EPJ Web of Conferences **379**, 04001 (2026) https://doi.org/10.1051/epjconf/202637904001
ICS2MT-2026

Determining Risk Factors for Vulnerable Road User Crashes: Integrating GIS and Binary Logistic Regression in a Medium-sized City of a Developing Nation

Srinivasa Rao Gandupalli[1,2*], *Purnanandam* Kokkeragadda[2], *Mukund R.* Dangeti[1] and *Praveen* Samarthi[3]

[1]Department of Civil Engineering, GITAM School of Core Engineering, GITAM Deemed to be University, Visakhapatnam, Andhra Pradesh, India
[2]Department of Civil Engineering, University College of Engineering Kakinada (UCEK), JNTUK, Kakinada, Andhra Pradesh, India
[3]Department of Civil Engineering, CVR College of Engineering, Hyderabad, Telangana, India

Abstract. The population increase in urban areas leads to heightened vehicle usage, resulting in more interactions among vehicles, pedestrians, and bicyclists, thereby raising substantial road safety issues. Infrastructure must be appropriately constructed for both motorized and non-motorized vehicles to mitigate safety problems for all road users. This study identifies high-crash-prone road segments and critical risk factors for vulnerable road user (VRU) crashes, encompassing pedestrians, bicyclists, and motorcyclists. This is accomplished by creating heatmaps from black-spot research through Geographic Information Systems (GIS) and by determining risk factors with a binary logistic regression (BLR) model. Crash data from Visakhapatnam, India, for the years 2014-2016 and 2019-2021 reveals that more than 50% of fatal incidents involved VRUs, and 75.6% of identified blackspot road segments are situated along a 62-km stretch of National Highway traversing the city. Roadways and land-use factors were collected during road-safety audits along the designated segment. The BLR Model identifies risk factors such as segment length, crash time, season, land use, driver sight distance, and vehicle type. The severity of crashes increases by 17.6% per unit increase in segment length, whereas insufficient visibility increases it by 43.1%. This integrated approach directs targeted interventions to improve the safety of VRUs.

1 Introduction and Background

Road traffic fatalities are a major public health issue in urban areas of developing countries. In 2021, the World Health Organisation reported [15] 1.19 million fatalities due to road traffic accidents. VRUs, such as motorcyclists (21%), pedestrians (23%), cyclists (6%), and e-scooter users (3%), accounted for 53% of the deaths. Approximately 80% of global roads do

* Corresponding author: sgandupa@gitam.edu

not meet VRU requirements. Approximately 25% of countries have adopted policies promoting walking, cycling, or public transportation.

India, with a population exceeding 1.3 billion, holds the second-largest road network globally. Highways account for less than 5% of road length yet are responsible for over 56% of road crashes and 61% of traffic fatalities [15]. In 2022 [5], there were more than 461,000 crashes, leading to 168,491 fatalities and over 443,000 injuries. Over speeding, negligent driving, and poor enforcement remain leading causes. Two-wheelers, a significant segment of India's vehicle fleet, accounted for approximately 40% of fatalities.

Deaths resulting from VRUs are both foreseeable and avoidable [15]. In 2022, 67% of road fatalities in India involved VRUs, with motorcycles accounting for 44.5% and pedestrians comprising 19.5% of the total. Urban areas account for 80% of road fatalities, with VRUs predominantly affected in areas lacking adequate pedestrian and cycling infrastructure, a situation attributed to insufficient planning. City-level studies substantiate this concern. In Delhi, more than 85% of traffic deaths are VRUs [4], whereas in Kolkata, pedestrians make up almost 50% of deaths [9]. A study [7] found that VRU fatalities varied between 84% and 93% across six medium-sized Indian cities, including Visakhapatnam.

In high-income nations, mortality rates for VRUs have markedly decreased. In 2016, VRUs accounted for 31% of traffic fatalities in the United States and 35% in Australia. In low- and middle-income countries (LMICs), the lack of systematic frameworks for VRU safety increases the risks associated with VRUs. Numerous studies have identified the determinants of crashes involving VRUs. A Study [13] identified that insufficient signalisation and improper bus stop placement in Thiruvananthapuram contributed to pedestrian crashes. Pedestrian behaviour and risk perception at signalised intersections are investigated [8]. The risk factors associated with motorcycle fatalities are investigated on rural roads [10]. BLR was employed to identify accident cause and location as key factors influencing severity in Riyadh [2]. In a study, a connection was established between fatality risk, vehicle type, and the absence of enforcement at intersections [14].

