, deployed on a Geographic Information Systems platform, namely spatial autocorrelation analysis, to identify and quantify statistically significant spatial patterns based on the number and severity of traffic incidents in Thanh Hoa province for the past years from 2020 to 2023. Using spatial autocorrelation allows a statistical analysis of incident space patterns, helping to distinguish high-incident clusters from low-incident clusters. Risk hotspots are analyzed, identified, and sorted by priority. The results obtained are in close agreement with the statistical report of Thanh Hoa Transportation Police Department from 2020 to 2023 and actual monitoring datasets. From there, the cause is found, and a plan to avoid accidents in dangerous locations comes up in the future. Assessing the statistical significance of hotspots in the past will help management agencies make decisions on costs and appropriately allocate labor safety resources, reducing property damage and human lives. The study results prove the correctness of Getis-Ord Gi* spatial statistical techniques to identify traffic incident hotspot locations and evaluate the statistical significance of hotspot clusters to support forecasting road traffic incidents in Thanh Hoa province of Vietnam. However, this study has not investigated the spatial dependence of incidents on road parameters such as slope, curvature, stopping distance, etc. In future studies, combining some clustering methods to improve accuracy is necessary.

References

[1] Truong, Q. T., (2023). Improving the Effectiveness of State Management of Road Traffic. *State Management Magazine*, Vol. 331, 76-80. https:/doi.org/10.59394/qlnn.331.2023.589.

[2] Haulle, E. and Kisiri, M., (2016). The Impact of Road Accidents to the Community of Iringa Municipality: Challenges in Reducing Risks. *International and Multidisciplinary Journal of Social Sciences*, Vol. 5(3), 253-280. http://doi.org/10.17583/rimcis.2016.1880.

[3] World Health Organization, (2015). Global Status Report on Road Safety, 2015. Geneva, Switzerland. https://www.afro.who.int/sites/default/files/2017-06/9789241565066_eng.pdf.

[4] Vu, A. T. and Nguyen, D. V. M., (2021). Potential Effectiveness of a Training Program to Improve the Ability to Recognize and Handle Dangerous Situations on the Road for Motorbike and Moped Drivers, *Vietnam Journal of Science and Technology*, Vol. 63(10), 29-33. https://doi.org/10.31276/VJST.63(10).

[5] Hsu, T. P., Sadullah, E. A. F. M. and Dao, I. N. X., (2003). A Comparison Study on Motorcycle Traffic Development in Some Asian Countries-Case of Taiwan, Malaysia and Vietnam, *The Eastern Asia Society for Transportation Studies (EASTS)*, 1-51. https://www.researchgate.net/publication/284685742_A_comparative_study_on_motorcycle_traffic_development_of_Taiwan_Malaysia_and_Vietnam.

[6] Lubis, H. A., (2009). Motorcycles Growth and Its Impacts to Urban Transportation. *Proceedings of the Eastern Asia Society for Transportation Studies*, Vol. 7, 1-14. https://www.jstage.jst.go.jp/article/eastpro/2009/0/2009_0_329/_pdf.

[7] Khatun, M. S., Hossain, M. A., Kabir, M. A. and Rahman, M. A., (2024). Identification and Analysis of Accident Black Spots Using Geographic Information System (GIS): A Study on Kushtia-Jhenaidah National Highway (N704), Bangladesh. *Heliyon*, Vol. 10, 1-16. https://doi.org/10.1016/j.heliyon.2024.e25952.

[8] Minale, A. S., (2015). Identifying the Hot Spot Areas of Road Traffic Accidents. *Jordan Journal of Civil Engineering*, Vol. 9(3), 358-370. https://www.researchgate.net/publication/280934126_Identifying_the_Hot_Spot_Areas_of_Road_Traffic_Accidents.

[9] Carson, J. L. and Powers, M., (2004). Before-After Crash Analysis: A Primer for Using the Empirical Bayes Method Tutorial. *Montana Department of Transportation*. Research Section. https://rosap.ntl.bts.gov/view/dot/24878.

