Japan Bilingual Publishing Co.

Transportation Development Research

https://ojs.bilpub.com/index.php/tdr

ARTICLE

A Composite Geospatial Index for Road Safety Risk: Integrating Crash Data with Roadway Characteristics

Siqing Chen

Melbourne School of Design, Faculty of Architecture Building and Planning, University of Melbourne, Parkville 3010, Australia

ABSTRACT

Road traffic accidents pose significant threats to public safety and urban infrastructure. While effective safety management is a critical component of sustainable transportation planning and public health, traditional approaches often rely heavily on identifying historical crash hotspots. These reactive methods frequently fail to account for the intrinsic environmental risk factors or overlook segments where crash frequency is low, yet the potential for severe outcomes remains high. To address this limitation, this paper presents a proactive, multi-factor geospatial model for calculating a comprehensive Road Safety Risk index at the individual road-segment level. Utilizing the City of Manningham, Victoria, as a case study, the research employs a GIS framework to synthesize official road network and historical crash data. The model incorporates four distinct risk dimensions: (1) accident frequency normalized by segment length; (2) a weighted accident severity index prioritizing serious incidents; (3) a normalized Speed Zone Factor; and (4) a Road Class Factor accounting for the functional hierarchy of the road network. The resulting risk map provides a granular and nuanced visualization of risk distribution, clearly identifying high-risk arterial corridors and intersections. Crucially, the analysis highlights road segments that, despite lower crash locations, pose significant threats due to a confluence of high speeds, road function, and crash severity. This replicable model serves as a powerful evidence-based tool for transport authorities, enabling a paradigm shift from reactive mitigation to proactive safety management.

Keywords: Road Safety; Risk Assessment; Geospatial Analysis; Traffic Accidents; Transport Planning; Manningham

*CORRESPONDING AUTHOR:
Siqing Chen, Melbourne School of Design, Faculty of Architecture Building and Planning, University of Melbourne, Parkville 3010, Australia; Email: chens@unimelb.edu.au

ARTICLE INFO
Received: 15 December 2025 | Revised: 19 February 2026 | Accepted: 26 February 2026 | Published Online: 4 March 2026
DOI: https://doi.org/10.55121/tdr.v4i1.1024

CITATION
Chen, S., 2026. A Composite Geospatial Index for Road Safety Risk: Integrating Crash Data with Roadway Characteristics. Transportation Development Research. 4(1): 12–27. DOI: https://doi.org/10.55121/tdr.v4i1.1024

1. Introduction

Safe and efficient flows of people and goods in cities are a foundational pillar of modern urban life. However, the persistent challenge of road traffic accidents represents a significant threat to this ideal, imposing immense social, economic, and human costs on communities worldwide. Road safety reports present a serious global problem: annually, around 1.35 million people are victims of fatal general traffic crashes [1]. In the context of contemporary urban planning, the pursuit of road safety extends far beyond mere incident reduction; it is intrinsically linked to the broader goals of creating sustainable, liveable, and healthy cities [2–4]. The perception of traffic danger, particularly for vulnerable road users such as pedestrians and cyclists, directly influences travel behavior [5,6]. Environments perceived as unsafe actively discourage active transport modes like walking and cycling, leading to increased car dependency, reduced physical activity, and negative environmental outcomes [7]. Consequently, a paradigm shift is underway in transportation planning, moving away from a primary focus on vehicle mobility towards a more holistic, systems-based approach that prioritizes the safety and accessibility of all road users [8]. However, this shift also necessitates the development of specialized, data-driven tools capable of diagnosing different types of risk across the complex hierarchy of the urban road network.

The current state of research and practice in road safety analysis has traditionally relied on reactive methods, most notably the spatial identification of crash hotspots and black spots [1,5,9,10] using techniques like Kernel Density Estimation [11,12]. These techniques, which map clusters of historical incidents, are valuable for identifying locations with a proven history of failure. However, this approach has critical limitations. It is inherently reactive, requiring a pattern of crashes to emerge before a location is flagged for intervention [8]. Furthermore, it often fails to explain the underlying causal factors, treating the road network as a passive backdrop rather than an active contributor to risk. This overlooks the reality that the intrinsic characteristics of a road segment, such as its designed speed, functional classification and geometry, are powerful determinants of both crash likelihood and severity [13]. This is particularly true for the primary road network, where high speeds and traffic volumes create a fundamentally different risk environment compared to local residential streets. The research field, therefore, faces a critical need for proactive models that can assess risk holistically by integrating historical crash data with the inherent risk factors of the road environment itself.

This paper addresses this gap by developing and applying a multi-factor geospatial model specifically designed to identify high-risk corridors within the primary road network, where the majority of fatal and serious injury crashes occur. The model's scope is therefore predominantly vehicle-centric, focusing on the systemic risks inherent in the design and operation of arterial roads and freeways. The model moves beyond simple incident counts to incorporate four distinct variables: normalized accident frequency, weighted accident severity, and two key environmental risk indicators—a Speed Zone Factor and a Road Class Factor. The primary aim of this work is to demonstrate the efficacy of this composite model as a tool for proactive safety management, using the City of Manningham, Victoria, as a case study. Through the successful application of this methodology, a nuanced and actionable risk map was developed, showcasing the model's ability to provide transport planners with a more comprehensive understanding of where and why risks exist on the urban road network.

2. Literature Review

Following the paradigm shift from reactive hotspot identification to proactive risk assessment, a substantial body of research has emerged to understand the multifaceted nature of road traffic safety. This literature can be broadly categorized into three key areas that directly inform the development of a multi-factor risk model: the application of geospatial techniques to uncover crash patterns, the identification of specific roadway and environmental risk factors, and the advancement of statistical methods for modeling crash data.

Geographic Information Systems (GIS) have become an indispensable platform for road safety analysis, providing the foundational tools to store, manage, and visualize the spatial components of crash data [14]. The most fundamental application of GIS in this field is the identification of accident hotspots. Early approaches relied on simple pin maps, but methodologies have become increasingly sophisticated. Kernel Density Estimation (KDE) is a widely used technique that moves beyond discrete points to create a continuous surface map of crash densities, offering an intuitive visualization of where incident concentrations are highest [9,15]. While powerful for visualization, KDE does not inherently confirm the statistical significance of these clusters. To address this, researchers frequently employ spatial statistical methods like the Getis-Ord Gi* statistic. This tool formally identifies statistically significant hotspots and cold spots (clusters of low values), allowing analysts to distinguish between true spatial patterns and random chance [10,16]. These geospatial techniques provide the critical "where" that guides further investigation into the "why".

A primary focus of proactive safety research is to understand how the physical design of the road network contributes to crash risk. A large body of work has established clear links between road geometry and accident frequency and severity [17]. Key factors include the horizontal and vertical alignment of the road, such as the radius of curves and the steepness of grades, which directly impact vehicle dynamics and driver behavior [18]. Cross-sectional elements have also been extensively studied; variables such as the number of lanes, lane width, shoulder width, and the presence or absence of medians are known to significantly influence safety outcomes [19,20]. Specific interventions, such as the application of perceptual road markings in curves, have been shown to positively affect driver speed and lateral position, thereby reducing risk [21]. From an operational perspective, traffic characteristics such as speed and traffic volume are fundamental. Area-wide average speed, for instance, has been found to be positively associated with total fatalities and serious injuries [22]. These metrics also serve as crucial measures of exposure; normalizing crash counts by vehicle-kilometers traveled is a standard practice that provides a more accurate assessment of risk by accounting for the level of activity on a given roadway [23].

