Jordan Journal of Civil Engineering, Volume **18**, No. **2**, **2024**
DOI: https://doi.org/10.14525/JJCE.v18i2.07

Jordan Journal of Civil Engineering

Journal homepage: https://jjce.just.edu.jo

Modeling of Risk Factors for Vulnerable Road Users' Crashes on National Highways Passing through Medium-sized Indian Cities

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

[1)] Department of Civil Engineering, University College of Engineering Kakinada (UCEK), Jawaharlal Nehru Technological University Kakinada (JNTUK), Kakinada, Andhra Pradesh, India.
[2)] Department of Civil Engineering, GITAM School of Technology, GITAM Deemed to be University, Visakhapatnam, Andhra Pradesh, India. [*] Corresponding Author. E-Mail: sgandupa@gitam.edu

ARTICLE INFO

Article History:
Received: 10/12/2023
Accepted: 5/3/2024

ABSTRACT

The safety of vulnerable road users, such as pedestrians, bicyclists and motorised two-wheelers, on national highways passing through urban areas has been a significant concern in developing countries like India. This study analyzes Visakhapatnam police crash data (2014-2017) to identify high-risk road segments to vulnerable road users. Establishing negative binomial models helps analyze risk variables for fatal and injury crashes concerning several susceptible road users on mid-block sections. Further, this study focuses on site-specific information collected from road-safety audits, speed, average daily traffic, length of road section and percentage of two-wheelers, also considered to fully help understand the likelihood of a crash. The statistical analysis of this paper identified the risk variables, like length of road segment, presence of service roads, land-use type (i.e., commercial/mixed land use), number of curves and average daily traffic associated with the incidence of fatal and injury crashes of vulnerable road users in Visakhapatnam city, India. The conclusions highlight the crucial safety measures that transportation planners and policymakers must take to create a more secure environment for road users.

Keywords: Vulnerable road users, Mid-block sections, Negative binomial models, Crash, Count-data models, Safety measures.

INTRODUCTION

The problem of road safety is a significant global concern. The phenomenon has no limitations regarding population characteristics or geographical distribution, affecting nations of all income levels. Worldwide, around 1.35 million people yearly lose their lives in traffic accidents (WHO,2018). In 2022, road accidents in India resulted in the loss of over 168,491 lives and caused injury to over 443,366 individuals. National Highways (NHs), comprising about 2.1% of the overall road network in the country, were responsible for around 32.9% of all accidents and more than 36.2% of deaths (MoRTH, 2023). This issue requires an immediate solution.

Vulnerable road users (VRUs), including motorcyclists, pedestrians and cyclists, comprise almost 50% of all road-traffic fatalities (WHO, 2018). In India, VRUs accounted for 67% of total road-traffic fatalities. Two-wheelers, with a share of 44.5%, constitute the most significant number of victims of road-accident deaths. Pedestrians accounted for 19.5% of the total fatalities in road accidents in 2022 (MoRTH, 2023). In urban India, almost 80% of the deaths were attributed to the inadequate planning of facilities for VRUs.

Multiple studies have identified VRU safety as a prominent road-safety issue in urban India. More than 85% of the people killed in traffic accidents in Delhi are VRUs (Goel et al., 2018). According to the Kolkata Police crash records, 80% of the fatally injured victims are VRUs. Pedestrians are significant, comprising almost 50% of fatal accidents (Mukherjee and Mitra, 2022). The modal proportion of road-traffic fatalities attributed to VRUs in Mumbai city is about 89% (Mohan et al., 2016). In the city of Bengaluru, VRUs consist of 94% of the individuals who have lost their lives in road accidents (Verma et al., 2021). A study by Mohan et al. (2016) presented statistics on fatal road-traffic crashes in six medium-sized cities in India: Agra, Bhopal, Vadodara, Amritsar, Ludhiana and Vishakhapatnam. The proportion of fatalities among VRUs, with the overall number of road fatalities, varies from 84% to 93% among all six cities. In Patiala, nearly 87% of road-traffic fatalities are VRUs. In Rajpura, VRUs account for over 75% of road-traffic deaths (Dhanoa, 2019). In contrast, the extent of fatalities caused by VRUs in high-income countries (HICs) is considerably lower. Specifically, according to the 2016 figures from the World Health Organization, the percentage of VRU mortality is as low as 31% in the USA and 35% in Australia (WHO, 2018).

The annual reports released by the Ministry of Road Transport and Highways (MoRTH) and the National Crime Records Bureau (NCRB) rely on police data gathered from 53 cities having a population of one million or above, as defined by the 2011 Census, commonly known as million-plus cities. Statistics about the cause of road accidents should not be employed for any analysis or policy formulation. All 53 million plus cities had an average mortality rate of 10.5 per 100,000 people in the period from 2019 to 2021, somewhat lower than the national average of 11.4. Visakhapatnam ranked fifth, along with Jaipur, among 53 million-plus cities, with 19.1 deaths per 100,000 people, considerably exceeding the national average.

Several factors contribute to road accidents involving vulnerable road users in India. One of the main issues is the inadequate planning of facilities for VRUs, such as grade-separated crossing facilities, dedicated lanes and specialized signal phases. Additionally, there is a lack of awareness and education regarding road safety among VRUs and other road users. The behaviours of road users, including speeding, drunk driving and distracted driving, also play a significant role in these accidents.

In addition to the loss of lives, road accidents have significant economic costs. They result in medical expenses, property damage and loss of productivity. Families of the victims also suffer emotionally and financially and the strain on healthcare systems can be overwhelming. Furthermore, road accidents can cause disruptions in transportation networks and increase insurance costs for individuals and businesses. Therefore, ensuring the security of VRUs remains a significant societal concern in urban India.

According to the Seventh Pay Commission of India, 2017, Visakhapatnam is classified as a tier-II smart city with a population of 1.73 million citizens, facing several road-safety challenges. The city's rapid changes and population growth have increased traffic congestion, putting VRUs at a higher risk of accidents. The proportion of VRU road-traffic fatalities in Visakhapatnam (2013-2015) was 84% (Tiwari et al., 2023). The road infrastructure may not be adequately equipped to handle the growing number of vehicles, leading to deficiencies in terms of safety measures for VRUs. Understanding the factors influencing fatal and injury crashes on the selected study corridor is crucial for developing effective countermeasures to minimize the number of road accidents and improve road safety. The present study's findings will provide valuable insights for policymakers, traffic authorities and urban planners in implementing targeted interventions and infrastructure improvements to enhance road safety in Visakhapatnam and similar urban areas.

