



Pedestrian crash causation analysis near bus stops: Insights from random parameters Negative Binomial–Lindley model

Mohammad Anis^a, Srinivas R. Geedipally^b, Dominique Lord^a

^a Zachry Department of Civil & Environmental Engineering, Texas A&M University, College Station, TX 77843, USA

^b Center for Transportation Safety, Texas A&M Transportation Institute, 111 RELLIS Parkway Bryan, TX 77807, USA



ARTICLE INFO

Keywords:

Pedestrian safety
Bus stops
Contributing factors
RPNB-L
Full Bayes PSI
Hazardous stops
Transit corridors

ABSTRACT

Pedestrian safety remains a pressing concern near bus stops along urban transit, where frequent pedestrian–vehicle interactions occur. While prior research has primarily focused on intersections and midblock locations, bus stops have often been treated as secondary contributors rather than as distinct sites requiring targeted safety assessments. This has left a critical gap in understanding how traffic exposure, roadway characteristics, and bus stop design features specifically influence pedestrian crash risks around bus stop locations. To address these gaps, this study develops a comprehensive framework focused on pedestrian safety in the vicinity of bus stops. The proposed approach employs a Random Parameters Negative Binomial–Lindley (RPNB-L) model to account for unobserved heterogeneity and site-specific variability. Using data from 596 bus stops in Fort Worth, Texas (2018–2022), the model identifies that higher pedestrian crash frequencies are significantly associated with increased AADT, elevated boarding activity, and the absence of key safety elements such as crosswalks, medians, and lighting. Conversely, far-side bus stop placement, signalized intersections, sidewalks, and mixed-use development are associated with lower crash risks. Roads near schools and those with speed limits of ≤ 35 mph show elevated crash risk. To support proactive safety management, the study integrates a Full Bayes-based Potential for Safety Improvement (PSI) metric, enabling the identification of hazardous stops and high-risk corridors. By unifying advanced count-based modeling with strategic risk prioritization, this research offers actionable, data-driven insights for improving pedestrian safety near bus stops.

1. Introduction

Pedestrian safety in the United States has emerged as a critical public health issue, with pedestrian fatalities increasing substantially even as overall vehicle-related deaths have declined. Between 2020 and 2021, pedestrian fatalities increased by 12.54%, and over the past two decades (2001–2021), pedestrian deaths increased by 50.74%, while vehicle occupant fatalities declined by 4.67% (Stewart, 2023). Since 2009, pedestrian fatalities have escalated by 59%, highlighting the need for enhanced safety measures (Sanchez Rodriguez and Ferenchak, 2024). In Texas, these trends are similarly concerning, with pedestrians accounting for 20% of traffic-related deaths in 2020 (Stewart, 2023). In 2021 alone, 5370 pedestrian crashes resulted in 843 fatalities and 1467 severe injuries. Due to their vulnerability, pedestrians are approximately 1.5 times more likely to suffer fatal injuries compared to vehicle occupants (Li et al., 2017; Shinar, 2012; Beck et al., 2007). Fig. 1 illustrates the widening gap between pedestrian and vehicle fatalities, emphasizing the need for targeted safety interventions.

Research on pedestrian crashes has predominantly explored contributing factors and hazardous locations (Xie et al., 2017; Lee et al., 2015), demonstrating that roadway design, vehicle characteristics, and user behaviors collectively shape crash outcomes. Such findings support evidence-based strategies for enhancing pedestrian safety, improving infrastructure, and informing policy decisions (Das et al., 2019). While intersections (Haleem et al., 2015) and mid-block areas (Toran Pour et al., 2017; Quistberg et al., 2015a) have received significant attention, pedestrian crashes occurring specifically at or near bus stops have been comparatively neglected. Typically, bus stops have been viewed merely as contributing factors rather than distinct safety assessment sites, leaving a notable gap in understanding the unique risk factors and safety conditions associated with these locations. Factors such as traffic exposure, roadway environment, and specific bus stop designs have not been comprehensively evaluated in relation to pedestrian crashes, despite their potential to significantly influence crash risk. Additionally, the lack of a standardized analytical framework limits the identification and prioritization of hazardous bus stops within transit corridors.

* Corresponding author.

E-mail address: mohammad.anis@tamu.edu (M. Anis).

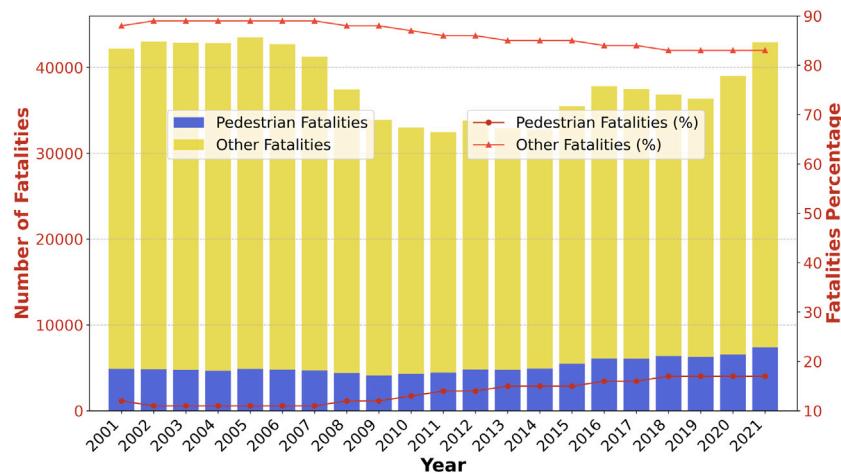


Fig. 1. Pedestrian and vehicle fatalities (2001–2021).

Bus stops inherently present distinct safety challenges due to high pedestrian volumes, frequent street crossings by pedestrians rushing to catch buses, and complex traffic dynamics (Yendra et al., 2024). Prior studies have indicated that the placement of bus stops near intersections or along arterial roads increases pedestrian crash risks significantly (Hu and Cicchino, 2018). Campbell et al. (2003) reported that approximately 2% of urban pedestrian crashes occur in the vicinity of bus stops. Meanwhile, AASHTO (2010) has noted that having one or two bus stops within 1000 ft of an intersection raises pedestrian crash risks by 2.78 times, and by 4.15 times when more than two stops are present. Moreover, the growing prevalence of larger vehicles further exacerbates these risks (Tyndall, 2021), particularly on busy transit corridors where multiple bus stops may be closely spaced (Hess et al., 2004), making transit corridors a focal point of concern (Craig et al., 2019; Ulak et al., 2021).

Moreover, existing analytical frameworks often rely on traditional statistical models such as Poisson and Negative Binomial (NB), which frequently fail to accommodate the unique characteristics of pedestrian crash data, including randomness and a high proportion of zero-crash observations. These models assume that all locations have a probability of experiencing crashes. Although Zero-Inflated Poisson (ZIP) and Zero-Inflated Negative Binomial (ZINB) models attempt to address excess zeros, these approaches often emphasize model fit over interpretability, potentially limiting their practical utility (Lord et al., 2005, 2007). Alternative models, such as Negative Binomial–Lindley (NB–L) (Geedipally et al., 2012; Lord and Geedipally, 2011) and Negative Binomial–Generalized Exponential (NB–GE) (Vangala et al., 2015), have demonstrated superior performance for highly skewed datasets. Few studies compared the performance of ZINB, NB, NB–L, NB–GE (Lord and Geedipally, 2011; Vangala et al., 2015), these have shown that NB–L provides superior statistical fit, particularly for highly skewed datasets. Another persistent challenge lies in accounting for unobserved heterogeneity caused by spatial correlations, temporal variability, and omitted variables (Mannering et al., 2016). Random Parameters (RP) methods directly address this issue by allowing model coefficients to vary across sites, with the RPNB–L model showing particular promise when excess zeros and unobserved heterogeneity overlap (Rusli et al., 2018).

Accordingly, the main objective of this study is to develop a comprehensive framework for assessing pedestrian safety near bus stops, one that explicitly accounts for traffic exposure, roadway characteristics, and bus stop design features. More specifically, this study (i) proposes a pedestrian crash analysis framework tailored to bus stops along urban transit corridors; (ii) applies a random parameters approach to capture unobserved heterogeneity driven by site-specific variations; and (iii) employs a Full Bayes (FB) based Potential for Safety Improvement (PSI)

metric to identify and prioritize high-risk transit corridors proactively. By addressing notable gaps in the existing literature, this work offers a systematic approach to understanding and mitigating pedestrian crashes near bus stops.

To accomplish these goals, the study integrates five years (2018–2022) of pedestrian crash data from 596 bus stop locations in Fort Worth, Texas, with complementary datasets on roadway conditions, traffic exposure, and bus stop design characteristics. Traffic exposure data were obtained from local agency records, while roadway and design attributes were extracted using the Rhino database and Google Earth & Street View. To effectively capture site-specific variation and unobserved heterogeneity in crash frequencies, the study adopts the RPNB–L model as its core analytical method. Baseline models, including NB–L and RPNB–GE, are also estimated for comparison. Additionally, the Full Bayes-based PSI metric is applied to systematically identify hazardous bus stops and transit corridors, enabling a proactive, data-driven safety management approach. By bridging these critical gaps in the literature, this study contributes a novel, transferable framework for evaluating and improving pedestrian safety near bus stops in urban settings.