Logistic regression was utilised to determine high speed, car type, and crash location as significant predictors in Garmian [1]. A study [12] conducted spatial analysis and surveys to evaluate pedestrian risk perception in Delhi, identifying built environment factors like speed, sidewalk conditions, and lane width as significant influences. GIS has been extensively utilised by researchers to identify accident hotspots [3,11].

Despite heightened research efforts, megacities continue to dominate, while tier-2 cities receive comparatively less attention. This research aims to fill a gap by assessing the severity of VRU crashes in a major highway corridor in Visakhapatnam, India. This study identifies critical risk factors by employing GIS-based hotspot detection and BLR modelling. It provides insights for urban planners, traffic authorities, and policymakers to significantly improve VRU safety in developing urban areas.

This study fills a significant gap in current research by examining Visakhapatnam, a tier-2 city with rapidly changing traffic dynamics that has received limited attention in road safety literature. By integrating spatial filtering via GIS heatmaps and BLR, this research establishes a novel framework for detecting risk patterns linked to VRUs in medium-sized urban areas of developing nations.

2 Methodology

This study utilises methodological approaches to identify and quantify the primary sources of risk variables impacting crashes involving VRUs across different locations within the road network, aiming to achieve the following objectives: a) Creation of heatmaps using GIS to identify blackspots based on historical crash data; b) Formulation of a crash severity model through BLR and identification of associated risk factors; c) Offering effective solutions to reduce the identified risks faced by VRUs.

2.1 Study area

In India, cities with populations ranging from 0.5 to 5 million are considered tier-II cities [6]. In this study, a tier-II city, Visakhapatnam, was selected as the study area. Figure 1 illustrates the location map of the city.

Fig. 1. Study area location map

2.2 Crash data

Crash data were collected from the Visakhapatnam Traffic Police for the years 2014 to 2021. Each crash record includes the types of vehicles involved, crash severity, date, time, geographic locations, and the number of deaths and injuries. Pre-processing of data was carried out to eliminate incomplete and duplicate records. GIS was used to analyse these records, focusing on crash dispersion and the creation of density-based heat maps.

2.3 Heat map generation using GIS

Crash hotspots were identified using the Kernel Density Estimation (KDE) technique. This method employs a Gaussian kernel function along with a bandwidth parameter to delineate the spatial influence of crash locations. The appearance of the density surface is influenced by bandwidth. Larger bandwidth values yield smoother spatial patterns, whereas smaller values exhibit more localised variations in crash concentration. Figure 2 presents KDE-based crash density maps for the periods 2014–2016 and 2019–2021.