[10] Campbell, J. R. and Knapp, K., (2005). Alternative Crash Severity Ranking Measures and the Implications on Crash Severity Ranking Procedures. *Proceedings of the 2005 Mid-Continent Transportation Research Symposium, Ames Iowa, United States*. https://trid.trb.org/view/760618

[11] Hauer, E., Harwood, D. W., Council, F. M. and Griffith, M. S., (2002). Estimating Safety by the Empirical Bayes Method: A Tutorial, Transportation Research Record. *Journal of the Transportation Research Board*, Vol. 1784(1), 126-131. https://doi.org/10.3141/1784-16.

[12] Manepalli, U. R., Bham, G. H. and Kandada, S., (2011). Evaluation of Hotspots Identification Using Kernel Density Estimation (K) and Getis-Ord (Gi*) on I- 630. 3rd International Conference on Road Safety and Simulation, 2011, Indianapolis. https://onlinepubs.trb.org/onlinepubs/conferences/2011/RSS/2/Manepalli,UR.pdf.

[13] Booth, A. J., (2004). Spatial Statistic and Aquatic Geographic Information Systems. *Proceedings of the Second International Symposium on GIS/Spatial Analysis in the Aquatic Sciences, Kawagoe-city, Japan*, 3-44.

[14] Kim, K. and Levine, N., (1996). Using GIS to Improve Highway Safety. Computers, Environment, and Urban Systems, Vol. 20(4), 289-302. https://doi.org/10.1016/S0198-9715(96)00022-13.

[15] Peled, A., Haj-Yehia, B. and Hakkert, A. S., (1996). ArcInfo-based Geographical Information System for Road Safety Analyses and Improvement. *1996 ESRI User Conference Environmental Systems Research Institute, Incorporated*. http://gis.esri.com/library/userconf/proc96/TO50/PAP005/P5.HTM.

[16] Majumder, R., Bhunia, G. S., Patra, P., Mandal, A. C., Ghosh, D. and Shit, P. K., (2019). Assessment of Flood Hotspot at a Village Level Using GIS-based Spatial Statistical Techniques. *Arabian Journal of Geosciences*, Vol. 12, 1-12. https://doi.org/10.1007/s12517-019-4558-y.

[17] Truong, L. T. and Somenahalli, S. V. C., (2011). Using GIS to Identify Pedestrian-vehicle Crash Hotspots and Unsafe Bus Stops. *Journal of Public Transportation*, Vol. 14(1), 99-114. https://doi.org/10.5038/2375-0901.14.1.6.

[18] Kuo, P. F., Zeng, X. and Lord, D., (2011). Guidelines for Choosing Hot-spot Analysis Tools Based on Data Characteristics, Network Restrictions, and Time Distributions. *Proceedings of the 91 Annual Meeting of the Transportation Research Board, College Station, TX, USA, 14 November 2011*. https://www.yumpu.com/en/document/view/12102061/guidelines-for-choosing-hot-spot-analysis-tools-on-ceprofs-.

[19] Prasannakumar, V., Vijith, H., Charutha, R. and Geetha, N., (2011). Spatio-temporal Clustering of Road Accidents: GIS Based Analysis and Assessment. *Procedia-social and Behavioral Sciences*, Vol. 21, 317-325. https://doi.org/10.1016/j.sbspro.2011.07.020.

[20] Mhetre, K. V. and Thube, A. D., (2023). Road Safety, Crash Hot-spot, and Crash Cold-spot Identification on a Rural National Highway in Maharashtra, India. *Materials Today: Proceedings*, Vol. 77, 780-787. https://doi.org/10.1016/j.matpr.2022.11.447.

[21] Haghighi, F. and Karimi, E., (2018). Evaluation and Statistical Validation of Black-spots Identification Methods, *International Journal of Transportation Engineering*, Vol. 6(1), 1-15. https://www.ijte.ir/article_52977.html.

[22] Han, J. and Kamber, M., (2001). *Data Mining Concepts and Techniques*. San Francisco, CA: Morgan Kaufmann Publishers.