The structure of this study is organized as follows: Section 2 provides an extensive review of the relevant literature. Section 3 describes the methodology, model estimation process, inference, and validation. Section 4 covers the data description. The model estimation results and a brief discussion are presented in Section 5. Section 6 summarizes the key points and presents the conclusions.

2. Background

Bus stops are critical interaction points where pedestrians, cyclists, passengers, and vehicles frequently converge, increasing the likelihood of crashes (Torbic et al., 2010). Vulnerable road users, such as pedestrians, passengers, etc., face significant safety risks near these locations (Zhang et al., 2023). Studies indicate that 89% of high-crash locations are within 150 ft of bus stops, and 90% occurring within 70 ft of crosswalks (Walgren, 1998). These findings highlight the need for enhanced pedestrian safety near bus stops.

Various methods have been employed to analyze pedestrian safety, including historical crash data analysis (Craig et al., 2019; Hess et al., 2004), geography information system-based safety inspections (Yu, 2024), road user observations (Akintayo and Adibeli, 2022), and traffic conflict techniques (Zhang et al., 2023). Over time, crash-based analyses have utilized statistical models to establish relationships between crash frequency and explanatory variables (Thakali et al., 2015). These approaches provide a comprehensive understanding of influencing factors, such as exposure-related behaviors, roadway conditions, bus stop design elements (Clifton et al., 2009; Craig et al., 2019;

Haleem et al., 2015; Hu and Cicchino, 2018; Zahabi et al., 2011; Zegeer and Bushell, 2012). However, prior studies have often lacked a holistic approach, potentially leading to misleading conclusions about the effects of certain factors.

Pedestrian activity is particularly high near bus stops, especially in high-traffic areas such as schools and commercial zones, which heightens the risk of collisions (Geedipally, 2021; Craig et al., 2019; Hess et al., 2004). In addition to pedestrian–vehicle interactions, bus stops often serve as high-conflict areas for pedestrian–cyclist interactions, particularly where segregated bike paths or shared-use lanes exist. Afghari et al. (2014) examined such conflicts in Montreal using automated video-based conflict analysis. Their study found that pedestrians frequently took evasive actions (e.g., stopping or running) to avoid cyclists, whereas cyclists generally maintained their speeds. Risky pedestrian behaviors, such as crossing midblock or rushing to board buses, further increase crash risks. Risky pedestrian behaviors, such as crossing without waiting for signals or rushing to board buses, further contribute to crash risks (Quistberg et al., 2015b). Additionally, crossing roads at midblock instead of using designated crosswalks significantly increases crash potential (Pessaro et al., 2017). A study found that 59% of pedestrian crashes occurred when individuals rushed to catch a bus (Wretstrand et al., 2014). Additionally, waiting in undesignated areas may cause buses to stop improperly, increasing the likelihood of crashes (Chand and Chandra, 2017). These behaviors underscore passenger activity as a critical factor influencing crash frequency, with boarding and alighting volumes serving as key indicators of pedestrian risk near bus stops.

Bus stop sign infrastructure is essential to pedestrian safety. Clear and visible signage ensures that pedestrians wait in designated areas and helps bus drivers identify stopping points. When signage is obstructed by trees or advertisements, the area becomes more hazardous (Rossetti et al., 2020). Bus shelters provide accessible waiting spaces, and the absence of such facilities has been linked to the most hazardous bus stop locations (Salum et al., 2024). These risks are further exacerbated by inadequate pedestrian infrastructure, such as marked crosswalks or pedestrian signals (Pulugurtha and Vanapalli, 2008). Several studies underscore the importance of lighting for safety at bus stops, especially at nighttime (Rossetti et al., 2020; Mukherjee et al., 2023), with poor lighting linked to higher pedestrian fatality rates, especially in sparsely populated areas (Lakhotia et al., 2020). Many studies recommend improved lighting at bus stops and their approaches to enhance safety (Salum et al., 2024; Pessaro et al., 2017).

The location of bus stops also influences pedestrian crash risks due to factors such as driver distraction and poor yielding behavior (Craig et al., 2019). Fitzpatrick and Nowlin (1997) examined the safety impacts of bus stop placement and found that far-side stops (positioned after an intersection) experience fewer crashes than near-side stops due to reduced conflicts with right-turning vehicles. However, far-side stops can obstruct intersections during peak hours, diminishing visibility for both drivers and pedestrians (Institute and Foundation, 1996).

Safe and accessible pedestrian infrastructure around bus stops strongly correlates with increased transit access and safety (Sukor and Fisal, 2020). The absence of sidewalks forces pedestrians to walk on roadways, increasing their exposure to traffic and crash risk (Rossetti et al., 2020). Similarly, the presence of crosswalks improves pedestrian safety (Ulak et al., 2021), yet over 30% of bus stops lack them, making these areas more hazardous (Tiboni and Rossetti, 2013). Studies indicate that a substantial proportion of pedestrian crashes occur due to non-use of crosswalks, with two-thirds of pedestrian fatalities occurring under such conditions (Clifton et al., 2009). At unsignalized intersections, pedestrian injury severity is significantly influenced by crosswalk availability (Clifton et al., 2009). Additionally, bus stops near signalized intersections exhibit higher pedestrian crash rates due to increased pedestrian volume and transit activity (Pulugurtha and Sambhara, 2011). Several studies acknowledge this association, emphasizing that bus stops near intersections pose heightened pedestrian risks (Harwood

et al., 2008; Kuşkapan et al., 2022; Walgren, 1998). Zegeer and Bushell (2012) found that the likelihood of pedestrian crashes increases significantly when bus stops are located within 1000 ft of an intersection, especially when combined with high traffic and pedestrian volumes, multiple lanes without a median, and nearby schools. Similarly, Torbic et al. (2010) developed a method to predict vehicle–pedestrian crashes, showing that crashes are four times more likely when a bus stop is within 1000 ft of a signalized intersection, mainly when schools or alcohol retail stores are present.

Several studies have examined the relationship between the built environment, traffic characteristics, and crash frequency, considering factors such as traffic volume, speed, and roadway geometry (Hess et al., 2004; Ulak et al., 2021; Lakhotia et al., 2020; Tiboni and Rossetti, 2013). For example, high average annual daily traffic (AADT) near bus stops increases potential conflicts between vehicles and pedestrians. High traffic speeds are also a major safety risk, particularly in urban areas where pedestrian activity is high (Rossetti et al., 2020). Furthermore, bus stops located along high-speed roads are linked to increased pedestrian fatalities (Hess et al., 2004; Lakhotia et al., 2020; Pulugurtha and Vanapalli, 2008). Roadway width also plays a role in pedestrian risk, as wider roads complicate pedestrian crossings and increase crash likelihood, whereas narrower roads may result in more pedestrian–vehicle conflicts (Zhang et al., 2023). The presence of a median near bus stops enhances safety by providing pedestrians with a refuge area, allowing them to cross in stages (Pessaro et al., 2017; Yu, 2024). Land use also plays a crucial role in pedestrian safety near bus stops. High-activity areas, such as shopping centers, schools, and workplaces, attract significant pedestrian and vehicular traffic, increasing collision risks (Hess et al., 2004). Areas with high vehicular activity, mainly commercial zones like big box stores, are associated with more pedestrian crashes due to increased traffic flow and parking-related movements (Yu, 2024).

Several researchers have conducted investigations about identifying hotspots with a high incidence of pedestrian–vehicle collisions and bus stops that pose safety risks (Truong and Somenahalli, 2011; Ulak et al., 2021). Truong and Somenahalli (2011) proposed a GIS-based approach utilizing spatial autocorrelation analysis and severity indices to rank unsafe bus stops based on pedestrian crash data. Ulak et al. (2021) introduced the Bus Stop Safety Index (SSI), a quantitative measure that incorporates injury severities, pedestrian crash proximity, socio-demographic factors, traffic conditions, and transit stop attributes to identify high-risk bus stops in Palm Beach County, FL. This metric facilitates targeted interventions to enhance pedestrian safety and improve urban transit system accessibility.

While previous studies have acknowledged the significance of bus stops in pedestrian safety, they often fail to provide a comprehensive framework that integrates exposure characteristics, roadway environments, and bus stop design elements in assessing crash risk. Most studies consider bus stops as secondary variables rather than directly examining how infrastructure, traffic, and environmental factors interact to influence the frequency of crashes. To address this gap, the present study employs advanced statistical models that integrate diverse categorical variables to quantify crash risks and identify hazardous transit corridors. These models aim to generate actionable insights, supporting transportation agencies in developing targeted interventions to enhance pedestrian safety near bus stops.

3. Methodology

The primary objective of this study is to develop and apply a robust statistical modeling framework to investigate the relationship between pedestrian crashes and key contributing factors near bus stops. Based on the research gaps and objectives outlined in the previous sections, multiple statistical modeling approaches have been employed, informed by established crash frequency modeling research. This section provides a concise overview of the selected modeling methodologies and describes the evaluation criteria used to compare their performance and interpretability.

