International Journal of Basic and Applied Sciences, 14 (7) (2025) 213-218

SPC

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/1b78mm15
Research paper

Integrating Z-Score and GIS-Based Methods for Road Accident-Prone Segment Identification

Wahyu Satyaning Budhi [1], Wahyu Naris Wari [1], M. Shofi'ul Amin [1], Ahmad Utanaka [1] *, Hera Widyastuti [2]

[1] Department of Civil Engineering, Politeknik Negeri Banyuwangi, Banyuwangi, East Java, Indonesia
[2] Department of Civil Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, East Java, Indonesia
**Corresponding author E-mail: ahmad.utanaka@poliwangi.ac.id*

Received: October 20, 2025, Accepted: October 30, 2025, Published: November 6, 2025

Abstract

The rapid development of road infrastructure today presents significant challenges in ensuring road safety. The National Road III section from Kabat to the Banyuwangi City Boundary has shown a concerning trend in traffic accidents, highlighting the need to identify accident locations. This study employed secondary data obtained from the Banyuwangi Police traffic unit accident reports. The research analysis was conducted in two phases. First, the Z-score method was utilized to identify accident road segments through statistical evaluation. Subsequently, Hotspot Analysis (Getis-Ord Gi*) integrated with Geographic Information System (GIS) was applied to map and verify the spatial clustering of accidents. Both methods consistently identified two critical segments out of 25 road sections (each 200 meters in length). These were located at STA 5+200 – 5+400 with a 95% risk level and STA 8+400 – 8+600 with a 90% risk level, while the remaining segments did not show statistically significant clustering of accidents. The consistency of results across both methods strengthens the reliability of the findings, underscoring the importance of targeted safety interventions and accident mitigation strategies in these high-risk zones.

Keywords: *Geographic Information System (GIS); Hotspot Analysis; Traffic Accident; Z-Score Method.*

1. Introduction

Nowadays, infrastructure development in Indonesia is experiencing significant acceleration. [1]. This development trend aligns with global projections, which predict that developing countries will add 3.0–4.7 million kilometers of new roads by 2050 [2]. Enhanced connectivity, economic growth, and rapid industrial development serve as the main factors driving the expansion of road construction and its networks. Consequently, reductions in transportation costs and improvements in distribution efficiency contribute to greater mobility, the expansion of trade routes, and the advancement of economic growth. Between 1996 and 2021, Indonesia experienced rapid growth in road infrastructure development. This expansion not only increased the length of roads but also significantly enhanced the quality of existing road conditions. By 2021, 91.81% of the roads in Indonesia were reported to be in good condition. [3]. As a highly diverse country, Indonesia regards road connectivity as a key factor in linking rural areas with populated metropolitan centers. Therefore, improving road infrastructure is essential not only for sustaining an efficient and reliable transportation system but also for fostering long-term economic growth. [4].

The National Highway III, particularly the segment extending from Kabat District to the border of Banyuwangi City, is characterized by moderate traffic flow. This segment of road is utilized by several vehicle types, including unmotorized vehicles, motorcycles, light vehicles, and heavy vehicles, all of which significantly contribute to traffic volume. Nonetheless, this traffic volume exposes this segment to a heightened risk of accidents. Statistical data indicate an increasing trend in the number of accidents each year along this segment, which emphasizes the necessity for well-directed interventions. Safety is a fundamental element of sustainable road infrastructure, especially as the expanding road network increases the potential for accidents and conflicts between modes of transportation. Therefore, road safety management needs to be prioritized, particularly at intersections, railway crossings, and areas with moderate to high traffic volumes. Analyzing accident-prone segments is crucial for identifying vulnerable areas and formulating appropriate mitigation strategies to improve road safety.

In traffic safety studies, several terms are used to describe accident-prone areas, including blackspot, blacklink, and blackarea. Blacklink is one of the criteria in identifying and handling traffic accident-prone locations. Blacklink focuses on longer road segment with high accident rates, meanwhile blackspot only focuses on short segment such as intersections, bridges, or short road segments, meanwhile blackspot only focuses on short segment such as intersections, bridges, or short road segments. Blacklinks cover a larger amount of road segments vulnerable to accidents caused by traffic movement, geometric configurations, or road surface conditions, whilst Blackspots indicate locations of higher risk. While blackspot addresses high-risk areas, blacklink coverage is continuous road segments with similar characteristics that contribute to higher occurrences. Besides blacklink, the term black area is also used in road safety. The black area refers to a larger road network area where accident frequencies exceed the minimum limit [5]. Several factors contribute to the increase in traffic accidents on the road segments, including traffic volume, road design, and surface conditions [6]–[8].