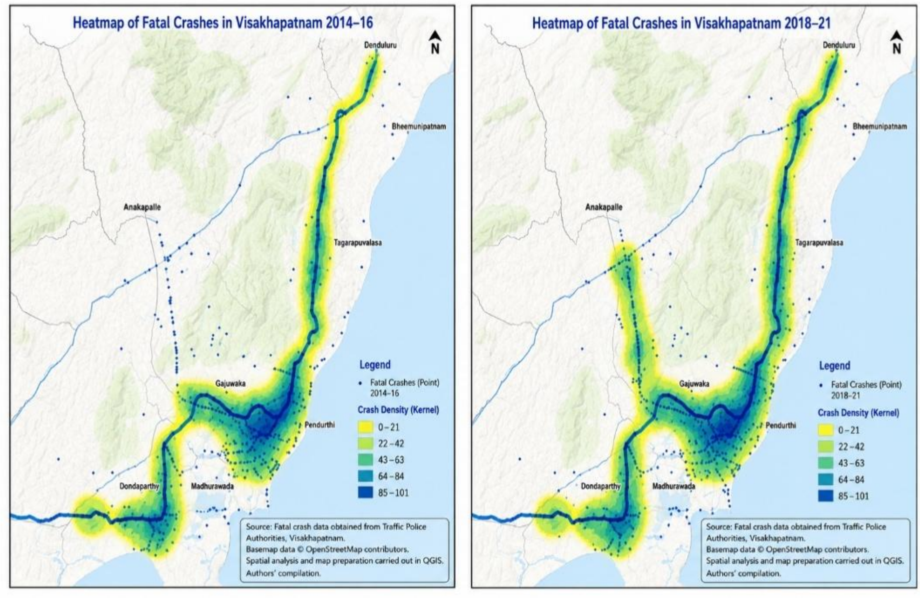

Fig. 2. Kernel density heatmaps of crash locations in Visakhapatnam for 2014-2016 and 2019-2021

2.4 Blackspot identification

As per MoRTH, a "black spot" is a designated segment of road measuring 500 meters in length. Over the past three years, there have been 5 traffic accidents resulting in fatalities or serious injuries. Alternatively, it can be defined as a geographic region where fatalities have reached or exceeded ten within three years. The heat maps indicated that crash density on the national highway segment traversing the city was higher during 2014-2016 and 2019-2021 than on other urban roads in the town. Figure 3 demonstrates that 75.6% to 67.2% of the road segments classified as black spots were concentrated along the city's 62-kilometre section of the national highway.

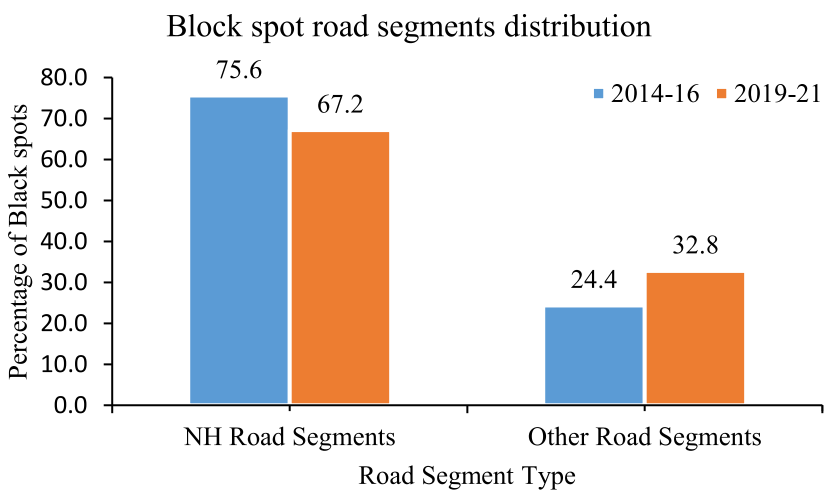

Fig. 3. Identified blackspot segments on urban roads and national highway (62km) within Visakhapatnam

In comparison, 32.8% to 24.4% of black spots were identified on the remaining urban road network of the city. The analysis indicates that the 62-kilometre national highway segment requires additional investigation to determine the risk factors affecting crash severity. Figure 4 presents the distribution of fatal VRU crashes along the study corridor.