3.1. Negative Binomial-Lindley (NB-L)

The NB-L model (Geedipally et al., 2012; Rahman Shaon and Qin, 2016) is an extension of the standard NB model designed to handle highly overdispersed count data. Geedipally et al. (2012) demonstrated that the NB-L model performs better than the NB model when the data contains many zeros or is characterized by a long tail. The NB-L model introduces an additional layer of flexibility by combining the NB distribution with the Lindley distribution to better account for unobserved heterogeneity and excess variation in the data. Therefore, the mixture of the NB and Lindley distributions is as follows:

$$P(Y_i = y_i | \mu_i, \phi, \theta) = \int \text{NB}(y_i | \phi, \lambda_i \mu_i) \text{Lindley}(\lambda_i | \theta) d\lambda_i \quad (1)$$

The random error term is represented by λ_i , while the Lindley distribution is denoted by θ . In Eq. (1), both λ_i and Y_i are assumed to follow a NB distribution, with the mean defined as $\lambda_i \mu_i$. The corresponding inverse dispersion parameter is expressed as $\phi = \frac{1}{\alpha}$, where λ_i is governed by the Lindley distribution parameter θ .

Here, the Lindley distribution can be presented as Lord et al. (2021) is defined as:

$$\text{Lindley}(x|\theta) = \frac{\theta^2}{\theta+1} (1+x)e^{-\theta x}, \quad \theta > 0, x > 0 \quad (2)$$

The mean response function for the NB-L can be expressed as:

$$E(Y_i = y_i) = \mu_i \cdot E(\lambda_i) \quad (3)$$

where $\mu_i = \exp(b_0 + \sum_{j=1}^q b_j X_j)$ represents the expected number of crashes, and $E(\lambda_i) = \frac{\theta+2}{\theta(\theta+1)}$ is the expected value of the Lindley-distributed term. Replacing the values for μ_i and $E(\lambda_i)$, the mean response function becomes:

$$\begin{aligned} E(Y_i = y_i) &= \exp \left(b_0 + \sum_{j=1}^k b_j X_j \right) \cdot \frac{\theta+2}{\theta(\theta+1)} \\ &= \exp \left(b_0 + \log \left(\frac{\theta+2}{\theta(\theta+1)} \right) + \sum_{j=1}^k b_j X_j \right) \end{aligned} \quad (4)$$

To simplify, we incorporate the term $\log \left(\frac{\theta+2}{\theta(\theta+1)} \right)$ into a new intercept term, b'_0 , yielding:

$$E(Y_i = y_i) = \exp \left(b'_0 + \sum_{j=1}^k b_j X_j \right) \quad (5)$$

where:

$$b'_0 = b_0 + \log \left(\frac{\theta+2}{\theta(\theta+1)} \right)$$

Thus, the NB-L mean response function accounts for over-dispersion and excess zeros by adjusting the intercept term to include the effects of the Lindley distribution.

The Lindley distribution is a mixture of two Gamma distributions as follows (Zamani and Ismail, 2010):

$$\lambda_i \sim \frac{1}{1+\theta} \text{Gamma}(2, \theta) + \frac{\theta}{1+\theta} \text{Gamma}(1, \theta) \quad (6)$$

The Eq. (6) alternatively, it can be restructured as according to Lord et al. (2021)

$$\lambda_i \sim \sum \left\{ \text{Gamma}(1+z_i, \theta) \text{Bernoulli} \left(z_i; \frac{1}{1+\theta} \right) \right\} \quad (7)$$

Therefore, the NB-L model can be written as the following hierarchical model (Geedipally et al., 2012; Lord et al., 2021; Khodadadi et al., 2023):

$$P(Y_i = y_i; \phi, \mu_i | \lambda_i) = \text{NB}(y_i; \phi, \lambda_i \mu_i) \quad (8)$$

$$\mu_i = \exp \left(b_0 + \sum_{j=1}^q b_j X_j \right) \quad (9)$$

$$\lambda_i \sim \text{Gamma}(1+z_i, \theta) \quad (10)$$

$$z_i \sim \text{Bernoulli} \left(\frac{1}{1+\theta} \right) \quad (11)$$

Prior distributions must be specified to estimate the model parameters using a Bayesian framework. These priors reflect past knowledge or assumptions about the parameters. The random error λ term is given a non-informative prior from a gamma distribution, while the gamma shape parameter is governed by a Bernoulli distribution with probability $1/(1+\theta)$. When weakly informative priors are used, the Lindley distribution can influence the model more than the NB component. One challenge in Bayesian estimation is the poor mixing of the MCMC process caused by the intercept and error term correlation. Geedipally et al. (2012) recommend employing informative priors when possible to mitigate this issue. A prior should be chosen to ensure that $E(\epsilon) = 1$ limits the influence of the Lindley distribution. Geedipally et al. (2012) advocate for the use of a prior on $1/(1+\theta)$ that follows a beta distribution, specifically Beta($n/3$, $n/2$), where n represents the number of observations. This approach enhances model stability and improves the accuracy of parameter estimation.

3.2. Random Parameters Negative Binomial-Lindley (RPNB-L)

Random parameters (RP) models an alternative way to account for unobserved heterogeneity by allowing parameters to vary from one observation to another (Anastasopoulos and Mannerling, 2009; Barua et al., 2016; Shaon et al., 2018; Rusli et al., 2018; Khan et al., 2023). To form this randomized variable, let us assume the coefficient β_{ij} associated with the j th covariate for the i th site and which can be expressed as:

$$\beta_{ij} = b_j + v_{ij} \quad (12)$$

where b_j represents the fixed mean estimate of the parameter, and v_{ij} is the random term, which typically follows a normal distribution with a mean of zero and a variance of σ^2 . This v_{ij} is included in the model if its standard deviation is significantly different from zero; otherwise, a fixed parameter b_j is applied uniformly across all observations (El-Basyouny and Sayed, 2009; Anastasopoulos and Mannerling, 2009). The probability mass function (pmf) of the RPNB model is given by:

As mentioned above, the NB-L model itself is the RP model since it accommodates the Lindley distribution by introducing site-specific effects, which coefficient fully transform to randomized by introducing RPNB-L, meaning that the coefficients of the covariates vary across sites, not just the Lindley portion. This leads to a hierarchical model structure, according to Shaon et al. (2018):

$$\begin{aligned} P(Y_i = y_i; \phi, \mu_i | \lambda_i) &= \text{NB}(y_i; \phi, \lambda_i \mu_i) \\ \mu_i &= \exp \left(b_0 + \sum_{j=1}^k b_{ij} X_{ij} \right) \end{aligned} \quad (13)$$

In this hierarchical structure, the random effect λ_i follows a gamma distribution, $\lambda_i \sim (\lambda|1+z_i, \theta)$, and z_i follows a Bernoulli distribution $z_i \sim \text{Bernoulli}(z_i | \frac{1}{1+\theta})$. The random parameter $b_{ij} = b_j + v_{ij}$ captures site-specific variability, where $v_{ij} \sim \mathcal{N}(0, \sigma_j^2)$.

A known issue with RPNB-L models, especially in a Bayesian framework, is poor mixing of the MCMC chains, mainly due to correlations between the intercept and coefficients, both of which vary across observations. One solution is to standardize the coefficients before using them in the model. After the MCMC chains converge, the standardized coefficients can be transformed back to their original scale using the following transformations for a more straightforward interpretation:

$$x_{ij}^* = \frac{x_{ij} - \mu_j}{\sigma_j} \quad (14)$$

where μ_j and σ_j are the mean and standard deviation of the j th covariate. After the analysis, the estimated coefficients are transformed back to the original scale using (Gelfand et al., 1995):

$$b_1 = \frac{b_1^*}{\sigma_1}, \dots, b_q = \frac{b_q^*}{\sigma_q}, \quad b_0 = b_0^* - \sum_{j=1}^q \frac{b_j \mu_j}{\sigma_j} \quad (15)$$

The random parameters b_{ij} are modeled using a normal prior:

$$b_{ij} \sim \mathcal{N}(b_j, \sigma_j^2), \quad \frac{1}{\sigma_j^2} \sim \Gamma(0.01, 0.01) \quad (16)$$

Previous research has shown that normal distributions typically provide the best fit for modeling the RP compared (Anastasopoulos and Manning, 2009) to using lognormal, uniform, gamma, triangular, etc., which helps ensure adequate MCMC mixing and improves the robustness of the model's estimates.

3.3. Random Parameters Negative Binomial – Generalized Exponential (RPNB-GE)

The RPNB-GE is an extension of the standard NB model, formulated to account for overdispersion, excess zeros, and unobserved heterogeneity in crash history data (Vangala et al., 2015). The GE distribution is used to introduce an additional layer of flexibility for capturing these variations. For a deeper understanding, readers are directed to Aryuyuen and Bodhisuwann (2013). The NB-GE model assumes that the mean crash rate is adjusted by a Generalized Exponential-distributed random effect, leading to the following formulation:

$$P(Y_i = y_i | \mu_i, \phi, a, b) = \int \text{NB}(y_i | \phi, \theta_i) \text{GE}(\delta_i | a, b) d\delta_i \quad (17)$$

where:

$$\theta_i = \mu_i \delta_i$$

$$\delta_i \sim \text{Generalized Exponential}(a, b)$$

Here, θ_i is the product of μ_i and δ_i , representing the mean crash rate adjusted by the Generalized Exponential-distributed term δ_i . The random effect δ_i distribution follows as *Generalized Exponential(a, b)*, where a and b are the distribution parameters.