For analyzing the outcome of a crash, injury severity models are employed. These models must account for the discrete and ordered nature of the dependent variable (e.g., no injury, slight injury, serious injury, fatal). Ordered-response models, such as the ordered logit and ordered probit, are frequently used for this purpose as they explicitly recognize this inherent ranking [13]. More advanced techniques, like mixed logit models, further enhance the analysis by accounting for unobserved heterogeneity across the crash population [24]. To address the fact that the influence of certain risk factors may vary geographically, researchers have also adopted Geographically Weighted Regression (GWR). This local form of regression allows model parameters to differ across space, providing a more nuanced understanding of how relationships, such as that between speed and crash risk, might change from an urban to a rural context [25]. A persistent challenge across all methodologies, however, is the quality and completeness of official accident data, which is often subject to significant under-reporting, especially for non-injury or minor-injury crashes [26].

The literature provides a robust foundation for proactive road safety analysis, confirming that GIS is an essential platform, that specific roadway and environmental factors are proven to influence risk, and that advanced statistical models are required to analyze crash data accurately. While many studies have provided deep insights into individual risk factors or specific methodologies, a gap often exists in synthesizing these elements into a single, comprehensive, and practical framework for use by transport planners. The contribution of the proposed research, therefore, is not necessarily the discovery of a novel risk factor, but rather the development of a transparent and replicable multi-factor model that integrates several well-established indicators—Accident Frequency, Accident Severity, Speed Zone Factor, and Road Class Factor. By combining these variables into a composite Road Safety Risk index, this approach offers a holistic tool that can be practically applied to proactively identify hazardous corridors and support data-driven decision-making in urban road safety management [27].

3. Materials and Methods

This study develops and applies a multi-factor geospatial model to assess road safety risk in an urban environment. The methodology comprises three primary stages: (1) data acquisition and pre-processing, (2) modeling of the road safety risk components, and (3) calculation and visualization of a final composite risk index. The entire workflow was implemented using ArcGIS Pro geoprocessing tools and its integrated Python environment.

3.1. Study Area

The model was applied to the City of Manningham, a municipality located approximately 15 km east of Melbourne's Central Business District in Victoria, Australia (**Figure 1**). The City of Manningham was selected as an ideal case study for several key reasons. Primarily, its urban form is representative of many suburban municipalities globally, featuring a heterogeneous road network. This network includes a diverse mix of high-speed, high-volume freeways and arterial roads (such as the M3 Eastern Freeway), intermediate collector roads, and an extensive network of low-speed local streets in residential areas. This variety in road classification and design speed is essential for rigorously testing a model that explicitly uses a Road Class Factor and a Speed Zone Factor as core components. Furthermore, the availability of high-quality, publicly accessible, and spatially accurate datasets from the Victorian Government Data Directory was a critical factor. The VicMap Transport and Victoria road crash datasets provided the necessary granular data on road geometry, speed limits, and historical incidents required to calculate each of the model's four input variables. This combination of a typical suburban road hierarchy and reliable, open-source data makes Manningham a suitable and powerful testbed for a methodology intended to be replicable and applicable in other similar jurisdictions that face common road safety challenges.

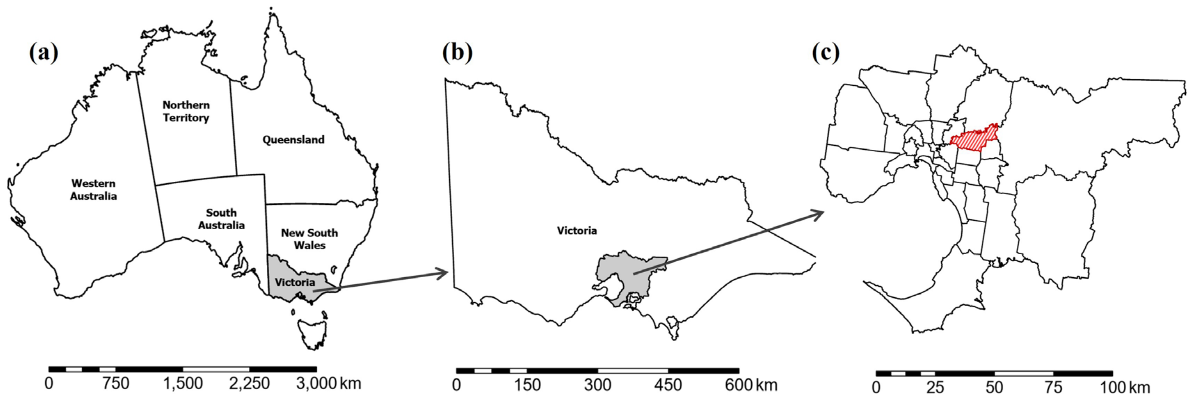

Figure 1. Location of the Study Area: (**a**) Location of Victoria in Australia; (**b**) Location of Melbourne Metropolitan in Victoria; (**c**) Location of the City of Manningham in Metropolitan Melbourne.

3.2. Data Sources and Pre-Processing

Three primary datasets, all publicly available from the Victorian Government Data Directory, were used for this research. All spatial datasets were projected into the GDA 1994 VICGRID94 coordinate system to ensure spatial consistency.

1) Administrative Boundary: The polygon shapefile for the City of Manningham was sourced from the VicMap Admin dataset. This dataset defines the primary boundary used to create the buffered study area.
2) Road Network: A polyline shapefile of the metropolitan road network was obtained from the VicMap Transport dataset. This dataset contains the geometric information and key attributes for each road segment, including the Class_Code used for the Road Class Factor.
3) Traffic Accident Data: A point shapefile of historical traffic incidents was derived from the Victoria road crash dataset. This dataset includes location-specific data for each incident, including the SEVERITY of the crash and the SPEED_ZONE in which each incident occurs.

Initial pre-processing involved clipping the road network and traffic accident datasets to the buffered study area boundary using the Clip tool in ArcGIS Pro. Following this, a data cleansing step was performed on the clipped traffic accident dataset. Records containing invalid SPEED_ZONE values (specifically codes 777, 888, and

999, which denote unknown or inapplicable speed limits) were removed to ensure the integrity of subsequent calculations. This was achieved using the Make Feature Layer tool with a SQL expression, followed by the Copy Features tool to create a cleansed, permanent dataset.

3.3. Modelling Road Safety Risk

Four factors associated with urban road safety, namely Accident Severity, Accident Frequency, Road Class Factor, and Speed Zone Factor in Manningham (**Figure 2**), are used to calculate Road Safety Risk.