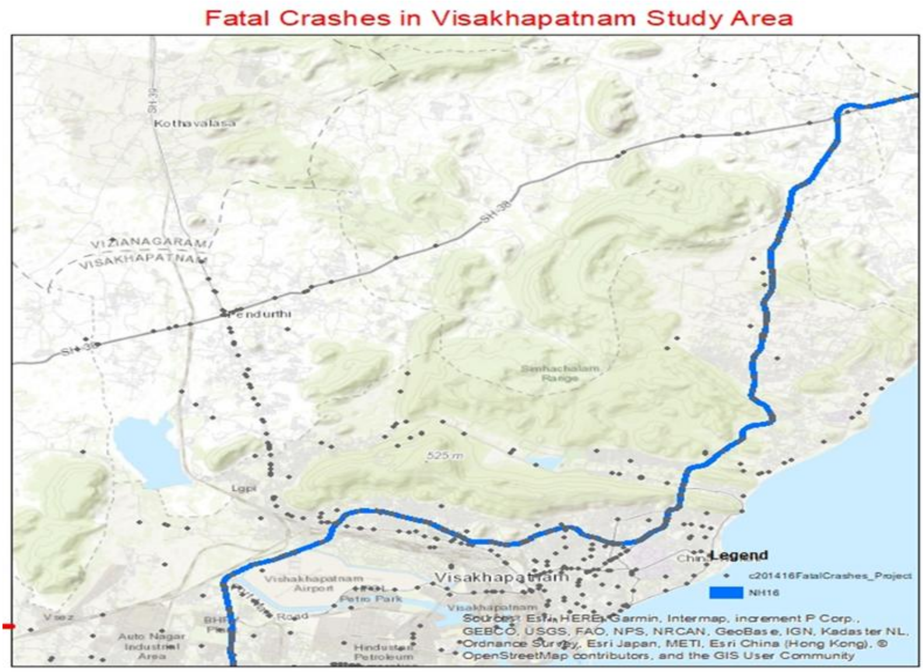

Fig. 4. Distribution of fatal VRU crashes along the study corridor (NH Segment)

Following the establishment of the study stretch, a road safety audit was performed to gather data on the road infrastructure. Road safety audits provide insights into road geometry, encompassing the number of pedestrian crossings, curves, median openings, intersections, side access roads, segment length, presence of service roads, earthen shoulders, proper signage, adequate sight clearance for drivers, road markings, and type of land use.

3 Model Development

BLR is employed to construct the crash severity prediction models. This study classifies accident severity into two categories for BLR models: fatal and non-fatal crashes. BLR models are suitable for scenarios where the output variable is binary (e.g., VRU fatality occurrence: Yes=1; No=0 at a specific location), with predictor variables being either categorical or continuous. The coefficients for independent variables are estimated using either the odds as the response metric. The probability of a fatal VRU crash at a specific segment is determined using the function in equation (1) and the odds in equation (2).

$$Y_{in} = Logit_i = ln\left(\frac{Prob_i}{1-Prob_i}\right) = a_i + b_{i1}X_1 + b_{i2}X_2 + \cdots + b_{in}X_n \qquad (1)$$

$$Odds_i = \frac{Prob_i}{1-Prob_i} = e^{a_i + b_{i1}X_1 + b_{i2}X_2 + \cdots + b_{in}X_n} = e^{Y_i} \qquad (2)$$

Y_{in} is the utility function for the outcome i of observation n; ai is a constant, n is the number of predictor variables, and b_{i1} to b_{in} are the coefficients. The utility of outcome 'i must be converted into a likelihood to establish the probability of a specific outcome occurring at a particular location.

From 2014 to 2021, 2,871 crash records involving VRUs were recorded along a 62 km segment of a national highway designated as the study corridor. 2,294 samples and their corresponding location details were employed for subsequent analysis. Of the records, 80% of 1837 were utilised for model construction, whereas 20% were allocated for validation. This study utilised correlation analysis to evaluate the relationship between response and predictor variables. The study identified predictive factors associated with fatal and injury crashes, with correlation coefficients indicating a moderate to strong association between these variables and accident incidence at 90% to 95% confidence intervals.

Table 1 displays the correlation coefficients for each predictor factor with the response variables. The test results demonstrate a strong correlation between segment length, crash season, vehicle type, crash time, number of pedestrian crossings, and sight clearance with crash type at a 95% confidence level. The curves and appropriate road markings moderately correlate with crash type at 90% confidence intervals. Additional predictor variables, such as crash day, presence of service road, land use, median openings, intersections, side access roads, presence of earthen shoulder, and proper signage, were excluded from model development because of their weak correlation with crash type.