3.4. Model estimation

The previous sections outlined the posterior estimation framework for all models, which was implemented using a Bayesian approach through Markov Chain Monte Carlo (MCMC) simulation in OpenBUGS. Given the bias and auto-correlation inherent in MCMC samples, this study ran numerous iterations to mitigate these effects. Multiple simulation chains were executed to ensure robust convergence; each initialized with different starting values. Specifically, three chains were run for 80,000 iterations per parameter, with the first 30,000 samples discarded as burn-in to eliminate potential bias. The remaining 50,000 iterations were used to estimate posterior distributions. Convergence was evaluated by visually inspecting trace plots, which show the progression of chains from various starting points, and by calculating the Brooks–Gelman–Rubin (BGR) statistic for each parameter (El-Basyouny and Sayed, 2009). A BGR value stabilized near 1 (below 1.1) indicated effective convergence (Gelman and Rubin, 1992; Mitra and Washington, 2007). Additionally, each parameter's Monte Carlo (MC) error was kept below 3% of the posterior standard deviation to ensure reliable parameter estimates.

The dependent variable considered pedestrian–vehicle crashes with all severity levels (KABCO) included. The study highlights that only explanatory variables with a 5% significance level were included in the results, indicating the use of a 95% CI to identify key crash-contributing factors. In Bayesian modeling, the random parameter framework introduces an additional layer of complexity by allowing parameters to vary across observations, capturing unobserved heterogeneity. However, when the variability in these parameters is minimal, their inclusion as random parameters can unnecessarily complicate the model without adding meaningful value. To balance complexity and performance, carefully determining which parameters benefit from this hierarchical structure is crucial.

3.5. Model performance

The Deviance Information Criterion (DIC) is widely used within the Bayesian framework for model selection, given the availability of numerous models. The principle of DIC is parsimony, aiming to find the simplest model that explains the most variation in the data. The DIC is calculated as:

$$DIC = \bar{D}(\Theta) + P_D \quad \left[\bar{D}(\Theta) = E[-2 \log L], \quad P_D = \bar{D}(\Theta) - D(\bar{\Theta}) \right] \quad (18)$$

where $\bar{D}(\Theta)$ is the posterior mean deviance indicating model fit, Θ is the total number of parameters, P_D is the adequate number of parameters, L is the likelihood. Generally, a model with a lower DIC is preferred. A difference greater than 10 in DIC values between models strongly favors the model with the lower DIC. Differences between 5 and 10 are considered substantial, while differences less than 5 suggest that the models are competitive.

The evaluation process relied on metrics, Mean Absolute Error (MAE), and Root Mean Square Error (RMSE) to assess the performance of the models on the test set and identify the best combination of hyperparameters for each model. By comparing the performance of all models, we can locate the superior results in capturing the relationship between pedestrian crashes and risk factors (Lord et al., 2021). The MAE is calculated as:

$$MAE = \frac{1}{n} \sum |y_{\text{observed}} - y_{\text{predicted}}| \quad (19)$$

where n represents the number of observations, and y_{observed} and $y_{\text{predicted}}$ correspond to the observed and predicted crash counts, respectively. The RMSE is calculated as:

$$RMSE = \sqrt{\frac{1}{n} \sum (y_{\text{observed}} - y_{\text{predicted}})^2} \quad (20)$$

These metrics are commonly employed to assess the accuracy of the model predictions by comparing the observed and predicted values of crash counts. The MAE provides the average magnitude of the errors in a model's predictions. At the same time, the RMSE gives greater weight to more significant errors, offering a more sensitive measure of model performance.

3.6. Marginal effects

Usually, the marginal effect (ME) is computed to explore a deeper understanding of the effects of explanatory variables (significant) on crashes. Which determines the unit change in expected crash frequency due to a unit changing explanatory variable X_j , while all other explanatory variables must be constant. For explanatory variables whose values range from 0 to 1 (Williams, 2012), calculating ME according to Washington et al. (2020):

$$ME_{X_{ij}} = \frac{\partial \lambda_i}{\partial X_{ij}} \quad (21)$$

Here, $ME_{X_{ij}}$ is defined as the marginal effect of the j th explanatory variable at the i th observation, λ_i denotes the expected mean crash frequency, and X_j is the j th explanatory variable of interest.

4. Data description

This section outlines the data foundation of the study. Section 4.1 describes the five years of pedestrian crash data obtained from TxDOT's Crash Records Information System (CRIS) database. Section 4.2 details the complementary data on bus stop design, roadway features, and traffic exposure. Together, these datasets enable a comprehensive and context-rich assessment of pedestrian safety near bus stops.

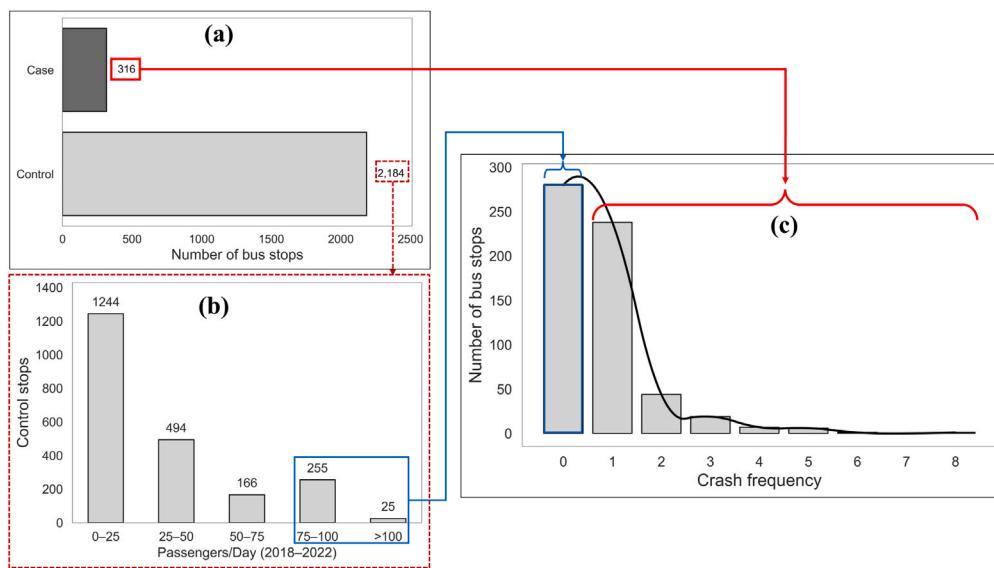


Fig. 2. Distribution of “Case” and “Control” stops: (a) stops classification, (b) daily ridership at “Control” stops, (c) frequency distribution among selected stops.

4.1. Pedestrian crash data

This study utilized bus stop data obtained from Trinity Metro, the public transit agency serving the Fort Worth metropolitan area. The data were accessed through the agency’s official website (<https://www.trinitymetro.org>) and included Geographic Information Systems (GIS) shapefiles containing spatial and attribute information for approximately 2500 bus stops. Each stop was identified by its corresponding main street and nearby cross-street and included details on amenities and wheelchair accessibility. In addition to the spatial data, Trinity Metro provided two supplementary spreadsheets reporting passenger boarding and alighting counts, disaggregated by stop location and service day type (weekday, Saturday, and Sunday). In the absence of direct pedestrian activity data, these counts served as a proxy for estimating relative crowding levels at each stop.

In alignment with the study’s objective, historical pedestrian crash data were systematically integrated with the stop dataset. The crash records for 2018 to 2022 were obtained from CRIS. This dataset included only reportable crashes, defined as incidents involving a fatality, injury, or property damage exceeding \$1000, and restricted to public roadways. To spatially associate pedestrian crashes with bus stop locations, a 250 ft buffer was established around each stop. This distance was informed by prior literature (e.g., [Craig et al. \(2019\)](#)), which noted that the majority of pedestrian crashes related to transit access occur within 150 ft of a stop. Building on this, and through a trial-and-error procedure aimed at generating reliable vehicle–pedestrian crash concentration maps, a 250 ft buffer was determined to provide an optimal balance. The threshold was sufficiently wide to capture pedestrian movements and crossing behaviors associated with bus stop access, yet narrow enough to reduce the likelihood of including unrelated incidents on adjacent roadways. To maintain spatial precision, crashes located within the buffer but occurring on cross-streets were manually reviewed and excluded from the analysis. Given the challenges in collecting detailed geometric and operational data for all bus stops, a targeted sampling strategy was adopted to enable a comprehensive yet manageable evaluation. Stops were selected using two clearly defined criteria to represent both high and low-risk environments:

1. Bus stops with at least one recorded pedestrian crash.
2. Bus stops exhibiting reasonable pedestrian activity (as indicated by boarding and alighting data) but with no recorded crashes.