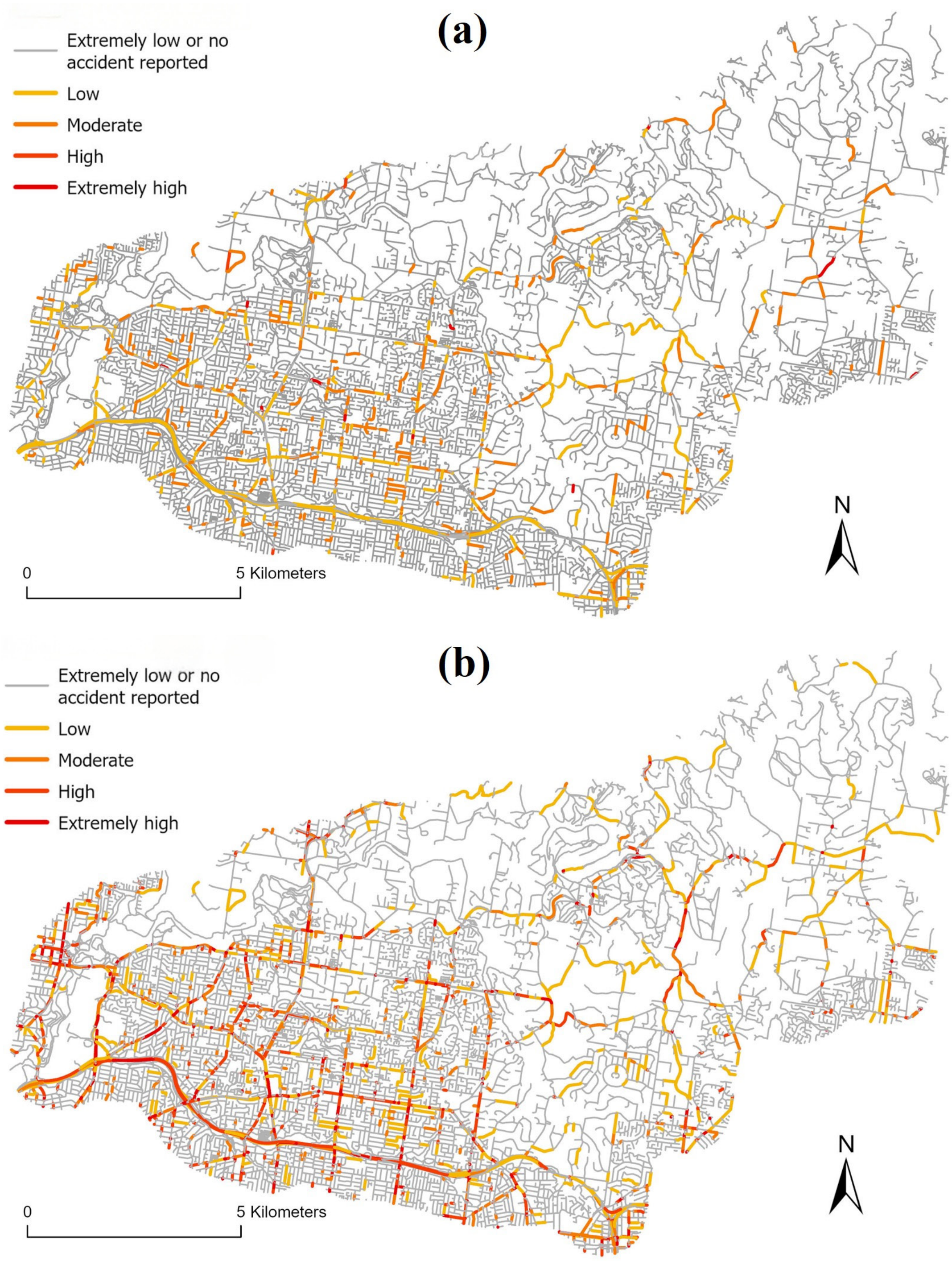

Figure 2. *Cont.*

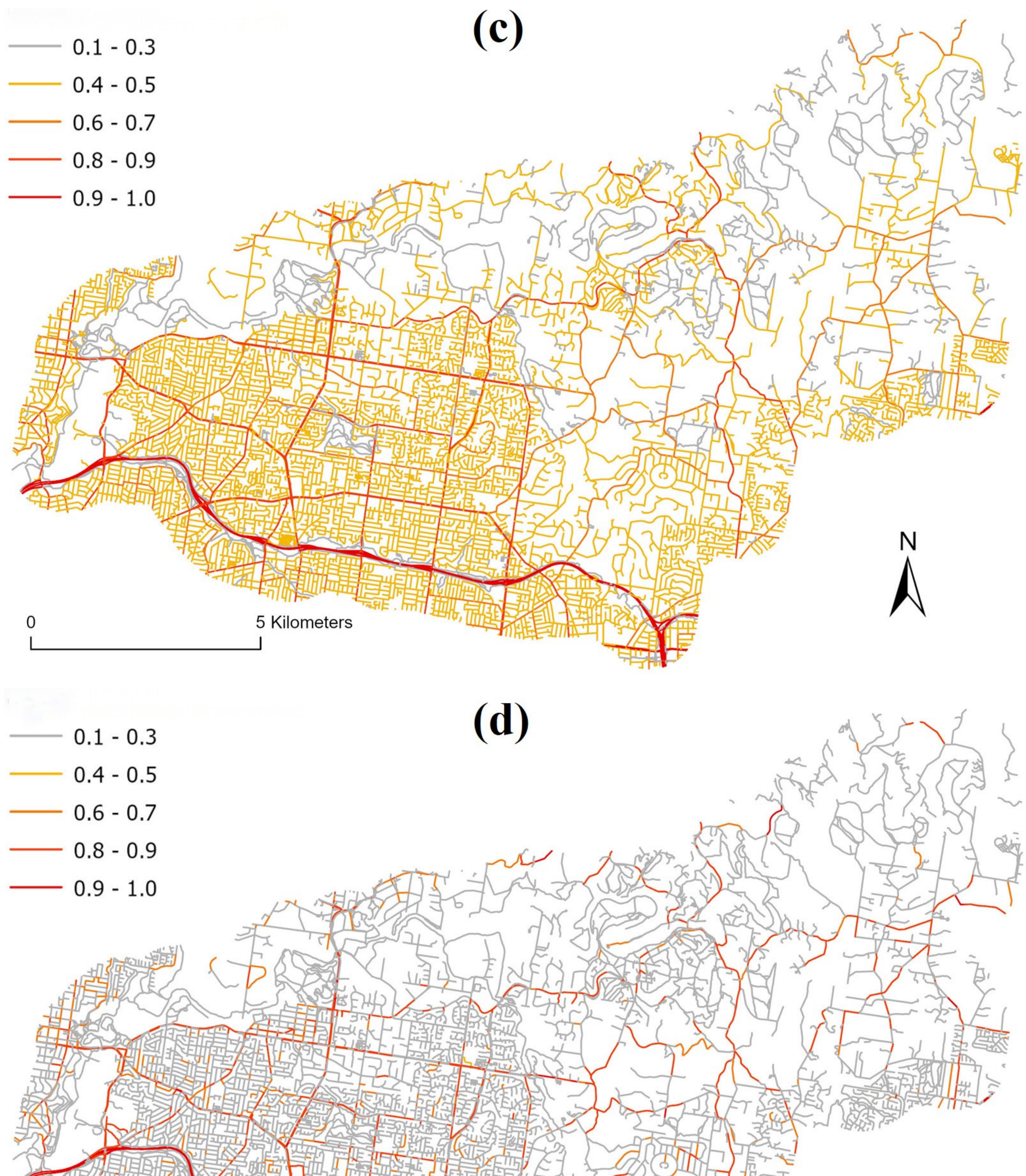

Figure 2. The four input factors in the City of Manningham for calculating the Road Safety Risk index: (**a**) Accident Severity; (**b**) Accident Frequency; (**c**) Road Class Factor; (**d**) Speed Zone Factor.