Research on accident-prone areas, especially along national highways in developing countries, often relies only on statistical methods and ignores spatial analysis, which limits the ability to fully visualize accident clusters. In this study, the Z-score method was employed to statistically assess deviations in accident frequencies from the average accident rate of other segments along the same route. To complement this approach, Geographic Information Systems (GIS) were applied for spatial visualization and analysis. By combining Z-score and GIS-based Hotspot Analysis (Getis-Ord Gi), the study identifies and confirms accident-prone segments both statistically and spatially.

2. Method

2.1. Experimental objectives

The research location is along National Highway III, specifically from Kabat to the Banyuwangi City boundary, spanning a total length of 5 kilometers. The overall characteristics of the road are described in detail. Additionally, the location of the experimental site for this research can be seen in Fig. 1 and supplementary road data presented in Table 1.

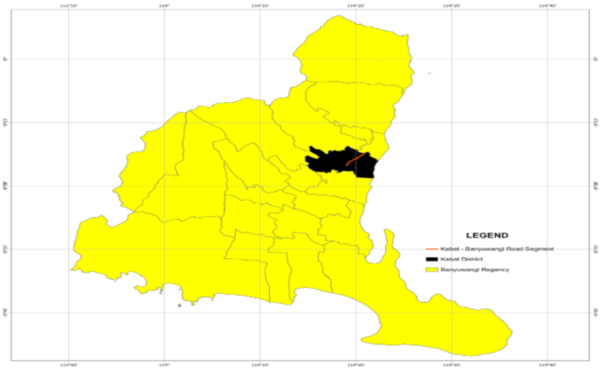

Fig. 1: Location of Experimental Objective.

Table 1: Road Data Details for Kabat - Banyuwangi City Boundary Segment

Data	Explanation
Road Type	Predominantly four-lane-two-way-undivided road (4/2UD)
Road Function	Primary Arterial
Road segment length	5 km
Average Lane Width	3,25 m
Median	Predominantly no median; the median is only present between STA 7+400 – 7+800
Road markings	Yes
Traffic signs	Yes

2.2. Z-score method

The Z-score, also known as the Z value or standardized value, represents a statistical measure indicating the distance of a data point from the sample mean. For a sample of size n with data points (X_1, X_2, X_3,..., X_n), with a certain mean and standard deviation S, as shown in equation (1), that have a mean of 0 and a standard deviation of 1. The Z value can then be calculated using equation (2) [9], [10].

$$S = \sqrt{\frac{\Sigma(X_i - \bar{X})^2}{n}} \quad (1)$$

$$Z_i = \frac{X_i - \bar{X}}{S} \quad (2)$$

Where S is the standard deviation, Xi is the average number of accidents per road segment, X is the overall average number of accidents, and n is the amount of data. Z_i is the Z-score value of accidents per road segment. Equation (3) is used to calculate the accident-prone class interval based on the Z-score, where I is the class interval, Z_H is the highest Z-score, Z_L is the lowest Z-score, and ΣI is the total number of intervals. [10].

$$I = \frac{Z_H - Z_L}{\Sigma I} \quad (3)$$

In addition, criteria based on Z-score values are used to assess accident risk. If the Z-score value is negative (below 0), the segment is considered not prone to accidents. [10]. Conversely, if the Z-score value is positive (above 0), the segment is considered accident-prone. After obtaining the Z-score value, the next step is to carry out an analysis using a quadrant graph, as can be seen in Fig. 2. This analysis aims to determine and measure the relationship between two variables [9].

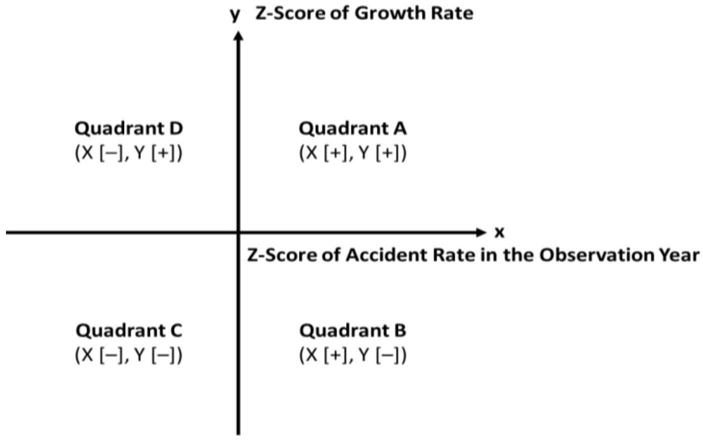

Fig. 2: Quadran Graph of the Relationship between the Z-Score of Accident Rate in the Observation Year and the Z-Score of Accident Growth Rate.