LITERATURE REVIEW

Safety Performance Functions (SPFs) are crucial in road-safety research. They provide a systematic way for transportation engineers and planners to evaluate road-infrastructure safety. SPFs link crash frequency to various factors, such as traffic operational parameters, spatial aspects and built environmental characteristics. For SPF development, traffic-crash modeling traditionally uses Poisson's and Negative Binomial

(NB) regression models. However, Poisson's models are unsuitable for over-dispersion data (i.e., variance > mean); so, negative binomial models are commonly used to address this issue. These models effectively manage over-dispersion in data related to traffic-crash frequency. Using well-established NB regression models, SPF development allows researchers to identify the key factors contributing to crashes and assess their impact on road safety (Lord and Mannering, 2010).

Mukherjee and Mitra (2022) evaluated VRU risk using Poisson's and NB crash-prediction models. The absence of bus stop, ADT, sight distance and speed significantly affect fatal pedestrian accidents. Overtaking study revealed that speeding and lack of sight cause motorcycle deaths. Insufficient street illumination, speed discrepancies and visibility endanger NMT users. Verma et al. (2022) constructed VRU "who-hits-whom" matrices using Bengaluru collision records. The findings consistently support citywide sustainable transportation rules to decrease road-user fatalities. Using discrete modeling, Gandupalli et al. (2023) predicted 60 kilometres of NH VRU-accident severity. Crash timing, season, segment length, accused vehicle type, driver eye clearance, land usage, bends, median openings and pedestrian crossings all impact safety.

Mukherjee and Mitra (2020) examined Kolkata Police crash information to identify pedestrian-hazardous road elements. Factors, such as high approach speed, low pedestrian traffic, overtaking, footpath encroachment, road width, land use, on-street parking, inadequate sight distance, illumination and pavement markings are associated with mid-block road fatalities. Saheli and Effati (2021) used count-based regression models to examine pedestrian accident frequency and road-side land uses on a rural multi-lane in Guilan province, Iran. Santhosh et al. (2020) developed a Poisson's model to predict urban mid-block pedestrian accidents and their severity, including severe, mild and fatal injuries, on NH-66 in Thiruvananthapuram, India. The number of bus stops, traffic volume, vehicle speed, pedestrian volume on walkways and crosswalks, side roads, signalized crossings and crossing quality affect pedestrian accidents.

Naqvi and Tiwari (2020) created SPFs using NB regression models to examine fatal crashes in India's three NH sections. Statistics show that segment length, land use, service roads and terrain type matter. Mitra et al. (2017) examined how access, geometric design and heterogeneous traffic affect multi-lane highway mid-block safety using random parameter panel-data models. Segment length, ADT, horizontal alignment, access-control methods, vehicle type, motorized two-wheeler proportion, higher truck traffic, segments with service roads and median-opening presence are significant.

Dhanoa et al. (2019) constructed NB regression models to explore how geometric design and traffic features impact fatal accident rates in the small Indian towns of Patiala and Rajpura. The paved-road width, number of lanes, segment length and ADT substantially predict fatal accidents. Al-Masaeid and Khaled (2023) employed regression, ANN and other approaches to simulate road-traffic accidents in Jordan during COVID-19 and found that entire and partial travel restrictions cut accidents, injuries and fatalities by 35%, 37% and 50%. Wang et al. (2021) found that driver experience, speed, surface quality, accident timing and seasonal impact have influences on Chinese expressway-truck crash severity. Chengye and Ranjitkar (2013) studied Auckland highway safety with accident-prediction models. AADT per lane, segment length and lane count were found to affect crash frequency the most.

From the above-cited literature, it is evident that significant research has been conducted on pedestrian safety in India and other countries. The literature states that the VRU traffic is significantly restricted in India, primarily emphasizing metropolitan cities. Thus, the current study aims to fill the gap analyzed by previous-research contributions by focusing on NHs passing through medium-sized Indian cities considered tier-II. Furthermore, this study aids in achieving the objective of creating a structural framework which improves the safety of VRUs in Visakhapatnam, a tier-II city in India. By identifying the main factors contributing to VRU crashes and suggesting practical measures to address these risks, the study has the potential to improve road safety and reduce fatalities significantly.

RESEARCH METHODOLOGY

The research methodology encompasses various components, such as crash-data collection, selecting the appropriate study corridor, performing road-safety audits, collecting site-specific data, conducting descriptive analysis, developing suitable statistical models, identifying factors contributing to crashes and suggesting effective countermeasures.

Crash-data Collection and Choosing the Study Corridor

Crash data recorded by the police in many countries may not accurately reflect the true number of traffic fatalities and injuries. Traditionally, it has been presumed that in India, however, many injury cases were treated at private hospitals without being officially documented by the police. In most of the tier-II cities, the majority of fatal and severe injury cases are reported to the police, hence being reliable. The data recorded by the police serves as the exclusive source for the yearly crash reports published by the MoRTH and NCRB. This study used crash data from the Visakhapatnam traffic police department in Andhra Pradesh, India, from 2014 to 2017. The police reports provide valuable information on the vehicles involved, crash type, time, location and count of fatalities and injuries. These reports lack specific details, such as collision diagrams, causes of accidents, socio-demographic information about the crash victims and road infrastructure and geometric conditions at the location.

Geospatially plotting crash sites using Geographic Information Systems (GIS) led to a heat map showing city-wide accident concentrations. According to MoRTH, a "black spot" refers to a specific section of road 500 meters long, where five road accidents resulting in fatalities or severe injuries have occurred within the past three calendar years. Alternatively, it can also be defined as an area where ten or more deaths have occurred over the preceding three calendar years. In the present study, crashes during 2015-2017 were analyzed and it was observed that 80% of the black spots are concentrated on a 62-kilometer stretch of NH-16 passing through Visakhapatnam city. Further research will focus on this corridor. Figure 1 shows a heatmap of Visakhapatnam's accident-prone zones.