The first criterion captures stops with documented pedestrian crash history during the study period, enabling a reactive analysis of contributing factors. To ensure inclusivity across the full spectrum of crash severities, 316 bus stops with one or more pedestrian crashes were designated as “case” stops. Among these, 75.3% had exactly one crash, 13.9% had two crashes, and a small number had three or more, with a maximum of eight at a single stop. This distribution, shown in Fig. 2(a & c), follows a long-tail pattern, most crash-involved stops experience low crash frequencies. Applying a stricter threshold (e.g., two or more crashes) would have excluded nearly 90% of these locations, greatly limiting sample diversity and statistical power. Including all stops with at least one crash thus ensures a more inclusive and representative crash risk analysis.

For the second criterion, a quantitative threshold was applied to define “reasonable pedestrian activity”. As illustrated in Fig. 2(b), most non-crash (control) stops had low boarding and alighting volumes: 1244 stops had 0–25 passengers/day, and 494 had 25–50 passengers/day. Including these would have introduced a major imbalance in exposure relative to the 316 case stops. Although higher-activity categories (50–75 and ≥100 passengers/day) contained fewer stops (n = 166 and 25, respectively), they still offered limited comparability. To mitigate this imbalance and adhere to resource constraints during data collection and fieldwork, the study limited control stops to those with ≥75 daily passengers. This threshold captured the upper end of the exposure distribution and provided a control sample reasonably aligned in size and exposure level with the case group. Although alternative thresholds (e.g., ≥50 or ≥100 passengers/day) were not tested in this study, future research could assess the sensitivity of results to these variations. The adopted threshold in this study reflects a practical balance between analytical rigor and feasibility, ensuring a matched, exposure-informed comparison between crash-prone and non-crash locations. By incorporating both types of stops, the analysis identifies overrepresented risk factors at hazardous sites and potential protective features at safer ones, strengthening the overall understanding of pedestrian safety dynamics near bus stops.

Fig. 3 presents a flowchart summarizing the data preparation process, including the integration of pedestrian crash records with bus stop and roadway information. Based on the sampling criteria described earlier, a total of 596 bus stops were selected for detailed analysis (Fig. 4). To support this analysis, pedestrian crash data were integrated with roadway geometric and traffic characteristics extracted from TxDOT’s 2019 Roadway Highway Inventory Network Offload (RHIno) database, a comprehensive GIS-based dataset.

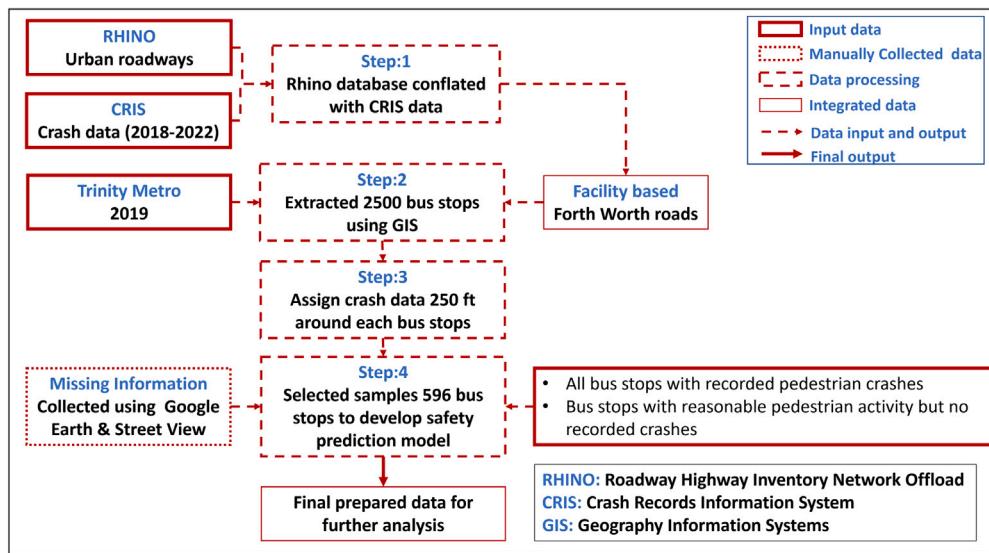


Fig. 3. Flowchart for data preparation.

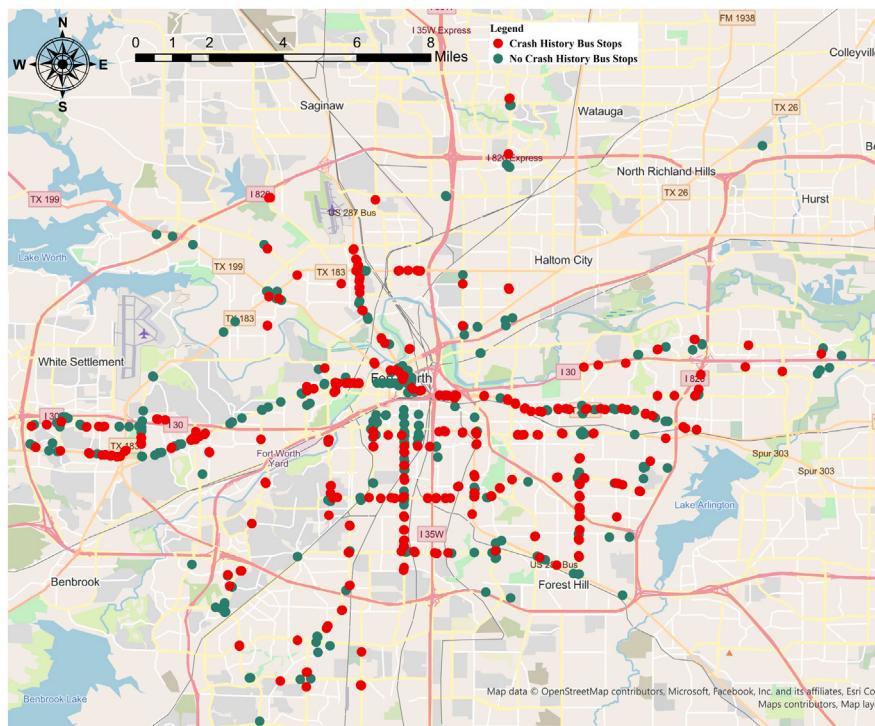


Fig. 4. Sample bus stop locations for data collection in Fort Worth, Texas.

The RHiNo database contains location-specific roadway information for both state and local roads, including system type, facility type, number of lanes, surface width, shoulder width, median type, and traffic volumes (current and historical). When key geometric features, such as the presence of the sidewalk, the availability of the crosswalk, or the lighting conditions, were missing, additional data were collected through Google Earth® and Google Street View®, which served as effective tools for field validation and remote assessment.

The final dataset included 448 pedestrian crashes, categorized by severity. Among these, 46 (10.3%) were fatal (K-level), 394 (87%) involved injuries (A, B, or C-level), and 8 crashes (1.8%) were classified as property damage only (PDO). The yearly distribution of crash severity is shown in Fig. 5, while Fig. 6 illustrates general crash characteristics,

including speed limit, day of the week, lighting, severity distribution, and weather conditions.

These underscore the substantial risk to pedestrians near bus stops—with nearly 98% of crashes resulting in injuries or fatalities, the data highlight an urgent need for targeted safety improvements. The low proportion of PDO crashes further emphasizes the vulnerability of pedestrians in traffic environments, reinforcing the critical importance of interventions aimed at reducing both fatal and injury-related incidents.

4.2. Bus stops environment

Among the 596 bus stops analyzed, 280 (46.97%) had no recorded pedestrian crashes during the study period. These non-crash locations

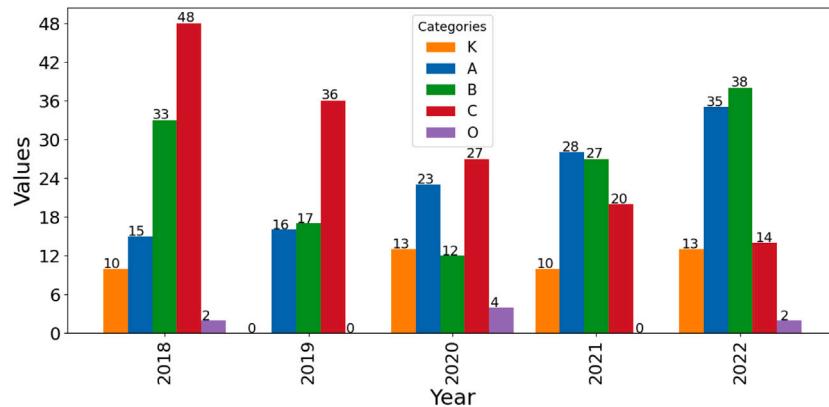


Fig. 5. Yearly different severity levels of pedestrian crashes.