2.2. Getis-Ord GI* method

In this study, the Hotspot Analysis method in ArcGIS is used to test statistical significance using the Gi* or Local G method. Local G or Getis-Ord Gi* statistical method is widely applied to identify spatial clusters with statistically significant concentrations of high values (hot spots) and low values (cold spots) [11]. This local spatial statistic, commonly referred to as Gi*, evaluates the degree of spatial association between features, thus allowing the detection of areas where values exhibit clustering patterns rather than random distribution. Gi* is defined as (4) [12].

$$G_i(d) = \frac{\sum_j w_{ij}\,(d)x_j}{\sum_j x_j} \tag{4}$$

Where Gi* is a statistical method used to describe the spatial dependency of an incident i in relation to all n accident locations. In this approach, d represents the distance threshold, while x_j denotes the number of accidents at location j. Meanwhile, w_{ij} serves as the spatial weight that quantifies the degree of association between two accident locations, i and j, based on the Fixed Distance Band approach [11].

3. Result and Discussion

3.1. Traffic accident characteristics

The analysis of traffic accident characteristics categorized incidents into six groups, as shown in Table 2. Table 2 presents the distribution of traffic accident characteristics based on various parameters, including victim type, accident type, subjects involved, demographic factors such as victim age and gender, and time of accident occurrence.

Table 2: Accident Characteristics

No.	Parameters		Percentage (%)	No.	Parameters		Percentage (%)
1.	Accident Victims			4.	Gender		
	–	Death	16.3%		–	Male	75.8%
	–	Serious injuries	3.9%		–	Female	24.2%
	–	Slight injuries	79.7%	5.	Age		
2.	Types of Traffic Accidents				–	0 - 9 years old	0.0%
	–	Hit-and-run collision	13.9%		–	>9 - 19 years old	25.9%
	–	Pedestrian collision	9.9%		–	>19 - 29 years old	20.3%
	–	Multiple-vehicle collision	9.9%		–	>29 - 39 years old	11.6%
	–	Front-front collision	11.9%		–	>39 - 49 years old	14.2%
	–	Front-rear collision	12.9%		–	>49 - 59 years old	16.4%
	–	Side collision	32.7%		–	>59 - 69 years old	9.5%
	–	Single vehicle collision	8.9%		–	>69 - 79 years old	1.7%
3.	Subject				–	>79 years old	0.4%
	–	Pedestrian	6.9%	6.	Time of Accident Occurrence		
	–	Two-Wheeled Vehicles	71.6%		–	00.00-06.00	9.8%
	–	Four and More Wheeled Vehicles	20.6%		–	>06.00-12.00	38.2%
	–	Bicycle	0.5%		–	>12.00-18.00	35.3%
	–	Pedicab	0.0%		–	>18.00-00.00	16.7%
	–	Other subjects	0.5%				

Further analysis of traffic accident data reveals a high incidence rate among motorcyclists, with most injuries classified as minor. While motorcycles are overrepresented in accident statistics, the severity of injuries sustained is generally low. The highest percentages of traffic accidents were observed in the age categories 10–19 years and 20–29 years. In the 20–29-year age group, 20.3% of accidents were recorded, while the 10–19-year age group accounted for 25.9%. This indicates that most traffic accidents occur in the 10–19-year age range, which primarily includes students. It can be inferred that these accidents may result from limited driving experience, and drivers in this age group may also lack valid driving licenses. Furthermore, this data provides deeper insight into accident patterns, which is crucial for supporting risk analysis and the formulation of more targeted traffic safety policies.

3.1. Z-score value

The Z-score analysis for each road segment was conducted using traffic accident data from Jalan Nasional III, Kabat to Banyuwangi City Boundary, for three years. A total of 25 segments were analyzed using the Z-score method, which helped identify accident-prone locations, known as blacklinks. Table 3 provides a list of the road segments, Z-scores, and corresponding criteria.