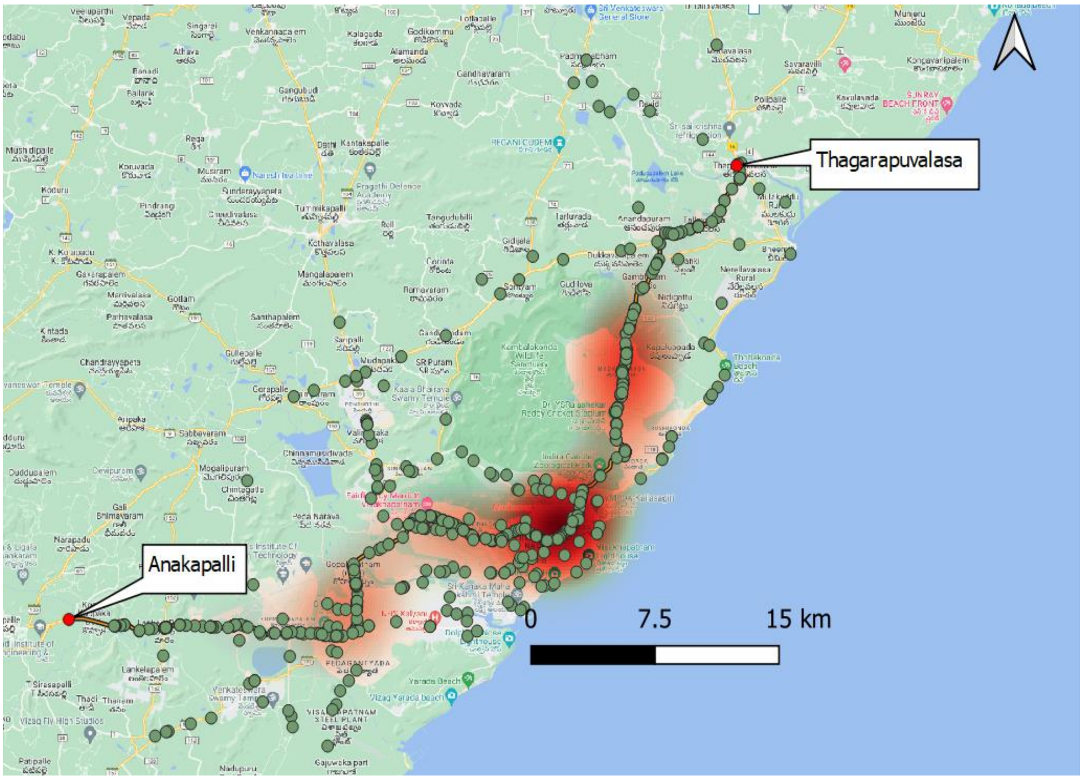

Figure (1): Heatmap of study corridor: accident-prone zones

This study focuses on fatal and injury collisions. This is because fatal and injury collisions in India are less likely to be underreported than crashes with no-death, no-injury crashes. There were 92 road-segments ranging from 0.1 to 2.9 miles in the 62-kilometer examination. A roadway "segment" is the distance between two junctions. This research excluded junction collisions and concentrated on non-intersection and mid-block crashes. According to the literature, intersections within 75-100m of their centres cause collisions (Mitra et al., 2017). The present research excluded crashes inside a 100-meter buffer zone around the junction. The subsequent study included mid-block accident data. A total of 2424 crashes were recorded on a 62-km stretch from Tagarapuvalasa to Anakapalli from 2014 to 2017. Among these, 1771 crashes have been classified as non-intersection mid-block crashes. A total of 372 crashes involving zero deaths and zero injuries were excluded from the analysis and a sample size of 1399 crashes was chosen for subsequent analysis.

Road-safety Audit Data

Road-safety audits were conducted on the entire stretch in 2017; the same data was used for the current study. Data from road-safety audits provide details about road-geometric features, such as curves, median openings, pedestrian crossings, intersections, presence of service roads, side access roads, land-use type, presence of earthen shoulders, proper signage, adequate road markings and sight clearance to the driver. Figure 2 illustrates a handful of the observations made during the road-safety audits with location coordinates.

 Location:17.88452, 83.37737	 Location:17.88452, 83.37737
The unauthorized approach road, speed-breakers missing on the approaching side road, vision obstruction to the driver on the highway, missing signage and road markings	Illegal median cuts for two-wheelers, vision obstruction to the driver on the highway, missing signage and road markings
 Location:17.86963, 83.36867	 Location: 17.85619, 83.36326
Service-road merging at “gap in the median,” insufficient height and width of median, significant number of pedestrians crossing highway during peak hours, absence of speed breakers on the service-road merging, un-authorized stopping of buses on the main carriageway	Illegal crash barrier removal on the curve, zero visibility for the approaching side road, elevation difference between highway and approach road, insufficient earthen shoulder on the curve between the crash barrier and paved shoulder

	
Location:17.825491, 83.35669	Location: 17.83390, 83.35891
A large number of pedestrian movements, improper bus passenger facilities, bus stop located on the curve, bus stops are observed on both the left and right sides of the carriageway	Continuous development of the built-up area on both sides of the carriageway with multiple direct approach roads to the National Highway without any speed breakers or proper signage, frequent cutting of the crash barriers