The core of the methodology is the calculation of a composite Road Safety Risk index for each individual road segment. The model is based on the following conceptual framework:

$$RSR = AS \times AF \times f_{SZ} \times f_{RC} \quad (1)$$

where RSR is road safety risk index; AS is traffic accident severity; f_{SZ} is road speed zone factor; and f_{RC} is road class factor. Each of the four components of the risk index was calculated as a new field in the attribute table of the joined road feature class using the Add Field and Calculate Field tools in the ArcGIS Pro 3.5 environment.

1) Accident Severity: The original SEVERITY attribute (1: Fatal, 2: Serious injury, 3: Other injury, 4: Non-injury) was recalculated to create a scoring system where higher values reflect greater severity. This inversion ensures that more severe accidents contribute more heavily to the final risk index. The final Accident Severity (AS) score for each road segment is calculated as the average of the recalculated severity scores of all accidents that occurred on it. This approach was chosen to provide a stable measure of the typical incident outcome on a segment, preventing its risk score from being skewed by a single, anomalous event.
2) Accident Frequency: To account for exposure, the raw accident count on each segment (FREQUENCY) was normalized by the segment's length. This variable, Accidents_per_km, prevents longer road segments from being unfairly penalized simply due to their length.
3) Speed Zone Factor: The average speed limit of each road segment (MEAN_SPEED_ZONE) was normalized by the highest speed limit in the study area (100 km/h). This creates a standardized factor between 0 and 1 that represents the influence of the speed environment.
4) Road Class Factor: To incorporate the inherent risk associated with different road types in the hierarchy, the Class_Code attribute was converted into a normalized factor. The formula assigns a higher risk weight to higher-order roads (e.g., freeways) and a lower weight to local roads, tracks and trails (see details in **Table 1**). The values assigned to f_{RC} reflect empirically observed differences in crash risk across functional classes and contexts. Similar methodology in the literature also includes roadway functional class [28], road classes [29], or road type [30] as a key indicator and assigns different scores/weights to motorways, provincial roads, and local roads in studies related to transport safety.

Table 1. Road Class Factor weights assigned to different road hierarchy levels.

Road Hierarchy	Class Code	Assigned Road Class Factor (f_{RC})
Freeway	0	1.0
Highway	1	0.9
Arterial	2	0.8
Sub-Arterial	3	0.7
Collector Road	4	0.6
Local Road	5	0.5
Minor Road	6	0.4
Major Track	7	0.3
Minor Track	8	0.2
Trail	9	0.1

The selection of a multiplicative model, as expressed in Equation (1), is grounded in the conceptual understanding that road safety risk arises from the synergistic interaction of multiple factors. Unlike an additive model, which is inherently compensatory (where a low score in one factor could be offset by a high score in another), a multiplicative framework ensures that risk factors amplify one another. For instance, a road segment with a high accident frequency becomes exponentially more dangerous when combined with a high-speed environment; the risks do not simply sum. Crucially, this approach logically handles instances where a key risk component is absent. If a road segment has zero recorded accidents (AF = 0), the composite RSR score correctly becomes zero, reflecting that no matter the inherent environmental risks from speed or road class, the demonstrated risk is non-existent. An additive model would fail to capture this important nuance, potentially assigning a moderate risk score to a location with no history of safety issues.

3.4. Final Road Safety Risk Index Calculation

With all four input variables—Accident Severity, Accident Frequency, Speed Zone Factor, and Road Class Factor—calculated for each road segment, the final composite Road Safety Risk (RSR) index is computed. This is achieved by multiplying the four normalized factors as defined in Equation (1), a process executed using a Python script within the GIS environment.

A critical aspect of this calculation is the model's handling of "zero-value" scenarios, most commonly occurring on road segments with no historical accidents. In such cases, the values for both Accident Frequency and Accident Severity are zero. Due to the multiplicative nature of the formula, the resulting composite RSR index for that segment is logically and correctly calculated as zero. This outcome is not merely a technical step to prevent calculation errors, but a deliberate and fundamental feature of the model's design. This approach ensures that the RSR index prioritizes locations where inherent risk (posed by high speeds or road class) has already translated into demonstrated risk (in the form of historical crashes). While a segment on a high-speed arterial road might have high values for its Speed Zone and Road Class factors, if it has no crash history, the model correctly assigns it a negligible risk score. This logic allows the model to effectively distinguish between latent potential for danger and locations with proven, empirical evidence of safety failures, thereby providing a more focused and actionable tool for prioritizing interventions. The Python function simply operationalizes this logic, ensuring that any segment with a null or zero value for its accident-related inputs is assigned a final RSR index of 0.

4. Results

The application of the multi-factor geospatial model and its associated data processing workflow, as described in the Materials and Methods section, resulted in the calculation of a composite Road Safety Risk index for each road segment within the City of Manningham. This section presents a description and interpretation of the key findings, culminating in a geospatial visualization of the final risk distribution.

4.1. Spatial Distribution of Road Safety Risks

Based on Equation (1), the primary output of this study is a geospatial risk map that visualizes the final calculated Road Safety Risk index for every road segment. **Figure 3** presents this spatial distribution, with road segments categorized and color-coded into five quantile classes ranging from "Extremely Low Risk" to "Extremely High Risk". The insert map in **Figure 3** shows details of in the RSR score in a major intersection, clearly demonstrating the granular details of RSR scores of different road segments with different road classes, speed limit, accident frequency, and accident severity. For example, the RSR of a section of the Eastern Freeway within the insert map is calculated as follows: RSR = 2.286 (Accident Severity) × 1.380 (Accident Frequency) × 0.610 (Speed Zone Factor × 1.000 (Road Class Factor) = 1.924, which is "Extremely High" (See **Table 2** for the RSR range for each risk category).

A distinct spatial pattern is immediately evident from the map. The highest-risk scores (depicted in dark red and orange) are predominantly concentrated along the freeway, major arterial, and collector roads that form the primary transportation corridors within the study area. These high-risk segments create clear, linear patterns that trace the main thoroughfares. Conversely, an extensive and dense network of local and residential streets consistently displays low to extremely low risk scores (depicted in blue and green). These low-risk areas form the fabric of the residential neighborhoods between the major corridors. Segments with moderate risk scores (depicted in yellow) typically appear on secondary collector roads that serve as transitions between the major arterials and the local street network, or on sections of arterial roads with fewer contributing risk factors.

A statistical summary of the total road length within each of the five risk categories is presented in **Table 2**. The analysis reveals a highly skewed distribution of risk across the network's total length of 1539.55 km. The vast majority of the road network, totaling 1370.77 km or 89.04% of the total, is classified as "Low Risk". In stark contrast, the categories representing elevated risk are geographically concentrated within a much smaller portion of the network. The "Moderate", "High", and "Extremely High" risk categories collectively account for just over 10% of the total road length, comprising 46.77 km (3.04%), 55.86 km (3.63%), and 46.83 km (3.04%), respectively. The "Extremely Low" risk category is similarly limited, covering only 19.32 km (1.25%). This distribution underscores a key experimental conclusion: while the highest-risk areas are of significant concern for safety interventions, they represent a small and geographically targeted fraction of the overall road network in Manningham. This implies that safety resources can be directed with a high degree of spatial precision to the relatively few corridors where risk is most concentrated.