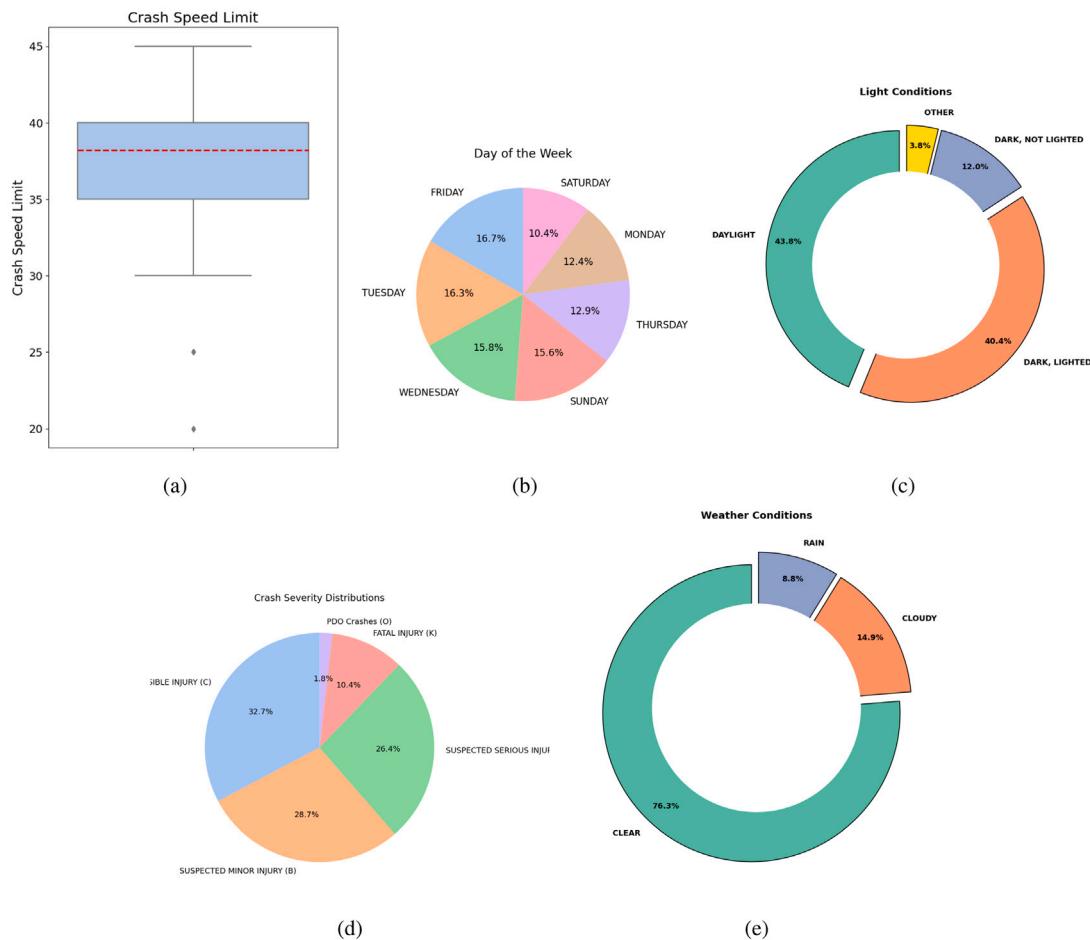


Fig. 6. Crash distribution characteristics.

were selected based on the inclusion criteria described in Section 4.1. As illustrated in Fig. 2(c), the distribution of crash frequency is heavily right-skewed, with a skewness value of 2.87, confirming the dominance of non-crash stops in the dataset.

This study distinguishes itself by integrating a diverse set of explanatory variables that capture spatial, geometric, and contextual dimensions often omitted in traditional modeling approaches. Table 1 presents descriptive statistics for the response variables and associated predictors. Data were compiled from multiple sources, including CRIS, RHINO, Trinity Metro, Google Earth®, and Google Street View®, to ensure a comprehensive understanding of bus stop environments. The

explanatory variables are organized into four overarching categories: exposure-related features, bus stop design characteristics, roadway attributes, and contextual elements. Exposure-related features reflect traffic and pedestrian activity near bus stops, incorporating metrics such as Average Annual Daily Traffic (AADT), boarding, and alighting counts. Bus stop design characteristics describe the physical layout and configuration of the stops, including design type and proximity to intersections. Roadway attributes capture infrastructure elements such as speed limits, lane counts, median type, and intersection distance. Contextual elements, such as land use type, lighting conditions, and nearby schools

or parks, provide additional insight into the operating environment and pedestrian safety implications.

To evaluate the influence of crash severity, the crash data were stratified into three groups. The first group includes all pedestrian-involved crashes (KABCO). The second group (KABC) narrows the focus to fatal and injury crashes, excluding property damage only. The third group (KAB) concentrates on the most severe incidents, including fatalities and incapacitating or non-incapacitating injuries. This stratification enables a more nuanced investigation of the factors contributing to crash severity.

As shown in [Table 1](#), the mean number of crashes per bus stop is 0.752, with a maximum of 8 crashes at a single location. The mean KABC crash frequency is 0.526 (maximum = 6), while the KAB category shows a mean of 0.314 (maximum = 5). These values demonstrate considerable variation in crash occurrences across the study area, underscoring the importance of site-specific analysis.

In addition to crash data, several continuous variables were examined. AADT values ranged from 166 to 42056 vehicles per day, with a mean of 13558.2, reflecting the variation in traffic exposure near bus stops. Although direct pedestrian counts were unavailable, boarding and alighting volumes served as reasonable proxies for pedestrian activity. The distance to the nearest intersection ranged from 0.4 to 2106 ft, enabling analysis of spatial relationships between crashes and network geometry. Other continuous variables include median width, posted speed limit, number of lanes, number of nearby schools and parks, and the count of adjacent transit stops within a 0.25-mile radius. Several categorical variables were also included to account for environmental and design-related influences. These include intersection type (signalized vs. non-signalized), presence of crosswalks and sidewalks, roadway curvature, lighting conditions, area type (commercial, residential, or mixed), bus stop design (curbside vs. other), proximity of the stop to intersections (near, far, or midblock), and shelter availability. These features may collectively influence pedestrian behavior, visibility, and driver response, thus affecting crash risk. [Table 1](#) summarizes these variables, illustrating the diverse conditions under which pedestrian crashes occur and providing a foundation for subsequent modeling and risk analysis.

5. Results and discussion

This section is divided into four subsections. First, it presents the model performance evaluation based on the KABCO crash severity scale. Second, it examines the key factors associated with pedestrian crashes near bus stops. Third, it explores the specific effects of these factors on severe crashes (KAB). Finally, the fourth subsection outlines the systematic identification and prioritization of high-risk bus stops using the proposed framework.

5.1. Model performance evaluation

Although predictive accuracy was not the primary objective of this study, a preliminary assessment was conducted to evaluate the comparative performance of the models. [Table 2](#) summarizes these evaluation results, highlighting the RPNB-L model as superior, evidenced by its lowest DIC value. A difference of 10 or more in DIC values is generally regarded as significant ([Lord et al., 2021](#)); in this study, the RPNB-L model demonstrated a DIC value approximately 3.27% lower than the NB-L model and 3.88% lower than the RPNB-GE model, signifying a meaningful improvement in model performance.

Additionally, the RPNB-L model exhibited the lowest posterior mean deviance (Dbar), reflecting a better overall fit to the observed crash data compared to the competing models. Despite its higher complexity, indicated by the adequate number of parameters (PD), the RPNB-L model achieved a better balance between complexity and fit, resulting in superior predictive performance without substantial penalties. This balance underscores the model's effectiveness in capturing unobserved heterogeneity and site-specific variability.

Regarding prediction accuracy, the RPNB-GE model displayed the highest MAE, signifying relatively less accurate predictions. Conversely, the RPNB-L model produced the lowest MAE for both training and testing datasets, as well as the lowest RMSE. Specifically, the RPNB-L model's MAE and RMSE were approximately 3.98% and 7.93% lower, respectively, than the NB-L model with fixed coefficients. This superior predictive performance of the RPNB-L model can be attributed to its flexible structure, which accommodates mixed distributions and random parameters, thereby effectively capturing variations across bus stop locations. Consequently, the RPNB-L model emerges as the most robust and reliable analytical tool, offering enhanced accuracy and adaptability in modeling pedestrian crashes near bus stops.

Given the superior performance of the RPNB-L model across key metrics, an additional validation using the Cumulative Residual (CURE) plot was conducted to assess potential biases in model predictions. The CURE plot provides valuable insights into prediction accuracy by visually examining the accumulation of residuals across critical explanatory variables ([Fig. 7](#)). Specifically, the cumulative residual behavior was evaluated for primary exposure variables, including Ln(AADT) and AvgOn. A desirable characteristic of the CURE plot is that cumulative residuals should consistently oscillate near the zero line, indicating unbiased predictions. However, in this analysis, the CURE plots exhibit prolonged smooth runs instead of frequent oscillations around the zero reference line. Rather than reflecting model overfitting, this pattern arises from the skewed distributions and uneven densities within the ranges of the exposure variables. Consequently, residuals accumulate smoothly over certain ranges of data points rather than fluctuating sharply. To further validate the assessment, 95% were computed by estimating residual variances over ascending values of explanatory variables and subsequently accumulating these variances. The cumulative residual lines for the RPNB-L model consistently remained close to the zero line, falling well within these confidence intervals. Overall, the CURE plot evaluation supports the robustness and predictive accuracy of the RPNB-L model, reinforcing its effectiveness in reliably modeling pedestrian crash risks near bus stops.