Table 1. Correlation Coefficient Between Predictor Variables and Crash Severity Type

Predictor Variables	**BLR Coefficients**
Segment length	.077**
Crash Time	.106**
Crash Season	-.062**
Day	0.018
Accused vehicle type	.159**
Presence of Service Road	0.022
Type of Land Use	-0.036
Number of Median openings	0.018
Curves	.053*
Intersections	-0.032
Pedestrian Crossings	.070**
Presence of Earthen shoulder	-0.002
Adequate Sight clearance for the drivers	-.070**
Proper Signage	-0.021
Adequate Road Markings	-.051*
Side Access Roads (number)	0.017
*significant at 90% C.I; ** significant at 95% C.I;	

The BLR model was created to analyse the characteristics of crashes, roadways, and land use that influence the likelihood of fatal crashes involving VRUs at specific locations. Table

2 summarises the BLR model and the coefficients for the variables included. The use of coefficients and model summaries enhances the interpretation of findings. The model's findings demonstrate a positive relation between crash season, segment length, open land-use type, and insufficient sight clearance for drivers, and crash severity. Conversely, crash times outside midnight and early morning, and vehicle types other than heavy vehicles, negatively correlate with crash severity. All risk factors were statistically significant at 95% C.I. Odds ratios (Exp(B)) were used to evaluate the influence of specific risk factors on accident severity. The BLR model achieved a classification accuracy of 73.6% in predicting fatal and non-fatal VRU crashes.

Table 2. BLR Model Summary

<table>
<tr><th>Factor</th><th>Predictor Variables</th><th>Coefficient (B)</th><th>Sig.</th><th>Exp(B)</th></tr>
<tr><td></td><td>Constant</td><td>-0.34</td><td>0.11</td><td>0.712</td></tr>
<tr><td rowspan="2">Roadway factor</td><td>Segment Length</td><td>0.16</td><td>0.04</td><td>1.176</td></tr>
<tr><td>Sight clearance for the drivers (Not Available)</td><td>0.36</td><td>0.01</td><td>1.431</td></tr>
<tr><td rowspan="11">Crash factor</td><td>Crash Time (Morning 6 AM-12 PM)</td><td>-1.18</td><td>0</td><td>0.309</td></tr>
<tr><td>Crash Time (Afternoon 12 PM-6 PM)</td><td>-0.8</td><td>0</td><td>0.448</td></tr>
<tr><td>Crash Time (Evening/Night 6 PM to 12 AM)</td><td>-0.73</td><td>0</td><td>0.483</td></tr>
<tr><td>Crash Time (Midnight 12 AM to 6 AM)</td><td colspan="3">Reference category</td></tr>
<tr><td>Crash Season (Summer)</td><td>0.32</td><td>0.02</td><td>1.38</td></tr>
<tr><td>Crash Season (Rainy)</td><td>0.29</td><td>0.03</td><td>1.334</td></tr>
<tr><td>Crash Season (Winter)</td><td colspan="3">Reference category</td></tr>
<tr><td>Accused vehicle (2-Wheeler)</td><td>-0.77</td><td>0</td><td>0.464</td></tr>
<tr><td>Accused vehicle (3-Wheeler)</td><td>-1.23</td><td>0</td><td>0.293</td></tr>
<tr><td>Accused Vehicle (Car)</td><td>-0.8</td><td>0</td><td>0.449</td></tr>
<tr><td>Accused Vehicle (Heavy Vehicle)</td><td colspan="3">Reference category</td></tr>
<tr><td>Land use factor</td><td>Land use (Open Areas/ No Land use)</td><td>0.24</td><td>0.04</td><td>1.275</td></tr>
<tr><td colspan="5">Model Summary</td></tr>
<tr><td colspan="2">AIC</td><td colspan="3">1658.8</td></tr>
<tr><td colspan="2">BIC</td><td colspan="3">1725</td></tr>
<tr><td colspan="2">Pearson Chi-Square</td><td colspan="3">1234.1</td></tr>
<tr><td colspan="2">loglikelihood</td><td colspan="3">-817.4</td></tr>
<tr><td colspan="2">Accuracy</td><td colspan="3">73.6%</td></tr>
</table>

4 Conclusions

The study identified high-risk crash segments through heatmaps and examined risk factors associated with VRU crashes in a tier-2 Indian city. Influence of crash characteristics, vehicle type, road infrastructure, and temporal variables on the severity of VRU crashes analysed using the BLR Model. Key findings are:

- A unit increase in segment length has 17.6% higher risk of fatal collisions involving VRUs.
- Inadequate driver visibility led to a 43.1% increase in crash severity.
- The presence of open land-use resulted in a 27.5% increase in fatal crashes compared with mixed or commercial land-use.
- Early morning and late-night collisions tend to be more severe.
- Crashes during the summer and monsoon seasons were more severe than those in winter.
- Heavy vehicles, such as trucks, are involved in more severe collisions than other vehicles, including two-wheelers, three-wheelers, and cars.