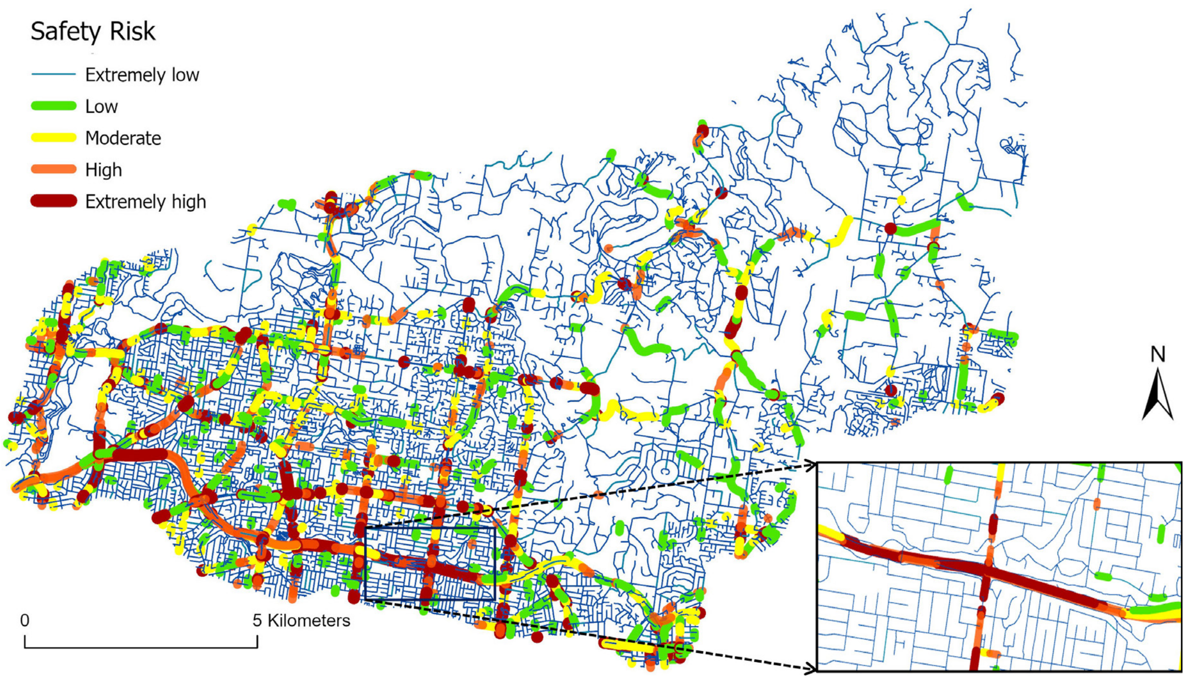

Figure 3. Road Safety Risk for City of Manningham.

Note: The inserted map clearly shows variation in risk scores at a segment-by-segment level for a major intersection (Eastern Freeway and Black Road), directly demonstrating the granular details.

Table 2. Length of roads of each risk category.

Road Safety Risk Category	RSR Score Range (Quantile)	Sub Total Road Length (km)	Percentage of Total Road Length
Extremely Low	<0.05	19.32	1.25%
Low	0.05–0.09	1370.77	89.04%
Moderate	0.09–0.17	46.77	3.04%
High	0.17–0.42	55.86	3.63%
Extremely High	>0.42	46.83	3.04%
Total	--	1539.55	100.00%

4.2. Interpretation of Results

The spatial patterns of the simulated Road Safety Risk scores indicate that the model successfully synthesized the underlying risk dimensions of the road network. The high-risk scores observed on arterial roads are a direct function of the confluence of all four input variables. These corridors inherently carry higher traffic volumes, which leads to a greater Accident Frequency. The higher operational speeds on these roads contribute to a higher average Accident Severity when crashes do occur. Furthermore, these same roads are assigned a greater weight from both the Speed Zone Factor and the Road Class Factor, amplifying their contribution to the final composite index.

In contrast, the low-risk scores on residential streets reflect the opposite conditions. These roads are characterized by near-zero or very low accident frequencies, lower posted speed limits, and a minimal weighting from the road class factor. Even if an isolated, minor incident were to occur on such a street, the low values of the other factors ensure that the segment is not incorrectly classified as high-risk.

Crucially, the model highlights that risk is not merely a function of historical accident counts alone. For example, a segment with a moderate number of non-injury accidents on a high-speed arterial road may ultimately receive a higher risk index than a segment with a slightly higher number of similar accidents on a low-speed local street. This demonstrates the model's ability to contextualize accident data within the inherent operational and hierarchical characteristics of the roadway.

4.3. Model Validation

Given that direct stakeholder feedback was beyond the scope of this study, the model's validity and the contribution of its multi-factor design were assessed through a comparative analysis against a classical hotspot identification method. To validate the added value of the composite RSR model, its results were compared against a hotspot map generated using a traditional method: Kernel Density Estimation (KDE). Using the Kernel Density tool in ArcGIS Pro, a continuous surface map was created from the same cleansed accident point data used in the RSR model. A search radius of 600 m was applied to identify localized crash clusters, with the output representing the density of crashes per square kilometer. This produced a classical hotspot map based solely on historical crash frequency.

The comparison, illustrated in **Figure 4**, reveals two key findings. First, there is a significant and logical overlap between the two methods. Major arterial corridors identified as 'High' or 'Extremely High Risk' by the RSR model consistently align with the high-density zones (hotspots) identified by the KDE analysis. This confirms that the RSR model effectively captures locations with a proven history of frequent crashes.

Second, and more importantly, the comparison highlights the unique contribution of the RSR model's proactive components. The model successfully identifies numerous road segments as high-risk that would be overlooked by the classical frequency-based approach, as evidenced by the insert maps in **Figure 4**. This provides a compelling example of a segment of the Eastern Freeway and Reynold Road. While the KDE analysis identifies the primary hotspot at a major intersection known for frequent but often minor collisions (Area A), the RSR model flags the entire connecting freeway segment as "Extremely High Risk". This is because, despite having a lower crash frequency, the incidents on this segment were more severe and occurred in a high-speed, high-hierarchy environment. A transport authority relying solely on classical hotspot mapping would miss the systemic risk present in Area B, demonstrating the RSR model's superior ability to provide a more holistic and proactive assessment for prioritizing safety interventions.