Figure (2): Observations of road-safety audits with location coordinates

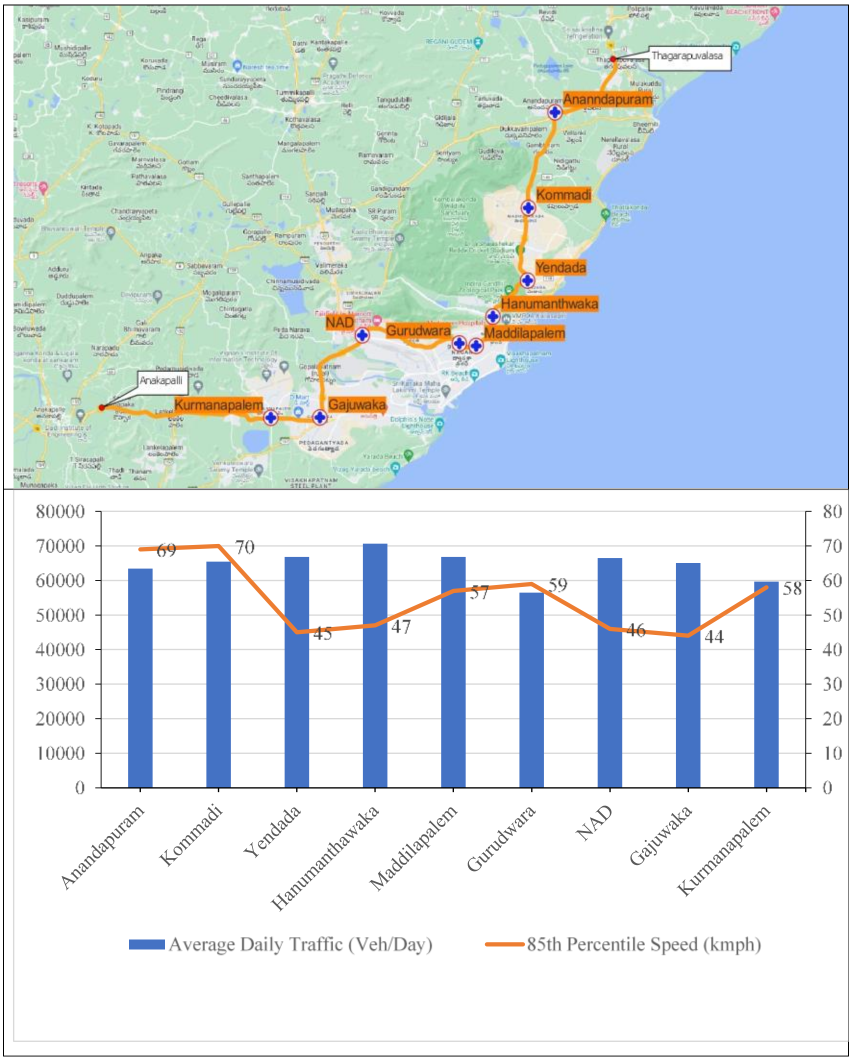

Figure (3): Traffic-volume and spot-speed data-collection locations

Traffic-volume and -speed Data

Traffic-volume studies were conducted at four locations on the stretch during the weekday with normal weather conditions using the videography method. Traffic volume was extracted from the video and the average daily traffic (ADT) and traffic composition were determined. In addition, spot-speed tests were carried out on various sections of the study area and cumulative frequency-distribution curves were created to ascertain the 85th-percentile speeds of each segment. Figure 3 shows traffic-volume and spot-speed data-collection locations.

The study's primary objective is to determine the factors influencing fatal and injury crashes of various victim types, such as pedestrians, two-wheelers, cars, heavy vehicles, three-wheelers and single-vehicles, as well as to propose effective measures to mitigate these risks. The proportion of various victim-type crashes is presented in Table 1. Different response and predictor variables were identified and Tables 2 and 3 display the descriptions of the predictor and response variables.

Table 1. The proportion of various victim-type crashes

	Proportion of various victim types crashes	
Victim Type	**Fatal Crashes**	**Injury Crashes**
Pedestrian involved	53.0	40.5
Two-wheeler	25.4	41.8
Car	2.2	4.2
Heavy vehicle	1.0	2.5
Three-wheeler	1.7	3.2
Two-wheeler single vehicle (TWSV)	12.9	4.5
Single vehicle involved- Others	3.9	3.7

Table 2. Description of predictor variables

Characteristics	**Variable**	**Variable Description**	**Type of Variable**	**Min.**	**Max.**	**Mean**	**Standard Deviation**
Traffic exposure and operational characteristics	Speed	Average 85th-percentile speed	Continuous	36.0	78.0	56.45	12.28
	Log (ADT)	Log (ADT) on segment	Continuous	10.82	11.23	11.11	0.15
	Percentage of Two-Wheelers	Proportion of two-wheelers	Continuous	0.53	0.57	0.55	0.02
	Percentage of Three-Wheelers	Proportion of three-wheelers	Continuous	0.13	0.20	0.17	0.03
	Percentage of Cars	Proportion of cars	Continuous	0.16	0.21	0.19	0.02
	Percentage of Heavy Vehicles	Proportion of heavy vehicles	Continuous	0.07	0.10	0.09	0.01
Infrastructure and roadway characteristics	Segment length	Length of road segment	Continuous	0.11	2.92	0.64	0.61
	Service roads	Service roads (0 if absent, 1 if present)	Categorical	0.00	1.00	0.25	0.44

	Median openings	Number of median openings over each segment	Continuous	0.00	3.00	0.70	0.74
	Curves	Number of curves over each segment	Continuous	0.00	8.00	1.23	1.90
	Pedestrian crossings	Number of pedestrian crossings over each segment	Continuous	0.00	8.00	0.79	1.37
	Earthen shoulder	Presence of earthen shoulder (0 if absent, 1 if present)	Categorical	0.00	1.00	0.41	0.50
	Adequate road markings	Presence of adequate road markings (0 if absent, 1 if present)	Categorical	0.00	1.00	0.83	0.38
	Sight clearance to the driver	Presence of sight clearance to driver (0 if absent, 1 if present)	Categorical	0.00	1.00	0.32	0.47
	Proper signage	Presence of proper signage (0 if absent, 1 if present)	Categorical	0.00	1.00	0.70	0.46
	Side access	Number of access roads over each segment	Continuous	0.00	17.00	3.93	3.50
Land-use characteristics	Land use	Type of land use (1 if commercial/ mixed, 0 if open area/no land use)	Categorical	0.00	1.00	0.61	0.49

Table 3. Description of response variables

Characteristics	Variable	Variable Description	Type of Variable	Min.	Max.	Mean	Standard Deviation
Safety-performance characteristics	2-W fatal crashes	Number of reported two-wheeler fatal crashes at the site (2014-2017)	Continuous	0	7	1.21	1.567
	2-W injury crashes	Number of reported two-wheeler injury crashes at the site (2014-2017)	Continuous	0	22	4.56	5.299
	Ped. fatal crashes	Number of reported pedestrian fatal crashes at the site (2014-2017)	Continuous	0	25	2.34	3.406

	Ped. injury crashes	Number of reported pedestrian injury crashes at the site (2014-2017)	Continuous	0	38	4.37	6.027
	TWSV fatal crashes	Number of reported TWSV fatal crashes at the site (2014-2017)	Continuous	0	4	0.55	0.807
	TWSV injury crashes	Number of reported TWSV injury crashes at the site (2014-2017)	Continuous	0	9	0.49	1.067

Data Modeling

The Poisson's regression model is a good choice for explaining the random occurrence of distinct events. The Poisson's distribution may be stated as (Lord and Mannering, 2010):

$$P[y_i] = \frac{e^{-m_i} m_i^{y_i}}{y_i!} \qquad y_i = 0, 1, 2, \ldots \qquad (1)$$

where $P[y_i]$ stands for the probability that y_i crashes will occur during the counting period on a road segment 'i' and m_i is the average number of incidents happening during a length of time t on a road segment 'i', which is the expected number of incidents throughout the counting period, $E[y_i]$.