Table 3: Accident Characteristics

Station							Z-score	Criteria
5	+	000	-	5	+	200	-0.20	Low-Risk Segment
5	+	200	-	5	+	400	0.75	High-Risk Segment with Major Accident
5	+	400	-	5	+	600	-0.40	Low-Risk Segment
5	+	600	-	5	+	800	-0.47	Low-Risk Segment
5	+	800	-	6	+	000	-0.06	Low-Risk Segment
6	+	000	-	6	+	200	-0.47	Low-Risk Segment
6	+	200	-	6	+	400	0.06	High-Risk Segment with Minor Accident
6	+	400	-	6	+	600	-0.24	Low-Risk Segment
6	+	600	-	6	+	800	-0.15	Low-Risk Segment
6	+	800	-	7	+	000	-0.33	Low-Risk Segment
7	+	000	-	7	+	200	-0.12	Low-Risk Segment
7	+	200	-	7	+	400	-0.22	Low-Risk Segment
7	+	400	-	7	+	600	0.55	High-Risk Segment with Moderate Accident
7	+	600	-	7	+	800	0.43	High-Risk Segment with Moderate Accident
7	+	800	-	8	+	000	-0.33	Low-Risk Segment
8	+	000	-	8	+	200	-0.47	Low-Risk Segment
8	+	200	-	8	+	400	0.38	High-Risk Segment with Moderate Accident
8	+	400	-	8	+	600	0.62	High-Risk Segment with Major Accident
8	+	600	-	8	+	800	0.24	High-Risk Segment with Minor Accident
8	+	800	-	9	+	000	0.32	High-Risk Segment with Moderate Accident
9	+	000	-	9	+	200	0.13	High-Risk Segment with Minor Accident
9	+	200	-	9	+	400	-0.01	Low-Risk Segment
9	+	400	-	9	+	600	0.13	High-Risk Segment with Minor Accident
9	+	600	-	9	+	800	-0.05	Low-Risk Segment
9	+	800	-	10	+	000	-0.12	Low-Risk Segment

The results above indicate that road segments with a high positive Z-score are likely to experience major accidents, such as those in STA 5+200 – 5+400 and STA 8+400 – 8+600. Therefore, these segments can be a priority for road safety interventions. Whereas, segments with a negative Z-score or close to zero indicate low risk but still require continuous monitoring.

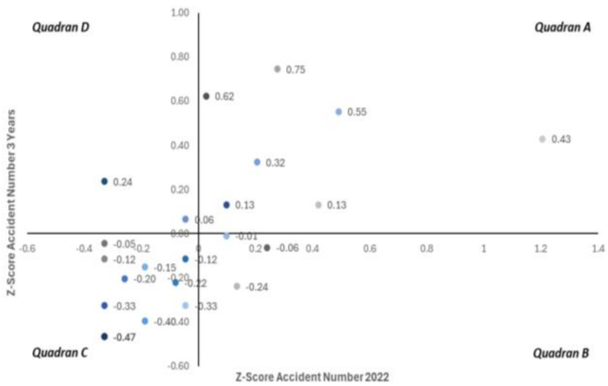

Fig. 3: Graph of The Relationship between Z-Score Value of Accident Number Based on 2022 with Z-Score of Accident Number Growth Per 3 Years.

Following the calculation of the Z-scores, a quadrant graph was used to examine the relationship between two variables, as shown in Fig. 3. These variables were the Z-score of traffic accidents in 2022 and the Z-score of the increase in traffic accidents over the three years. Quadrant A contains 7 locations with accident rates significantly higher than the average for the entire analyzed road section. These areas in Quadrant A may require immediate attention for traffic safety measures, as they represent areas with a high and growing crash risk. Quadrant B includes 3 locations with high accident rates, but the rate of accident increase is below the overall average for the reviewed road section. Although accidents in these areas remain high, the slower rate of increase indicates a relatively stable situation. The focus in these areas could be more on maintaining or improving existing safety measures. Quadrant C consists of 10 locations with low accident rates. While these areas do not require urgent intervention, ongoing monitoring is necessary to ensure that these segments do not develop into problems in the future. Quadrant D includes two locations with decreasing accident rates, but the rate of increase is above the overall average. This segment may require preventive measures to prevent future increases in accidents, even though the current accident rate is lower.