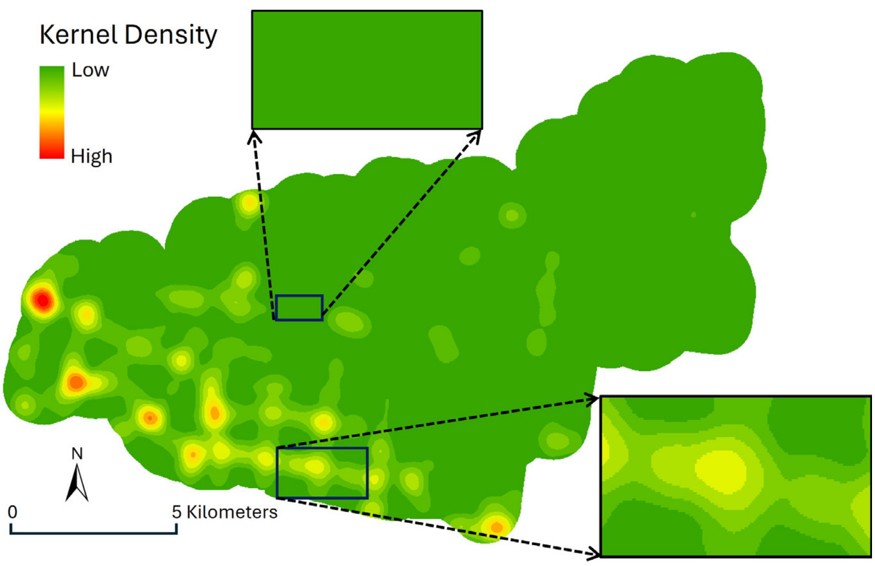

Figure 4. *Cont.*

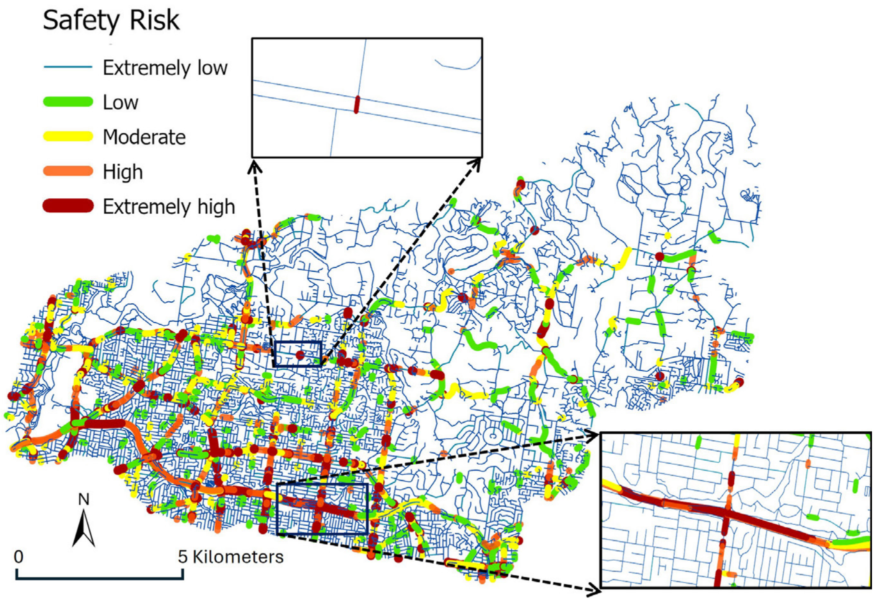

Figure 4. Comparative validation of the RSR model against classical hotspot analysis.

Note: Both the Eastern Freeway (bottom inset) and the Reynolds Rd corridor (top inset) are identified as 'Extremely High Risk' by the RSR model, while KDE classifies both of these dangerous road segments as having only low-to-moderate crash density, demonstrating the RSR model's superior ability to proactively identify systemic risk.

4.4. Summary of Findings

The application of the road safety risk model yields the following key outcomes for the Manningham study area:

1) Model Efficacy: The multi-factor geospatial model proved effective in identifying, quantifying, and spatially differentiating road safety risk across a heterogeneous urban road network. The methodology successfully transformed raw accident and road network data into an actionable and interpretable risk map.
2) Confirmation of Risk Concentration: The model confirms that the highest levels of road safety risk are concentrated on the primary arterial and collector road network. This outcome is driven not just by historical accident data but is significantly amplified by the compounding effects of higher speed environments and the road's functional classification.
3) Proactive Assessment Capability: The resulting risk map provides a clear, evidence-based visualization that moves beyond simple reactive hotspot mapping. By integrating multiple risk factors, the model offers a more proactive and nuanced assessment of road safety, providing a valuable tool for prioritizing safety interventions.

5. Discussion

5.1. Interpretation of Findings in the Context of Previous Studies

This study successfully developed and applied a multi-factor geospatial model to assess road safety risk in the City of Manningham. The results clearly indicate that elevated risk is spatially concentrated along a small fraction of the road network, specifically the major arterial and collector roads. This primary finding strongly aligns with the established literature. The working hypothesis that risk is a function of not only historical crashes but also the inherent characteristics of the roadway is strongly supported. Major arterial roads are defined by higher speed limits and greater traffic volumes, factors consistently linked to increased crash frequency and severity. Our results reconfirm the positive association between area-wide average speed and the number of fatalities and serious injuries [22]. The high-risk scores on these corridors are a product of this compounded risk: they not only have more crashes (Accident Frequency), but the crashes that do occur are more severe (Accident Severity), and this effect is amplified by the model's proactive weighting of the Speed Zone Factor and Road Class Factor. Conversely, the overwhelming proportion of the network classified as 'Low Risk' corresponds to residential streets, where lower speeds and traffic volumes create a safer, "self-explanatory" road environment that naturally mitigates risk [21].

The methodological approach of this paper also represents a practical step in the evolution of road safety analysis. Traditional hotspot identification relies on reactive clustering of historical crash data, often using techniques like Kernel Density Estimation (KDE) to visualize crash concentrations [9,15]. However, this approach can be limited by the often random and rare nature of crash events [31]. Our model addresses this limitation by integrating proactive risk indicators. It serves as a hybrid framework, using historical crash data to quantify demonstrated failure while incorporating the intrinsic risk posed by a road's design function and speed environment. This synthesis provides a more robust and holistic assessment than relying on either historical data or theoretical risk factors alone, a principle that aligns with the broader goal of developing more comprehensive and data-driven safety management tools [14,32].

5.2. Implications for Road Safety Policy and Practice

The findings of this study have significant implications for urban transport planning and safety management. The most immediate implication is the ability to support targeted and efficient resource allocation. The statistical result that approximately 90% of the road network length is classified as 'Low Risk', while the 'High' and 'Extremely High' risk categories constitute less than 7% of the network, provides a powerful message for policymakers. It demonstrates that road safety investments do not need to be spread thinly across an entire city but can be strategically focused on a small number of high-priority corridors where the greatest risk reductions can be achieved. This data-driven approach to prioritization is a cornerstone of modern safety management [33].

Furthermore, this study promotes a shift from a "spot-based" to a "corridor-level" approach for safety interventions. Instead of treating individual intersections as isolated black spots, the model visualizes risk as a continuous, linear phenomenon along major routes. This encourages planners to think systemically about corridor-wide treatments, such as speed management, access control, or infrastructure improvements for vulnerable road users, which are often more effective than localized fixes. The transparency and replicability of the methodology itself are key implications. By utilizing publicly available data and standard GIS tools, the model provides a practical and accessible framework for small to medium-sized municipalities that may lack the resources for more complex crash prediction modeling, thereby democratizing data-driven road safety analysis.

From the perspective of a local transport authority, the primary utility of this model lies in its function as a strategic screening and prioritization tool. While local authorities were not directly consulted in the development of this proof-of-concept model, its design—using open-source government data and standard GIS workflows—ensures it is both replicable and accessible. The resulting risk map serves as a powerful, evidence-based starting point for policy discussions and resource allocation. For instance, instead of relying solely on anecdotal evidence or politically driven requests, planners can use the risk map to objectively demonstrate why safety investments should be concentrated on specific arterial corridors. The analysis provides a clear justification for shifting from a purely reactive "black spot" treatment program to a more proactive, systemic corridor-level safety management strategy. The map does not replace the need for detailed on-site safety audits, but rather empowers authorities to conduct these expensive and time-consuming audits in the most critical locations, ensuring that limited public funds are directed where they can achieve the greatest impact on reducing fatal and serious injuries.

5.3. Limitations and Future Research Directions

While the proposed model provides valuable insights, it is important to acknowledge its limitations, which in turn highlight directions for future research. The model's primary limitation is its reliance on historical crash frequency as a proxy for traffic exposure, as it does not explicitly incorporate direct traffic volume data like Annual Average Daily Traffic (AADT). In road safety analysis, it is a well-established principle that accident counts must be corrected by a measure of exposure to determine a true risk rate. For instance, Aultman-Hall and Kaltenecker [23] normalized their bicycle accident data by "bicycle kilometer" to derive facility-specific event rates, demonstrating that simply counting incidents can be misleading without accounting for how much travel occurs.