The Poisson's distribution is inappropriate when the observed data shows a variance/mean ratio that deviates significantly from 1.0. From Table 3, VRUs were found to be over-dispersed; i.e., the percentage of variance/mean was greater than 1. NB models are frequently employed to address the issue of over-dispersion that typically arises in traffic collision-frequency data. The Poisson's model is extended to include the idea that the Poisson's parameter follows a gamma-probability distribution. This model is known as the NB (or the Poisson-gamma) model. The distribution may be stated as:

$$m_i = e^{(bX_i + \varepsilon_i)} \qquad (2)$$

$$\log(m_i) = b_0 + b_1 X_{i1} + b_2 X_{i2} + b_3 X_{i3} + \cdots + b_n X_{in} + \varepsilon_i \qquad (3)$$

The gamma-distributed error term ε_i has a mean of 1 and variance α. The accumulation of this term enables the variance to deviate from the mean.

$$\text{Variance} = \text{mean} + \alpha \times mean^2. \qquad (4)$$

The parameter α is frequently mentioned as the over-dispersion parameter. Obtaining a substantial value for the parameter α is crucial for an NB model. Ultimately, the log-likelihood ratio index is computed to assess the overall quality of the fit. Count-data models often employ a pseudo-R^2 statistic, commonly called the log-likelihood ratio index (ρ^2), as a standard practice. As stated by Mira et al. (2017), ρ^2 may be expressed as:

$$\rho^2 = 1 - \frac{LL(\beta)}{LL(0)}. \qquad (5)$$

LL(β) represents the log-likelihood value, whereas LL(0) represents the restricted log-likelihood value. Another approach to comparing models' adequacy is utilizing the Akaike Information Criterion (AIC). The AIC value is calculated using the formula:

$$AIC = -2LL(\beta) + 2k. \qquad (6)$$

k represents the model's estimated parameters and the model with the lowest AIC is preferred.

Correlation Analysis

Correlation analysis is an essential statistical tool to examine the relationships between variables. This study conducted correlation analysis to ascertain the extent of correlation between the predictor and responder variables. Significant predictive factors linked to fatal and injury crashes were determined by examining the strength of correlation coefficients. The correlation coefficients range from 0.3 to 0.7, indicating moderate to strong relationships between these variables and the occurrence of accidents at a 95% confidence interval. Table 4 presents the crucial findings from the correlation test. The test results reveal that several predictor variables, such as the length of the segment, speed, service road, the type of land use, median openings, curves, pedestrian crossings, the presence of an earthen shoulder, the proper signage, the presence of side access roads, ADT and the percentage of two-wheelers all show significant correlations with the-probability of both fatal and injury collisions. As part of the multi-collinearity analysis, predictor variables that exhibit a correlation coefficient lower than 0.3 with another predictor variable are consolidated in the model (Mitra et al., 2017). The exclusion of some predictor factors from the models, such as the number of median openings, pedestrian openings, side access roads, proportion of two-wheelers and proper signage, was based on their strong associations with other predictor variables. The graphical representation of inter-correlation among predictor variables is shown in Figure 4.

Table 4. Correlation between predictor and response variables

Predictor Variables	Response Variables					
	2-W Fatal Crashes	**2-W Injury Crashes**	**Pedestrian Fatal Crashes**	**Pedestrian Injury Crashes**	**Single-vehicle 2-W Fatal Crashes**	**Single-vehicle 2-W Injury Crashes**
Segment length	0.398***	0.361***	0.499***	0.267**	0.246**	0.093
Speed	-0.264**	-0.358***	-0.214*	-0.391***	-0.306***	-0.13
Service roads	0.364***	0.390***	0.477***	0.452***	0.147	0.014
Land use	0.266**	0.441***	0.253**	0.424***	0.289**	0.167
Median openings	0.292**	0.224*	0.209*	0.203*	0.195*	-0.011
Curves	0.131	0.232*	0.289**	0.195*	0.168	0.192*
Pedestrian crossings	0.208*	-0.037	0.044	-0.034	0.104	-0.032
Earthen shoulder	-0.093	-0.198*	-0.138	-0.216*	-0.231*	-0.141
Proper signage	0.202*	0.267**	0.331***	0.261**	0.14	0.167
Side access	0.224*	0.277**	0.386***	0.245**	0.102	0.072
Log (ADT)	-0.025	-0.271**	-0.17	-0.292**	-0.156	-0.202*
Proportion of two-wheelers	0.032	0.277**	0.193*	0.271**	0.163	0.209*

* Significant at 90% CI.
** Significant at 95% CI.
*** Significant at 99% CI.

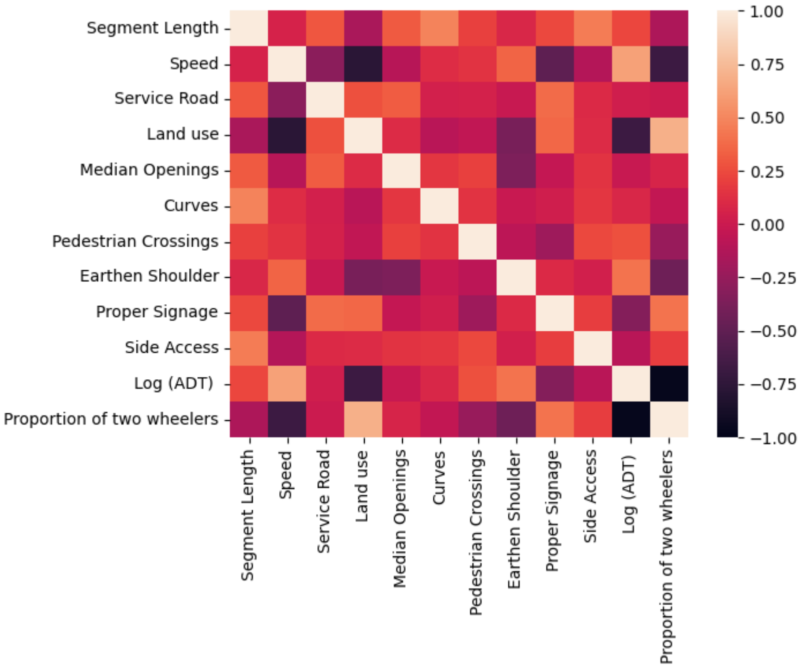

Figure (4): Graphical representation of inter-correlation among predictor variables

MODEL OUTCOMES AND DISCUSSION

In the current investigation, six crash-prediction models were developed to analyze the elements contributing to fatal and injury crashes involving VRUs, including pedestrians, two-wheelers and two-wheeler single-vehicle (TWSV) collisions. These models attempted to ascertain the risk variables linked to traffic exposure, operational attributes, infrastructural and roadway elements and land use. The 'backward-elimination' technique constructed a comprehensive model incorporating many predictor variables. The R Programming software package was used to estimate all statistical models. The importance of the model results from the viewpoint of a developing nation is demonstrated in the subsequent sub-sections.