3.2. Hotspot analysis

The analysis of traffic accident-prone locations was conducted using the Hot Spot Analysis (Getis-Ord Gi*) feature. This analysis was carried out spatially, producing clusters with either high or low values. The Gi* statistic evaluates data in the form of z-scores. This feature provides an assessment of the number of accident events grouped spatially. The results of the hotspot analysis using ArcGIS can be seen in Fig. 4. The image shows that accident-prone areas are highlighted in red, whereas areas that are not prone to traffic accidents are shown in blue.

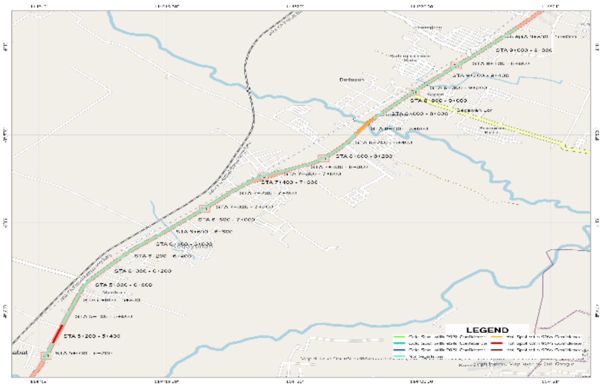

Fig. 4: Hotspot Map.

Hotspot analysis in GIS is a form of data exploration and analysis that falls under Exploratory Spatial Data Analysis (ESDA), where hotspot analysis specifically focuses on the characteristics of spatial autocorrelation and spatial heterogeneity. The Hot Spot Analysis (Getis-Ord Gi*) in GIS accurately identifies statistically significant spatial clusters of hot spots (high values) and cold spots (low values).
In this study, accident locations were divided into 25 segments, with each segment measuring 200 meters in length. The results of the GIS hotspot analysis revealed that two segments were categorized as accident hotspots. These segments were located at STA 5+200 – 5+400 and STA 8+400 – 8+600, while the remaining segments were not statistically significant as hotspots. The segment at STA 5+200 – 5+400 indicated a hotspot analysis value of 95%, representing a 95% level of traffic accident risk. Meanwhile, the segment at STA 8+400 – 8+600 showed a value of 90%, indicating a 90% level of accident risk. The other segments indicated either insignificant or low levels of traffic accident risk.

3.3. Discussion

In this study, the Z-score method and GIS-based Hotspot Analysis were used to identify accident-prone road segments. These two methods complement each other in the analysis process, providing a more holistic picture of the accident distribution along the analyzed road section. The Z-score method measures the extent to which the accident frequency on each road segment deviates from the average accident rate on other segments. A high Z-score indicates a segment with more accidents than other segments, and is therefore more prone to accidents. This process provides early identification of segments requiring increased attention. Conversely, Hotspot Analysis using the Getis-Ord Gi* method in GIS identifies and maps spatially significant accident clusters. This analysis clearly demonstrates the spatial distribution of accidents, both in the form of high-frequency accident clusters (hotspots) and low-frequency accident clusters (coldspots). Both methods demonstrate consistent results in identifying accident-prone segments. The high Z-score results for several segments were confirmed by Hotspot Analysis, which also indicated the presence of accident clusters in the same locations. Thus, the combination of these two methods provides stronger validation of the identification of accident-prone road segments, both from a statistical and spatial perspective. Furthermore, several potential factors may explain the high-risk segments identified in this study. Segments with a high crash frequency are typically characterized by limited visibility due to vertical curvature changes, inadequate lighting, vehicle speeds, and mixed traffic composition involving heavy vehicles and motorcycles. These conditions can reduce driver visibility and reaction time, increasing the likelihood of crashes. To mitigate these risks, targeted engineering and management interventions are recommended. For example, improving street lighting, adding warning signs, implementing traffic calming measures such as rumble strips, and stricter speed limit enforcement can significantly improve safety in these segments. Collaboration between local governments, transportation agencies, and law enforcement is crucial to ensure these measures are effectively implemented.
While the results provide valuable insights, several limitations should be acknowledged. This study has several limitations that need to be considered, in addition to the unreported accident data to the relevant authorities. The analysis was based on accident data without considering temporal variations such as weather or the timing of the incident, which may influence accident frequency. Furthermore, the level of detail in the available data is limited, and several external factors, such as the physical condition of the road infrastructure and road geometry, were not included in the analysis. These factors may affect the accuracy of identifying accident-prone locations. Future research is recommended to integrate more detailed temporal and environmental data and implement predictive models for real-time accident risk monitoring.