5.2. Key contributing risk variables

Given the superior predictive performance demonstrated in [Section 5.1](#), the RPNB-L model was selected for a more detailed interpretation and discussion. As previously described, the RPNB-L model incorporates random parameters within its hierarchical structure (see Eq. (12)), allowing the effects of explanatory variables to vary across bus stop locations. However, during the analysis, it became apparent that certain parameters exhibited minimal variability, contributing only negligible improvements to GOF and parameter mean estimates outside of 95% CI. To balance model complexity with interpretability, random parameters were therefore assigned only to the primary exposure variables (e.g., Ln(AADT) and AvgOn) while the remaining parameters were treated as fixed ([Hasan and Abdel-Aty, 2024](#)). This approach simplified the model without sacrificing significant predictive power.

As summarized in [Table 2](#), all parameter estimates remained statistically significant at the 95% confidence interval, including the standard deviations for the randomly specified exposure parameters. To provide practical insights into the effects of these variables, average marginal effects were calculated, quantifying how variations in each explanatory variable influence crash frequency. The procedure for calculating marginal effects is detailed in [Section 3.6](#), and the corresponding results are presented in [Table 3](#).

5.2.1. Annual average daily traffic (AADT)

AADT near bus stops significantly influences pedestrian crash frequency. The RPNB-L model confirms that Ln(AADT) is statistically significant at the 95% CI, indicating a positive relationship between traffic volume and crash frequency. Specifically, the estimated mean coefficient for Ln(AADT) is 0.345, suggesting that higher traffic volumes

Table 1
Descriptive statistics of explanatory variables.

Variable	Definition	Descriptive statistics			
		Min	Max	Mean	Std Dev
Total crashes	KABCO frequency	0	8	0.752	1.45
KABC crashes	KABC frequency	0	6	0.526	1.11
KAB crashes	KAB frequency	0	5	0.314	0.65
Continuous variables					
AADT	Traffic Volume on the street near bus stop	166	42056	13 540.6	8827.85
Avg on	Average boarding	0	769.0	53.20	66.78
Avg off	Average alighting	0	831.0	56.63	60.64
Dis to int	Distance between bus stop and nearest intersection corner	0.4	2106	198.94	241.04
Med w	Median width	0.00	131.9	7.10	12.38
Speed limit	Posted speed limit	20	65	36.63	7.78
Lane count	Total number of lanes	1	8	4.34	1.46
School	Number of schools within a half-mile walk of the bus stop	0	6	0.82	1.12
Park count	Number of parks within a half-mile walk of the bus stop	0	6	0.64	0.84
Stop count	Count of transit stops (bus or rail) within a quarter mile of the bus stop	0	11	3.45	2.19
Categorical variables		Category	No of stops	Crash frequency	
Int t	Nearest intersection type	Signalized (1) Non-signalized (0)	382 214	237 206	
Marked Xwalk	Presence of marked crosswalk within 250ft of bus stop	No (1) Yes (0)	234 362	176 267	
Med t	Median types	Undivided (1) Divided (0)	311 297	267 176	
Lighting	Lighting presence	No (1) Yes (0)	339 257	273 170	
Area	Area type	Com (0) Res (1) Mix (2)	307 118 171	216 116 111	
Sidewalks	Presence of sidewalks	Yes (0) No (1)	550 46	398 45	
Curve	Presence of curve near bus stop	Yes (1) No (0)	38 558	18 425	
Design	Bus stop design	Curbside (1) Other (0)	507 89	416 27	
Proximity	The proximity of intersection from bus stop location	Far (1) Near (0) Midblock (2)	301 226 69	207 236 0	
Cover	Presence of shelter	Covered (0) Uncovered (1)	302 294	236 207	

correspond to a greater likelihood of pedestrian crashes. This finding aligns with the expectation that increased vehicular flow elevates pedestrian exposure to collisions. Consistent with previous studies, these results highlight the risks associated with high traffic volumes near bus stops, which are often located in high-density passenger activity areas. Such areas experience significant traffic movement, increasing pedestrian exposure and crash risk (Tiboni and Rossetti, 2013). The absence of dedicated bus lanes exacerbates this issue, as buses operating in mixed traffic negatively impact pedestrian safety (Yendra et al., 2024). Additionally, studies have shown that higher AADT levels correlate with increased pedestrian fatalities in these environments (Ulak et al., 2021; Lakhotia et al., 2020; Hess et al., 2004; Amadori and Bonino, 2012). The model's random parameterization for $\ln(\text{AADT})$, with a standard deviation of 0.042, highlights the variability in traffic volume impact across bus stop locations. This suggests that while traffic volume generally increases crash frequency, the magnitude of this effect varies between sites. Some bus stops may experience a steeper rise in crash risk with growing traffic volumes, while others may be less affected. This variability underscores the importance of considering site-specific factors in evaluating vehicular traffic's impact on pedestrian safety. Furthermore, the marginal effects of the RPNB-L model indicate that a one-unit increase in the logarithm of AADT corresponds to approximately a 0.336-unit increase in pedestrian–vehicle crash frequency near bus stops.

5.2.2. Passengers density

In this study, pedestrian density near bus stops was represented by the mean number of passengers boarding and alighting at each stop, given the challenges of capturing real-time density. The model

results indicate a positive association between higher pedestrian exposure, reflected in increased boarding activity, and crash frequency. Greater pedestrian volumes elevate the likelihood of vehicle–pedestrian interactions, increasing crash risk. The estimated mean coefficient for the “Avg on” variable is 0.010, with a standard deviation of 0.008, highlighting variability in effect across different bus stops due to factors such as traffic conditions and surrounding infrastructure. These findings are consistent with prior research (Hess et al., 2004; Geedipally, 2021; Craig et al., 2019), which links high pedestrian volumes near bus stops to elevated crash risks, particularly in dense areas like schools and commercial zones where large groups of pedestrians gather. The marginal effect indicates that an increase in average boarding activity raises passenger–vehicle crash frequencies by approximately 0.018 near bus stops. Both the model and previous studies emphasize that higher pedestrian density increases exposure to traffic, thereby raising crash risks.

5.2.3. Speed limits

The results reveal a paradox: lower speed limits (35 mph or less) near bus stops are associated with increased pedestrian crashes. One potential cause is that pedestrians may feel safer with lower speed limits posted, which leads them to cross the road more frequently and increases their exposure to collisions. In addition, some drivers may exceed speed limits, resulting in unpredictable driving behaviors that further increase crash risks. Previous research has primarily focused on the dangers of high traffic speeds near bus stops, consistently linking higher speeds to an increased likelihood of pedestrian deaths (Rossetti et al., 2020; Pulugurtha and Vanapalli, 2008; Hess et al., 2004; Lakhotia

Table 2

Posterior model estimated parameters for pedestrian–vehicle (KABCO) crashes near bus stops.

Variables	NB-L			RPNB-L			RPNB-GE		
	Mean (Std. Dev.)	95% CI		Mean (Std. Dev.)	95% CI		Mean (Std. Dev.)	95% CI	
		LL	UL		LL	UL		LL	UL
Mean of parameters									
Constant	-0.776(0.076)	-0.879	-0.578	-0.488(0.031)	-0.668	-0.278	-0.812 (0.08)	-0.898	-0.545
Exposure variable									
Ln (AADT)	0.349(0.080)	0.095	0.411	0.345(0.083)	0.091	0.424	0.351(0.081)	0.096	0.040
Avg on	0.007(0.003)	0.002	0.013	0.010(0.004)	0.001	0.014	0.012(0.008)	0.007	0.019
Roadway environment									
Speed limit (1 if speed limit greater or equal to 35 mph, 0 otherwise)	0.457(0.163)	0.139	0.778	0.448(0.162)	0.132	0.760	0.461(0.16)	0.141	0.791
Med t (1 if nearest median is Undivided, 0 otherwise)	0.321(0.149)	0.027	0.621	0.326(0.148)	0.0425	0.624	0.323(0.150)	0.028	0.631
Int t (1 if nearest intersection is signalized, 0 otherwise)	-0.330(0.098)	-0.622	-0.004	-0.329(0.108)	-0.675	-0.017	-0.334(0.010)	-0.624	-0.010
Bus stop design features									
Sidewalk (1 if sidewalk is present, 0 otherwise)				-0.379(0.235)	-0.843	-0.011	-0.446(0.312)	-0.976	-0.234
Marked Xwalk (1 if no presence of marked xwalk, 0 otherwise)				0.098 (0.075)	0.002	0.446			
Lighting (1 if no presence of lighting, 0 otherwise)				0.100(0.014)	0.060	0.184			
Area Mix (1 if bus stop located in mixed environment area, 0 otherwise)	-0.355(0.161)	-0.672	-0.037	-0.367(0.159)	-0.642	-0.029	-0.378(0.166)	-0.688	-0.039
School count (1 if the Number of schools within a half-mile walk of the bus stop is greater than 1, 0 otherwise)	0.562(0.181)	0.272	0.867	0.556(0.179)	0.200	0.908			
Proximity (1 if bus stop located far side of intersection, 0 otherwise)	-0.253(0.142)	-0.483	-0.023	-0.251(0.141)	-0.528	-0.027	-0.252(0.141)	-0.471	-0.024
Standard deviation of random parameters									
Ln (AADT)				0.042(0.005)	0.044	0.072	0.049(0.006)	0.046	0.089
Avg On				0.008(0.001)	0.003	0.011	0.009(0.001)	0.003	0.012
Dispersion parameter (α)	0.365(0.09)	0.124	0.544	0.137(0.083)	0.067	0.245	0.378(0.010)	0.112	0.576
Lindley parameter (θ)	1.414(0.166)	1.233	1.766	1.378(0.122)	0.899	1.589			
GE parameter (a)							2.022(0.778)	1.089	2.532
GE parameter (b)							1.442(0.378)	1.102	1.988
Performance measure									
DIC	1233.56			1193.23			1241.4		
Dbar	1114.06			992.13			1096.6		
Pd	119.5			201.1			144.8		
MAE (Train)	0.714			0.686			0.724		
MAE (Test)	0.805			0.773			0.812		
RMSE (Train)	0.956			0.881			0.968		
RMSE (Test)	1.072			0.987			1.104		