More advanced crash prediction models frequently use AADT or other measures of Total Traffic Activity (TTA) as a fundamental explanatory variable, as it provides a direct measure of the volume of vehicles exposed to potential risk on a given segment [22]. Therefore, future iterations of this model would be significantly enhanced by integrating AADT. This would allow for a more robust normalization of the accident data, enabling the calculation of true crash rates (e.g., crashes per million vehicle-kilometers traveled) rather than relying on crash frequency alone. As suggested by Rodrigues et al. [32], such normalization provides a more accurate reflection of risk, particularly when comparing high-volume arterial roads with low-volume local streets, and is a critical step in moving from a general risk index to a more refined safety performance function.

Additionally, the model's four factors represent a high-level assessment of the road environment. Future research could expand this framework to include more granular geometric and contextual variables. For instance, incorporating intersection density, the presence and quality of pedestrian and cycling facilities, and local land use data would provide a richer understanding of risk, particularly for vulnerable road users [6,24]. The general issue of crash under-reporting, especially for minor incidents, also remains a persistent challenge in all safety research and is an underlying limitation of any model based on official police records [26].

Methodologically, future studies could apply more advanced spatial statistical models to this type of framework. For example, the use of Geographically Weighted Regression (GWR) could reveal spatial non-stationarity, identifying specific areas within the municipality where the relationship between risk factors (like speed) and crash outcomes is particularly strong [19,25]. The analysis could be disaggregated by crash type or road user, allowing for the development of separate risk models for pedestrians, cyclists, and different collision types (e.g., run-off-road, head-on), providing more targeted insights for specific safety countermeasures [23]. Additionally, the current model does not disaggregate risk by time of day or other temporal factors, and future research could develop dynamic risk indices by incorporating temporal data to provide more targeted operational safety strategies. Furthermore, the model operates at a network-level and does not incorporate granular data on specific safety infrastructure (e.g., traffic signals, presence of cycling paths). Acquiring and integrating such detailed, network-wide data was beyond the scope of this study, but would be a valuable enhancement for future, more refined models. Finally, the model's lack of focus on Vulnerable Road Users (VRUs) means that the methodology is less effective for identifying risks on local streets where VRU activity is higher. It is proposed that future work should develop separate or modified indices specifically for VRU risk, incorporating factors like pedestrian crossing facilities, bicycle lane quality, and land use data. Future work should also explore variable weighting schemes based on local policy priorities or expert opinion.

6. Conclusions

Traditional road safety management often relies on the reactive identification of historical crash hotspots, an approach limited by the random nature of accident data and a failure to account for the inherent risk characteristics of the road environment. To address these shortcomings, this study developed and applied a multi-factor geospatial model that provides a more proactive and holistic assessment of road safety risk. The model's framework, which combines reactive measures (historical crash frequency and severity) with proactive indicators of roadway function and speed environment, was successfully implemented in the City of Manningham using publicly available data and standard GIS tools.

The primary conclusion drawn from the application of this model is that road safety risk is not uniformly distributed but is instead highly concentrated on a small, identifiable portion of the urban road network. The results demonstrated that while nearly 90% of the road network length was classified as 'Low Risk,' the 'High' and 'Extremely High' risk categories were confined to less than 7% of the total road length, predominantly along major arterial and collector corridors. This finding confirms the working hypothesis that a holistic risk assessment, which looks beyond mere crash counts to include factors like speed and road hierarchy, can effectively pinpoint corridors of greatest concern [31].

The importance of this research lies in its direct applicability to urban transportation planning and evidence-based policy-making. By clearly identifying the small fraction of the road network where risk is most concentrated, the model provides a robust framework for the targeted and efficient allocation of limited safety resources. This enables a strategic shift from treating isolated "black spots" to implementing corridor-level interventions where they are most needed. Furthermore, the transparency and replicability of the methodology make it a valuable and accessible tool for municipalities that may lack the extensive data or technical resources required for more complex statistical crash prediction models. In conclusion, this paper contributes a practical and conceptually sound framework for advancing from reactive hotspot identification to proactive risk management. By successfully integrating historical accident data with intrinsic roadway characteristics, the model provides an actionable tool for planners and engineers to make more informed decisions, ultimately supporting the development of safer transportation systems.

Funding

This work received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments

I thank Lisa Wang from ESRI Australia for her help in combining the raw traffic crash data in the early stage of the research. I acknowledge the Faculty of Architecture Building and Planning at the University of Melbourne for its administrative and technical support.

Conflicts of Interest

The author declares no conflict of interest.

AI Use Statement

During the preparation of this manuscript, the author used Google AI Studio to assist in refining the phrasing of methodological descriptions and clarifying the presentation of data analysis results. The author has reviewed and edited the output and takes full responsibility for the content of this publication.

References

[1] Oestreich, L., Pereira, B.M., Ruiz-Padillo, A., 2022. Identification of Contributing Factors and Hotspots of Pedestrian-Vehicle Collisions in Urban Areas: An Alternative Hybrid Method for Small Cities. Traffic Injury Prevention. 23(7), 416–421. DOI: https://doi.org/10.1080/15389588.2022.2104838

[2] Jayasinghe, B.C., Withanage, N.C., Mishra, P.K., 2025. Evaluating Geographical Variations of Road Traffic Accidents in Matara, Sri Lanka: A Geospatial Perspective to Policy Decisions. Revue Internationale de Géomatique. 34, 707–729. DOI: https://doi.org/10.32604/rig.2025.067395

[3] Ibe, C.C., Nwokedi, T.C., 2025. Geospatial Probability Mapping of Road Incidents for Prioritizing Road Safety Awareness and De-Escalating Associated Death and Injury Burdens in Nigerian. Transportation Research Procedia. 89, 227–242. DOI: https://doi.org/10.1016/j.trpro.2025.05.058

[4] Litman, T., 2025. Evaluating Transportation Equity: Guidance for Incorporating Distributional Impacts in Transport Planning. Victoria Transport Policy Institute: Victoria, BC, Canada.

[5] Petch, R.O., Henson, R.R., 2000. Child Road Safety in the Urban Environment. Journal of Transport Geography. 8(3), 197–211. DOI: https://doi.org/10.1016/

S0966-6923(00)00006-5

[6] Hu, L., Wu, X., Huang, J., et al., 2020. Investigation of Clusters and Injuries in Pedestrian Crashes Using GIS in Changsha, China. Safety Science. 127, 104710. DOI: https://doi.org/10.1016/j.ssci.2020.104710

[7] Pucher, J., Dijkstra, L., 2003. Promoting Safe Walking and Cycling to Improve Public Health: Lessons from the Netherlands and Germany. American Journal of Public Health. 93(9), 1509–1516. DOI: https://doi.org/10.2105/ajph.93.9.1509

[8] Tiwari, G., 2023. Systems-Thinking-Based Road Safety Research: The Way Forward. International Journal of Injury Control and Safety Promotion. 30(4), 471–472. DOI: https://doi.org/10.1080/17457300.2023.2282001