4. Conclusion

This study identified several accident-prone road segment locations along the National Road III Kabat - Banyuwangi City Border section using a combined Z-score and GIS-based Hotspot Analysis approach. These findings also confirmed the correlation between the Z-score analysis and Hotspot Analysis, supporting the validity of the identification of accident-prone segments. The results of this study can be used by stakeholders to formulate appropriate measures to reduce accident risk, particularly major accidents, in these segments. In addition, it is important to conduct further review of factors that contribute to accidents, such as road pavement conditions, geometric design, and other factors that may affect traffic safety. As well as developing a real-time monitoring-based predictive modeling approach.

Acknowledgement

We express our gratitude to the Banyuwangi Police Traffic Unit for providing us with traffic accident data for this research.

References

[1] W. S. Budhi, A. Utanaka, I. K. H. Wiryasuta, and H. Widyastuti, "Identifying Traffic Accident Trends and Black Spot Locations on National Road (A Case Study: Rogojampi-Kabat, Banyuwangi)," in *International Conference on Architecture and Civil Engineering Conference*, 2023, pp. 683–695. https://doi.org/10.1007/978-981-97-0751-5_60.

[2] J. R. Meijer, M. A. J. Huijbregts, K. C. G. J. Schotten, and A. M. Schipper, "Global patterns of current and future road infrastructure," *Environ. Res. Lett.*, vol. 13, no. 6, p. 64006, 2018, https://doi.org/10.1088/1748-9326/aabd42.

[3] S. D. Prihapsari *et al.*, *Informasi Statistik Infrastruktur PUPR 2022*. Jakarta: Kementerian Pekerjaan Umum & Perumahan Rakyat, 1022.

[4] H. Widyastuti and W. S. Budhi, "Railway capacity analysis using Indonesian method and UIC code 405 method," in *IOP Conference Series: Materials Science and Engineering*, 2020, vol. 930, no. 1. https://doi.org/10.1088/1757-899X/930/1/012059.

[5] Direktorat Keselamatan Transportasi Darat, *Pedoman Operasi Accident Blackspot Investigation Unit/ Unit Penelitian Kecelakaan Lalulintas (Abiu/Upk)*. Jakarta: Direktorat Jenderal Perhubungan Darat, 2007.

[6] Machsus, I. Prayogo, Chomaedhi, D. W. Hayati, and A. Utanaka, "Road safety analysis on Achmad Yani frontage road Surabaya," in *IOP Conference Series: Materials Science and Engineering*, 2017, vol. 267, no. 1, pp. 1–11. https://doi.org/10.1088/1757-899X/267/1/012027.

[7] I. W. Agustin, C. Meidiana, and S. Muljaningsih, "Studi Simulasi Model Kecelakaan Pengendara Mobil untuk Meningkatkan Keselamatan Lalu Lintas di Daerah Perkotaan," *War. Penelit. Perhub.*, vol. 32, no. 2, pp. 93–102, 2020, https://doi.org/10.25104/warlit.v32i2.1513.

[8] A. Utanaka and W. S. Budhi, "Identifying Accident Severity Factor on Surabaya Secondary Arterial Roads," *Int. J. Transp. Infrastruct.*, vol. 6, no. 1, pp. 30–37, 2022.

[9] V. N. Arung and H. Widyastuti, "Penentuan Daerah Rawan Kecelakaan Lalu Lintas di Kota Surabaya," *J. Apl. Tek. Sipil*, vol. 18, no. 1, p. 17, 2020, https://doi.org/10.12962/j2579-891X.v18i1.5328.

[10] S. B. Pamungkas, A. MHM, and Setiono, "Analisis Lokasi Rawan Kecelakaan Di Jalan Arteri Primer Kota Surakarta," *Matriks Tek. Sipil*, vol. 5, no. 4, pp. 1199–1206, 2017.

[11] L. Srikanth and I. Srikanth, "A Case Study on Kernel Density Estimation and Hotspot Analysis Methods in Traffic Safety Management," in *International Conference on Communication Systems and NETworkS (COMSNETS)*, 2020, no. January, pp. 99–104. https://doi.org/10.1109/COMSNETS48256.2020.9027448.

[12] D. L. Srikanth, I. Srikanth, and D. M. Arockiasamy, "Identification of Traffic Accident Hotspots using Geographical Information System (GIS)," *Int. J. Eng. Adv. Technol.*, vol. 9, no. 2, pp. 4429–4438, 2019, https://doi.org/10.35940/ijeat.B3848.129219.