Modeling of Risk Factors

NB models were developed to determine the primary risk variables linked to fatal and injured two-wheeler, pedestrian and TWSV crashes. The model results are presented in Table 5. The model findings indicate that the length of the segment is in positive and significant correlation with fatal crashes of two-wheelers (β=0.47, $p<0.01$), pedestrians (β=0.76, $p<0.01$) and TWSV (β=0.5, $p<0.01$) at 99% confidence interval and injured crashes of two-wheelers (β=0.51, $p<0.05$) and pedestrians (β=0.51, $p<0.05$) at 95% confidence interval. The study revealed that the lack of service roads is strongly associated with a higher risk of fatal crashes involving two-wheelers (β=-0.66, $p<0.01$) and pedestrians (β=-0.86, $p<0.01$), as well as injuries to two-wheelers (β=-0.45, $p<0.1$) and pedestrians (β=-0.75, $p<0.01$). This absence of service roads reduces the probability of crashes occurring in non-intersection mid-block areas. Additionally, the model revealed that open-area land-use type is negatively correlated with fatal crashes of two-wheelers (β=-0.53, $p< 0.01$), TWSVs (β=-0.85, $p< 0.05$) and injured crashes (β=-1.05, $p< 0.01$), crashes of two-wheelers (β=-0.53, $p< 0.01$) and pedestrians ((β=-0.8, $p< 0.01$). Furthermore, the study discovered significant variables, such as ADT and number of horizontal curves, related explicitly to injury crashes involving TWSVs. The analysis revealed that the number of curves has a positive association with TWSV crashes (β=0.214, $p< 0.05$), while average daily traffic has a negative correlation with injured TWSV crashes (β=-3.3, $p< 0.05$). The adequacy of models was compared by utilizing the Akaike Information Criterion (AIC) and log-likelihood ratio index (ρ^2).

Table 5. Identification of risk factors for two-wheeler, Pedestrian and TWSV crashes

	Two-wheeler Fatal Crash Model		**Two-wheeler Injury Crash Model**		**Pedestrian Fatal Crash Model**		**Pedestrian Injury Crash Model**		**TWSV Fatal Crash Model**		**TWSV Injury Crash Model**	
	Model coefficient (B)	**p-value**	**Model coefficient (B)**	**p-value**	**Model coefficient (B)**	**p-value**	**Model coefficient (B)**	**p-value**	**Model coefficient (B)**	**p-value**	**Model coefficient (B)**	**p-value**
Attributes												
Intercept	0.44	0.067	1.712	0.000	0.723	0.001	1.724	0.000	-0.733	0.005	35.5	0.037
Presence of Service Roads-Not Available	-0.66	0.008	-0.449	0.153	-0.863	0.000	-0.752	0.004	-	-	-	-
Segment length	0.47	0.003	0.509	0.040	0.762	0.000	0.519	0.003	0.499	0.007	-	-
Land-use type- Open area/ no land use	-0.534	0.053	-1.054	0.000	-	-	-0.804	0.001	-0.849	0.034	-	-
Curves	-	-	-	-	-	-	-	-	-	-	0.214	0.055
Log (ADT)	-	-	-	-	-	-	-	-	-	-	-3.3	0.031
Model Summary												
Dispersion parameter	2.71		1.089		5.85		2.009		2.378		0.553	
Log-likelihood function	-100.276		-172.893		-123.488		-163.618		-64.81		-61.858	
Restricted log-likelihood function	-108.824		-225.468		-161.209		-226.621		-76.243		-68.288	
Akaike's Information Criterion (AIC.)	208.76		355.79		254.98		337.24		137.62		131.72	
ρ^2	0.079		0.233		0.234		0.278		0.15		0.094	

Segment Length

The model findings indicate a positive correlation between the segment length and crashes involving two-wheelers, pedestrians and TWSVs. Each unit increase in the segment length results in a 57% to 66% increase in the risk of fatal and injury accidents for two-wheelers, a 114% to 168% increase in the risk of fatal and injury accidents for pedestrians and a 65% increase in the chance of fatal accidents for TWSVs. It was due to road segments with longer stretch lengths being free from the intersection effect tending to facilitate higher vehicle speeds. This creates a high-speed corridor where two-wheelers and other vehicles are more likely to travel at faster speeds, increasing the severity of accidents and reducing reaction time for pedestrians, making them more vulnerable to fatal and injury accidents. Mukherjee and Mitra (2022) discovered that a unit increase in the average speed of a two-wheeler in km/h leads to a 5% increase in the likelihood of fatal crashes. Gandupalli et al. (2023) conducted a similar study using discrete methods and found that a unit increase in the segment length leads to a 17.6% increase in the chance of a VRU fatal crash. In their respective studies, Dhanoa et al. (2019) and Mitra et al. (2017) also identified a positive correlation between segment length and frequency of crashes. Several researchers supported this finding and made similar conclusions (Naqvi and Tiwari, 2020).

Presence of Service Roads

Models revealed that the presence of service roads has a significant relationship with fatal and injured crashes of both two-wheeler and pedestrian accidents, increasing the likelihood of collisions at mid-block portions. Consequently, road segments without service roads reduced the frequency of fatal and injury crashes from 48% to 36% for two-wheelers and from 58% to 53% for pedestrians. Urban regions often feature service roads with widths ranging from 5.5 to 7.5 meters on four- and six-lane NHs. Service roads are typically parallel to the main road, providing a dedicated space for vehicles to slow down, merge or exit the main traffic flow. This design feature helps reduce the likelihood of collisions at intersections by separating turning and merging vehicles from the main traffic stream. Multiple studies have acknowledged that service roads decrease the possibility of crashes at intersections in rural areas (Naqvi and Tiwari, 2020; Mitra et al., 2017). However, service roads in urban areas with significant pedestrian volume may have different effects than those in rural areas with lower pedestrian volume. The presence of service roads in mid-block portions promotes excessive speeds and elevates the likelihood of fatal and injury collisions involving two-wheelers. In addition, it is also seen that service roads increase the crossing width for pedestrians, thereby increasing their vulnerability to crashes.