[23] Anderson, T. K., (2009). Kernel Density Estimation and K-means Clustering to Profile Road Accident Hotspots. *Accident Analysis & Prevention*, Vol. 41(3), 359-364. https://doi.org/10.1016/j.aap.2008.12.014.

[24] O'Sullivan, D. and Unwin, D. J., (2010). *Geographic Information Analysis*, 2009, New York, NY: John Wiley and Sons.

[25] Xie, Z. and Yan, J., (2013). Detecting Traffic Accident Clusters with Network Kernel Density Estimation and Local Spatial Statistics: An Integrated Approach. *Journal of Transport Geography*, Vol. 31, 64-71. https://doi.org/10.1016/j.jtrangeo.2013.05.009.

[26] Yao, S., Loo, B. P. Y. and Yang, B. Z., (2016). Traffic Collisions in Space: Four Decades of Advancement in Applied GIS. *Annals of GIS*, Vol. 22(1), 1-14. https://doi.org/10.1080/19475683.2015.1085440.

[27] Manap, N., Borhan, M. N., Yazid, M. R. M., Hambali, M. K. A. and Rohan, A., (2019). Determining Spatial Patterns of Road Accidents at Expressway by Applying Getis-Ord Gi* Spatial Statistic. *International Journal of Recent Technology and Engineering*, Vol. 8(3S3), 345-350. https://doi.org/10.35940/ijrte.C1004.1183S319.

[28] Hazaymeh, K., Almagbile, A. and Alomari, A. H., (2022). Spatiotemporal Analysis of Traffic Accidents Hotspots Based on Geospatial Techniques. *International Journal of Geo-Information*, Vol. 11(4), 260-277. https://doi.org/10.3390/ijgi11040260.

[29] Geurts, K., Wets, G. and Brijs, T., (2004). Identification and Ranking of Black Spots: Sensitivity Analysis, *Transportation Research Record*, Vol. 1897(1), 34-42. http://dx.doi.org/10.3141/1897-05.

[30] Le, K. G., Liu, P. and 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*, Vol. 23(2), 153-164. https://doi.org/10.1080/10095020.2019.1683437.

[31] Van, D. H., (2022). Factors Affecting the Sustainable Development of the Logistics Industry: A Study in Thanh Hoa Province, Vietnam. *Mathematical Statistician and Engineering Applications*, Vol. 71(3), 1625-1636. https://doi.org/10.17762/msea.v71i3.1349.

[32] Thammaboribal, P., TRIPATHI, N., Junpha, J., Lipiloet, S., and Wongpituk, K. (2024). Examining the Correlation between COVID-19 Prevalence and Patient Behaviors, Healthcare, and Socioeconomic Determinants: A Geospatial Analysis of ASEAN Countries. *International Journal of Geoinformatics*, Vol. 20(3), 95–112. https://doi.org/10.52939/ijg.v20i3.3159.

[33] Tippayanate, N., Impool, T., Sujayanont, P., Muttitanon, W., Chemin, Y., and Som-ard, J. (2024). Spatial Distribution and Cluster Analysis of Road Traffic Accidents in Khon Kaen Municipality, Thailand. *International Journal of Geoinformatics*, Vol. 20(4), 43–55. https://doi.org/10.52939/ijg.v20i4.3149.

[34] Kanav, A., Yadav, B., Sharma, R., and Kumar, J. (2024). Spatio-temporal Analysis of COVID-19 Hotspots in India Using Geographic Information Systems. *International Journal of Geoinformatics*, Vol. 20(1), 72–87. https://doi.org/10.52939/ijg.v20i1.3027.

[35] Besag, J. and Newell, J., (1991). The Detection of Clusters in Rare Diseases. *Journal of the Royal Statistical Society*, Vol. 154, 143-155. https://doi.org/10.2307/2982708.

[36] Nguyen, V. M., (2012). *Handling Administrative Violations in the Field of Road Traffic in Thanh Hoa Province*. Doctoral thesis, Faculty of Law, Hanoi National University.