The results have substantial implications for the administration of traffic safety and urban planning in medium-sized cities in developing countries.

5 Recommendations

- Regulation of speed over extended segments requires Speed limits, traffic calming measures.
- Improving motorist visibility by grass removal, signposting, and convex mirrors can reduce crash severity.
- Improve road safety habits among VRUs and motorists through community-level initiatives, institutional programs, and periodic training sessions.
- Effective implementation of evidence-based safety measures requires collaboration among stakeholders, including law enforcement, city planners, transportation departments, and educational institutions.

Declaration

- The authors sincerely thank the Visakhapatnam Traffic Police Authorities for the data component.
- Funding is not applicable.
- The authors have no conflict of interest

References

1. L. A. Ahmed, Using Logistic Regression in Determining the Effective Variables in Traffic Accidents. Applied Mathematical Sciences. **11**(42), 2047–2058 (2017).
2. A. S. Al-Ghamdi, Using Logistic Regression to Estimate the Influence of Accident Factors on Accident Severity. Accident Analysis and Prevention. **34**, 729–74 (2002).
3. T. K. Anderson, Kernel Density Estimation and K-Means Clustering to Profile Road Accident Hotspots. Accident Analysis & Prevention. **41**(3), 359–364 (2009).
4. R. Goel, P. Jain, G. Tiwari, Correlates of Fatality Risk of Vulnerable Road Users in Delhi. Accident Analysis and Prevention. **111**, 86–93 (2018).
5. Ministry of Road Transport & Highways, Road Accidents in India 2022. Transport Research Wing, New Delhi (2023).

6. S.H. Mir, D. Kumar, Traffic Congestion in Tier-I and Tier-II Cities by Executing Intelligent Traffic System. International Journal of Progressive Research in Science and Engineering. **1**(3), 12–17 (2020).
7. D. Mohan, G. Tiwari, S. Mukherjee, Urban Traffic Safety Assessment: A Case Study of Six Indian Cities. IATSS Research. **39**(2), 95–101(2016).
8. D. Mukherjee, S. Mitra, Impact of Road Infrastructure Land Use and Traffic Operational Characteristics on Pedestrian Fatality Risk: A Case Study of Kolkata, India. Transportation in Developing Economies. **5**:6 (2019).
9. D. Mukherjee, S. Mitra, Development of a Systematic Methodology to Enhance the Safety of Vulnerable Road Users in Developing Countries. Transportation in Developing Economies. **8**:28, 1–17(2022).
10. H.M. Naqvi, G. Tiwari, Factors Contributing to Motorcycle Fatal Crashes on National Highways in India. Transportation Research Procedia. **25**, 2084–2097(2017).
11. V. Prasannakumar, H. Vijith, R. Charutha, N. Geetha, Spatio-Temporal Clustering of Road Accidents: GIS-Based Analysis and Assessment. International Journal of Geographical Information Science. **25**(8), 1243–1266 (2011).
12. S. Rankavat, G. Tiwari, Pedestrians' Risk Perception of Traffic Crash and Built Environment Features – Delhi, India. Safety Science. **87**, 1–7 (2016).
13. A. Santhosh, E. Sam, B. K. Bindhu, Pedestrian Accident Prediction Modelling—A Case Study in Thiruvananthapuram City. Transportation Research. edited by T. V. Mathew, G. J. Joshi, N. R. Velaga, S. Arkatkar, 637–645 (2020).
14. S. Sarkar, R. Tay, J. D. Hunt, Logistic Regression Model of Risk of Fatality in Vehicle–Pedestrian Crashes on National Highways in Bangladesh. Transportation Research Record. **2264**, 128–137 (2011).
15. World Health Organisation, Global Status Report on Road Safety 2023. Geneva: WHO (2023).