Land Use

Additionally, it was found from the model that open-area land-use type is negatively correlated with fatal and injury crashes of two-wheelers, injury crashes of pedestrians and fatal crashes of TWSVs. Open-area land-use type may have lower traffic volume, reduced congestion and better visibility, which can lead to fewer fatal and injury crashes for two-wheelers and pedestrians. According to the model, open-area land-use type has 43% and 65% fewer two-wheeler fatal and injury crashes than commercial/mixed land-use segments. Open-area land-use type segments have 55% and 57% less pedestrian injury and TWSV fatal crashes than commercial/mixed land-use parts. It was due to 75% to 80 % of total crashes being non-fatal; vehicles on road segments with commercial/ mixed land use interact more with other cars and pedestrians, leading to more injured crashes. Saheli and Effati (2021) conducted a comprehensive analysis of crash data in urban areas and found that residential and commercial land uses positively affect crash frequency. Their study suggests that mixed traffic, higher pedestrian volumes and increased vehicle interactions in these land-use types contribute to a higher incidence of crashes. Similarly, studies conducted by Mukherjee and Mitra (2020 and 2022) highlighted the significance of land use in fatal pedestrian crashes, emphasizing the need for effective land-use planning to improve road safety.

Number of Curves

The statistical models' findings justified that the number of curves positively correlates with two-wheeler single-vehicle injury crashes. Every unit increase in the number of curves raises the frequency of injury crashes by 24%. Gandupalli et al. (2023) found that a unit increment in the number of curves on a segment leads to a 16.5% rise in injury severity and a 16% rise in fatal crashes. Xin et al. (2017) investigated the impact of horizontal-curve design and related factors on the severity of injuries in single-motor cycle incidents. They showed sharp curves increased crash severity by 7.7% and reverse curves increased it by 26.2% compared to flat curves.

Average Daily Traffic (ADT)

From the model, it was determined that two-wheeler single-vehicle (TWSV) injury crashes exhibit a negative relationship with the logarithmic value of the average daily traffic. This suggests that the probability of two-wheeler crashes decreases as average daily traffic increases. Several researchers (Naqvi and Tiwari, 2020; Mukherjee and Mitra, 2022) studied the association between ADT and the frequency of crashes and found that average daily traffic significantly influences the probability of two-wheeler crashes.

Sensitivity analysis was conducted independently and cumulatively, examining traffic exposure, roadway and land-use factors for each victim-type fatal and injury model. Results were compared with performance metrics 'ρ^2 value' and 'AIC value'. Models with combined factors performed better when compared with those with individual characteristics.

The study was limited to Visakhapatnam city due to the ready availability of the data from the concerned authorities. Due to unavailability of data, it could not compare the models with those of other towns of similar size. However, 20% of the data was excluded from the modeling and was used as controlled data to test the validity of the models.

Model Validation

Validating the accuracy and applicability of a statistical model is an essential stage in the modeling process. It helps ensure that the model is precise on the training data and proficient in creating dependable predictions on new, unobserved data. The present study validated the pedestrian fatal NB model on 21 mid-block road segments in Visakhapatnam city, which accounted for 20% of the total data. Since the accident prediction models didn't include these 21 segments before, the prediction performance of the pedestrian fatal NB model was verified across the 21 segments using accuracy as a performance metric.

$$Accuracy = \frac{number\ of\ true\ prediction}{total\ number\ of\ predictions}. \quad (7)$$

The model underwent validation using crash data recorded by the police (i.e., observed values) and the validation results (i.e., predicted values) were given in the form of a table (Table 6) that indicates the likelihood of success. It was established that the overall accuracy for the number of fatal pedestrian crashes was 66.6%, calculated by summing the diagonal value and dividing it by the entire sample size. Mukherji and Mitra (2020) identified a 55% prediction accuracy for the frequency of fatal pedestrian accidents at intersections. Subsequently, the other models calibrated with the test data also predicted the likelihood of crashes to be from 35% to 48%.

Table 6. Model Validation

		Police recorded crash data (Observed Values)			
	Number of crashes	**0**	**1**	**2**	**3**
Model outcome (Predicted Values)	**0**	0	5	0	0
	1	0	12	1	0
	2	0	0	2	0
	3	0	1	0	0
Total sample size is 21.					
Overall accuracy of the model is 66.6%.					

CONCLUSIONS

Crashes on NH roads passing through urban areas constitute a significant share of traffic incidents in developing nations. This is especially true in medium-sized Indian cities like Visakhapatnam. However, due to the lack of regular crash-data gathering and maintenance, research on individual victim-type crashes at mid-block segments is limited. Segment length, land-use type, the presence of service roads, number of curves and average daily traffic are all recognized risk variables from previous literature and their findings are partly in line with the conclusions of the present study.

This study implemented various count-data models to analyze the relationships between different risk variables and specific victim-type crashes at mid-block segments. The findings of these models revealed exciting patterns and associations. Firstly, segment length was found to have a positive link with two-wheeler, pedestrian and TWSV accidents. Each unit increase in segment length increased the probability of fatal and injury accidents for two-wheelers by 57% to 66%, for pedestrians by 114% to 168% and for TWSVs by 65%. Secondly, the absence of service roads was associated with two-wheeler and pedestrian fatalities and injuries, contrary to previous studies. Thus, road segments without service roads reduced fatal and injury collisions from 48% to 36% for two-wheelers and from 58% to 53% for pedestrians. Although these service roads make merging into and out of local traffic safer and more accessible at intersections, urban mid-block sections tend to increase the likelihood of VRU crashes, as they perpetually become high-speed corridors and increase pedestrian crossing width.

Additionally, the study found that open-area land-use segments had significantly fewer two-wheeler death and injury collisions and pedestrian and TWSV fatal collisions compared to commercial/mixed land use. Because 75–80% of crashes are non-fatal, two-wheelers on commercial/mixed land use intersect with other cars and pedestrians, causing more significant injuries. Finally, the number of curves is positively correlated with TWSV injury collisions, with a 24% increase in injury crashes per unit increase of curves. A negative association was also observed between TWSV injury crashes and logarithmic average daily traffic. These findings enhance our understanding of the variables that impact road safety in urban areas.

For Visakhapatnam and other medium-sized towns in India, there may be specific risk factors for fatal and non-fatal collisions that are not present in larger cities. These risk factors could include higher levels of pedestrian traffic, inadequate pedestrian infrastructure or a higher prevalence of two-wheelers on the roads. Understanding these unique risk factors is crucial in developing adequate safety precautions to address the specific needs of different victim types. Based on many distinct findings from this investigation, the following immediate safety precautions are advised.