[9] Khatun, M.S., Hossain, M.A., Kabir, M.A., et al., 2024. Identification and Analysis of Accident Black Spots Using Geographic Information System (GIS): A Study on Kushtia-Jhenaidah National Highway (N704), Bangladesh. Heliyon. 10(3), e25952. DOI: https://doi.org/10.1016/j.heliyon.2024.e25952

[10] Alkaabi, K., 2023. Identification of Hotspot Areas for Traffic Accidents and Analyzing Drivers' Behaviors and Road Accidents. Transportation Research Interdisciplinary Perspectives. 22, 100929. DOI: https://doi.org/10.1016/j.trip.2023.100929

[11] Hu, Y., Wang, F., Guin, C., et al., 2018. A Spatio-Temporal Kernel Density Estimation Framework for Predictive Crime Hotspot Mapping and Evaluation. Applied Geography. 99, 89–97. DOI: https://doi.org/10.1016/j.apgeog.2018.08.001

[12] Forte, J.P., Brilha, J., Pereira, D.I., et al., 2018. Kernel Density Applied to the Quantitative Assessment of Geodiversity. Geoheritage. 10(2), 205–217. DOI: https://doi.org/10.1007/s12371-018-0282-3

[13] Savolainen, P.T., Mannering, F.L., Lord, D., et al., 2011. The Statistical Analysis of Highway Crash-Injury Severities: A Review and Assessment of Methodological Alternatives. Accident Analysis & Prevention. 43(5), 1666–1676. DOI: https://doi.org/10.1016/j.aap.2011.03.025

[14] Pinna, M., Serra, P., Porcu, C., et al., 2025. An Integrated Approach for Urban Road Safety Analysis. Transportation Research Procedia. 90, 702–709. DOI: https://doi.org/10.1016/j.trpro.2025.06.089

[15] Hashimoto, S., Yoshiki, S., Saeki, R., et al., 2016. Development and Application of Traffic Accident Density Estimation Models Using Kernel Density Estimation. Journal of Traffic and Transportation Engineering (English Edition). 3(3), 262–270. DOI: https://doi.org/10.1016/j.jtte.2016.01.005

[16] Aghajani, M.A., Dezfoulian, R.S., Arjroody, A.R., et al., 2017. Applying GIS to Identify the Spatial and Temporal Patterns of Road Accidents Using Spatial Statistics (Case Study: Ilam Province, Iran). Transportation Research Procedia. 25, 2126–2138. DOI: https://doi.org/10.1016/j.trpro.2017.05.409

[17] Macedo, M.R.O.B.C., Maia, M.L.A., Robbani, E.R.K., et al., 2022. Traffic Accident Prediction Model for Rural Highways in Pernambuco. Case Studies on Transport Policy. 10(1), 278–286. DOI: https://doi.org/10.1016/j.cstp.2021.12.009

[18] Zhang, H., Zhang, M., Zhang, C., et al., 2021. Formulating a GIS-Based Geometric Design Quality Assessment Model for Mountain Highways. Accident Analysis & Prevention. 157, 106172. DOI: https://doi.org/10.1016/j.aap.2021.106172

[19] Gu, Y., Cho, E., Oh, C., et al., 2025. Riding Safety Evaluation of Food Delivery Motor Scooters Based on Associating Sensor-Based Riding Behavior and Road Traffic Characteristics. Accident Analysis & Prevention. 211, 107871. DOI: https://doi.org/10.1016/j.aap.2024.107871

[20] Pokorny, P., Jensen, J.K., Gross, F., et al., 2020. Safety Effects of Traffic Lane and Shoulder Widths on Two-Lane Undivided Rural Roads: A Matched Case-Control Study from Norway. Accident Analysis & Prevention. 144, 105614. DOI: https://doi.org/10.1016/j.aap.2020.105614

[21] Babić, D., Brijs, T., 2021. Low-Cost Road Marking Measures for Increasing Safety in Horizontal Curves: A Driving Simulator Study. Accident Analysis & Prevention. 153, 106013. DOI: https://doi.org/10.1016/j.aap.2021.106013

[22] Wang, C., Quddus, M., Ison, S., 2009. The Effects of Area-Wide Road Speed and Curvature on Traffic Casualties in England. Journal of Transport Geography. 17(5), 385–395. DOI: https://doi.org/10.1016/j.jtrangeo.2008.06.003

[23] Aultman-Hall, L., Kaltenecker, M.G., 1999. Toronto Bicycle Commuter Safety Rates. Accident Analysis & Prevention. 31(6), 675–686. DOI: https://doi.org/10.1016/S0001-4575(99)00028-7

[24] Chen, C., Anderson, J.C., Wang, H., et al., 2017. How Bicycle Level of Traffic Stress Correlates with Reported Cyclist Accident Injury Severities: A Geospatial and Mixed Logit Analysis. Accident Analysis & Prevention. 108, 234–244. DOI: https://doi.org/10.1016/j.aap.2017.09.001

[25] Erdogan, S., 2009. Explorative Spatial Analysis of Traffic Accident Statistics and Road Mortality Among the Provinces of Turkey. Journal of Safety Research.

40(5), 341–351. DOI: https://doi.org/10.1016/j.jsr.2009.07.006

[26] Ahmed, A., Sadullah, A.F.M., Yahya, A.S., 2019. Errors in Accident Data, Its Types, Causes and Methods of Rectification: Analysis of the Literature. Accident Analysis & Prevention. 130, 3–21. DOI: https://doi.org/10.1016/j.aap.2017.07.018

[27] Jiang, F., Yuen, K.K.R., Lee, E.W.M., 2020. Analysis of Motorcycle Accidents Using Association Rule Mining-Based Framework with Parameter Optimization and GIS Technology. Journal of Safety Research. 75, 292–309. DOI: https://doi.org/10.1016/j.jsr.2020.09.004

[28] Wen, X., Xie, Y., Wu, L., et al., 2021. Quantifying and comparing the effects of key risk factors on various types of roadway segment crashes with LightGBM and SHAP. Accident Analysis & Prevention. 159, 106261. DOI: https://doi.org/10.1016/j.aap.2021.106261

[29] Al-Omari, M.M.A., Abdel-Aty, M., Cai, Q., 2021. Crash analysis and development of safety performance functions for Florida roads in the framework of the context classification system. Journal of Safety Research. 79, 1–13. DOI: https://doi.org/10.1016/j.jsr.2021.08.004

[30] Papadimitriou, E., Filtness, A., Theofilatos, A., et al., 2019. Review and ranking of crash risk factors related to the road infrastructure. Accident Analysis & Prevention. 125, 85–97. DOI: https://doi.org/10.1016/j.aap.2019.01.002

[31] Lord, D., Mannering, F., 2010. The Statistical Analysis of Crash-Frequency Data: A Review and Assessment of Methodological Alternatives. Transportation Research Part A: Policy and Practice. 44(5), 291–305. DOI: https://doi.org/10.1016/j.tra.2010.02.001

[32] Rodrigues, D.S., Ribeiro, P.J.G., da Silva Nogueira, I.C., 2015. Safety Classification Using GIS in Decision-Making Process to Define Priority Road Interventions. Journal of Transport Geography. 43, 101–110. DOI: https://doi.org/10.1016/j.jtrangeo.2015.01.007

[33] Matović, B., Damjanović, M., Ilić, V., et al., 2025. Identification and Ranking of High-Risk Road Segments and Intersections in the Municipality of Bačka Palanka. Transportation Research Procedia. 91, 806–813. DOI: https://doi.org/10.1016/j.trpro.2025.10.103