1. The longer mid-block road segments functioned as high-speed corridors. These corridors may be effectively regulated by imposing appropriate speed restrictions, implementing traffic calming measures and monitoring vehicle speed through speed cameras. The existing policies set by MoRTH prohibit the installation of speed bumps on National Highways and establish a maximum-speed limit of 60 km/h in urban areas. However, to enhance the safety of VRUs on highways passing through urban areas, it is beneficial to implement measures, such as decreasing the maximum-speed restriction from 60 km/h to 40 km/h in high-speed corridors, as well as installing speed bumps or roundabouts at regular intervals along these corridors.
2. Currently, MoRTH does not have explicit regulations on safety measures for pedestrians crossing mid-block road segments with parallel service roads in urban areas. Implementing stringent regulations, such as dedicated pedestrian crossings with refugee islands and distinct signal phases at regular intervals, particularly near bus stops and school zones, can enhance pedestrian safety. Furthermore, it is necessary to implement a policy that involves erecting barriers in the median to prevent undesignated pedestrian entry. This will also decrease the frequency of pedestrian-involved crashes.
3. Furthermore, it is necessary to establish a distinct classification of roads based on land uses in addition to the conventional functional categories. Installing traffic-control devices and infrastructure facilities ideal for a specific land-use road segment is necessary.
4. Our study deduced that the number of curves on a road segment that influence crashes involving two-wheelers, mandatory policies on implementation and regular maintenance of appropriate warning signs, increased lighting and chevron markers could reduce the vulnerability towards curve crashes.

In addition to the above-mentioned recommendations, there should be a heightened effort to enhance awareness among road users. To improve people’s perception of road safety, it is imperative to frequently conduct workshops, disseminate awareness messages on social media, launch awareness campaigns and arrange special events for schools, colleges, cooperate offices, industries, ... etc.

Future Research

The exclusion of factors, such as the number of median openings, pedestrian crossings, side-access roads, proportion of two-wheelers and proper signage, was based on multi-collinearity with other predictor variables. Although these factors were not incorporated into the models, it is crucial to recognize that they might still influence road accidents. Hence, future research should thoroughly examine these variables. Including socio-demographic data of crash victims and comparative analysis with other regions with similar urban characteristics should be considered for future crash studies on vulnerable road users (VRUs).

Acknowledgements

The authors thank the State of Andhra Pradesh Transport Department and the Traffic Police Authorities of Visakhapatnam city for their aid and support with the project’s data components.

REFERENCES

Al-Masaeid, H.R., and Khaled, F.J. (2023). "Performance of traffic-accident prediction models". Jordan Journal of Civil Engineering, 17 (1), 34-44.

Chengye, P., and Ranjitkar, P. (2013). "Modeling motorway accidents using negative binomial regression". Journal of the Eastern Asia Society for Transportation Studies, 10, 1946-1963.

Dhanoa, K.K., Tiwari, G., and Manoj, M. (2019). "Modeling fatal traffic-accident occurrences in small Indian cities, Patiala and Rajpura". International Journal of Injury Control and Safety Promotion, 26 (3), 225-232.

Gandupalli, S.R., Kokkeragadda, P., and Dangeti, M.R. (2023). "Analysis and modeling of crash severity of vulnerable road users through discrete methods: A case-study approach". Innovative Infrastructure Solutions, 8 (298), 1-14.

Goel, R., Jain P., and Tiwari, G. (2018). "Correlates of fatality risk of vulnerable road users in Delhi". Accident Analysis & Prevention, 111, 86-93.

Lord, D., and Mannering, F. (2010). "The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives". Transportation Research-Part A, 44, 291-305.

Ministry of Road Transport & Highways. (2023). "Road accidents in India 2022". Transport Research Wing, New Delhi.

Mitra, S., Haque, M., and King, M.J. (2017). "Effects of access, geometric design and heterogeneous traffic on safety performance of divided multi-lane highways in India". Journal of Transportation Safety & Security, 9 (1), 216-235.

Mohan, D., Tiwari, G., and Mukherjee, S. (2016). "Urban traffic-safety assessment: A case study of six Indian cities". IATSS Research, 39 (2), 95-101.

Mukherjee, D., and Mitra, S. (2022). "Development of a systematic methodology to enhance the safety of vulnerable road users in developing countries". Transportation in Developing Economies, 8 (28), 1-17.

Mukherjee, D., and Mitra, S. (2020). "Modeling risk factors for fatal pedestrian crashes in Kolkata, India". International Journal of Injury Control and Safety Promotion, 27 (2), 197-214.

Mukherjee, D., and Mitra, S. (2020). "Identification of pedestrian risk factors using negative binomial model". Transportation in Developing Economies, 6 (4), 1-14.

Naqvi, H.M., and Tiwari, G. (2020). "Safety-performance functions for fatal crashes on national highways in India". Procedia-Transportation Research, 48, 1185-1199.

Saheli, M.V., and Effati, M. (2021). "Segment-based count regression geospatial modeling of the effect of roadside land uses on pedestrian crash frequency in rural roads". International Journal of Intelligent Transportation Systems Research, 19, 347-365.

Santhosh, A., Sam, E., and Bindhu, B.K. (2020). "Pedestrian accident-prediction modeling: A case study in Thiruvananthapuram city". In: Mathew, T.V., Joshi, G.J., Velaga N.R., and Arkatkar, S. (Eds.). Transportation Research, Springer, Singapore, 637-645.

Tiwari, G., Goel, R., and Bhalla, K. (2023). "Road safety in India: Status report 2023". Transportation Research & Injury Prevention Centre, Indian Institute of Technology, Delhi.

Verma, A., Sasidharan, S., Bhalla, K., and Allirani, H. (2021). "Fatality-risk analysis of vulnerable road users from an Indian city". Case Studies on Transport Policy, 10 (1), 269-277.

Xin, C., Wang, Z., Lee, C., and Lin, P. (2017). "Modeling safety effects of horizontal-curve design on injury severity of single-motorcycle crashes with mixed-effect logistic model". Transportation Research Record: Journal of the Transportation Research Board, 2637 (1), 38-46.

Wang, Y., Zhang, H., and Shi, N. (2021). "Factors contributing to the severity of heavy-truck crashes: A comparative study of Jiangxi and Shaanxi, China". Jordan Journal of Civil Engineering, 15 (1), 41-51.

World Health Organization. (2018). "Global-status report on road safety 2018". ISBN 978-92-4-156568-4.