et al., 2020). As a result, lower speed limits have often been recommended to improve pedestrian safety (Ulak et al., 2021; Tubis et al., 2021; Pessaro et al., 2017). However, the findings of this study suggest that simply posting reduced speed limits may not adequately mitigate crash risks. To effectively improve safety, stricter speed enforcement and enhanced pedestrian infrastructure, such as crosswalks, pedestrian signals, and traffic calming measures, are essential. These measures can help ensure driver compliance with speed limits while providing pedestrians with safe crossing options and reducing vehicle–pedestrian conflicts near bus stops.

5.2.4. Median types

The findings show that roadways near bus stops without medians are significantly more likely to experience pedestrian crashes than those with medians. Medians serve as pedestrian refuges, allowing individuals to crossroads in stages and reducing their exposure to traffic. In contrast, undivided roadways require pedestrians to cross the entire width of the road at once, increasing crash risks due to unpredictable crossing behavior and limited driver visibility. These results align with previous studies highlighting the safety benefits of medians in reducing pedestrian–vehicle conflicts by providing protected spaces for pedestrians to pause mid-crossing (Yu, 2024; Salum

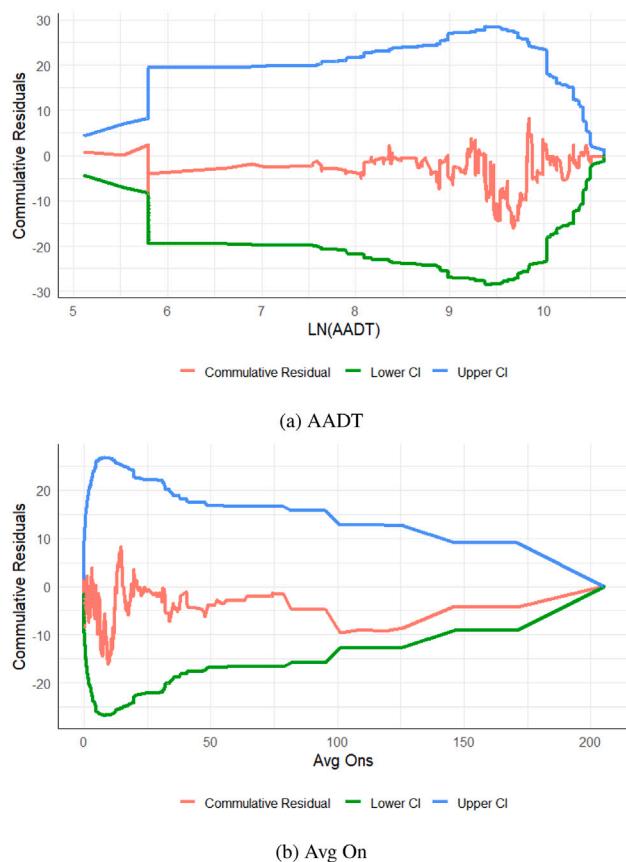


Fig. 7. CURE plots for RPNB-L model.

Table 3

RPNB-L based average marginal effects of crash contributing parameters.

Risk variables	Effects (Mean)
Ln (AADT)	0.336
Avg on	0.018
Speed limits (1 if speed limits greater or equal to 35 mph, 0 otherwise)	0.259
Med t (1 if nearest median is Undivided, 0 otherwise)	0.245
Int t (1 if nearest intersection is signalized, 0 otherwise)	-0.247
Sidewalk (1 if sidewalk is present, 0 otherwise)	-0.284
Marked Xwalk (1 if no presence of marked xwalk, 0 otherwise)	0.074
Lighting (1 if no presence of lighting , 0 otherwise)	0.075
Area Mix (1 if bus stop is located in mixed environment area, 0 otherwise)	-0.275
Proximity (1 if bus stop is located far side of the intersection, 0 otherwise)	-0.188
School count (1 if the Number of schools within a half-mile walk of the bus stop is greater than 1, 0 otherwise)	0.017

et al., 2024; Jeng et al., 2003). Medians offer an essential safety buffer, particularly on wide or high-traffic roads, by creating more predictable crossing points and minimizing pedestrian exposure to moving vehicles. The findings reinforce the importance of incorporating medians as a critical element in road design near bus stops to enhance pedestrian safety. Strategically placed medians ensure safer pedestrian movement and reduce potential conflicts with vehicles, ultimately contributing to lower crash risks (Pessaro et al., 2017).

5.2.5. Nearest intersection

Intersections are widely recognized as critical zones for pedestrian risk due to the convergence of multimodal traffic streams and complex crossing behaviors. Numerous studies (Harwood et al., 2008; Kuşkapan et al., 2022; Walgren, 1998; Geedipally, 2021; Quistberg et al., 2015b) have reported elevated pedestrian crash frequencies at or near bus stops located within 1000 ft of signalized intersections. However, the scope of these investigations has primarily focused on signalized intersections, often overlooking unsignalized intersections, which may present heightened risk due to the lack of formal traffic control mechanisms. To address this limitation, the present study examined the relationship between pedestrian crash risk and intersection type, specifically comparing signalized and unsignalized intersections. The RPNB-L model results (Table 2) indicate a negative and statistically significant association between bus stops located near signalized intersections and pedestrian crashes (-0.329). This suggests that relative to unsignalized locations, signalized intersections may offer a protective effect by reducing crash likelihood. These findings align with those of Samani and Amador-Jimenez (2023), who observed that even modest infrastructure elements such as pedestrian signals and stop signs at non-signalized intersections can enhance driver awareness and pedestrian visibility. Signalized intersections likely reduce pedestrian risk by facilitating more orderly vehicle movement, minimizing turning conflicts, and providing designated pedestrian crossing phases.

To further validate these findings and explore potential confounding effects, a stratified Mantel-Haenszel (MH) analysis (Santner and Duffy, 2012) was conducted. This method provides an adjusted estimate of the association between crash occurrence and intersection type by accounting for key contextual variables across multiple strata. These strata were defined based on traffic volume (AADT), bus stop proximity to intersections, presence of marked crosswalks, and land use type. The MH estimator was computed using stratum-specific 2×2 contingency tables, where each table cross-tabulated pedestrian crash occurrence (yes/no) with intersection type (signalized/unsignalized). The pooled Mantel-Haenszel odds ratio (MHOR) is given by:

$$\hat{OR}_{MH} = \frac{\sum_{s=1}^S \frac{a_s d_s}{n_s}}{\sum_{s=1}^S \frac{b_s c_s}{n_s}} \quad (22)$$

where a_s , b_s , c_s , and d_s represent the cell counts of the s th stratum, and n_s is the total number of observations within that stratum.

Results from the MH analysis corroborated the model-based findings. Signalized intersections were consistently associated with significantly lower crash risk across all stratified analyses. For example, when stratified by AADT with different ranges, the MHOR was 0.284, corresponding to a relative risk reduction of approximately 41% (RR \approx 0.59). Similarly, stratification by bus stop proximity on the far side yielded an MHOR of 0.313 (RR \approx 0.62); by marked crosswalk presence, 0.272 (RR \approx 0.56). All odds ratios were statistically significant at the $p < 0.001$ level. These consistent results across multiple strata highlight the presence of traffic control mechanisms, such as signal phases, pedestrian indications, and regulated turning movements, may contribute to reducing pedestrian exposure to conflicts and enhancing overall safety. Moreover, these findings help reconcile discrepancies in previous literature by illustrating how aggregated or unstratified analyses may obscure the protective influence of signalization.

5.2.6. Sidewalks

The model results from Tables 3 and 4 show that the absence of sidewalks near bus stops significantly increases pedestrian crash likelihood, as indicated by the negative coefficient (-0.379). This finding suggests that without sidewalks, pedestrians are more exposed to traffic, often forced to walk on the road, thereby increasing the risk of vehicle-pedestrian collisions. Sidewalks provide a designated pedestrian space, minimizing traffic interactions and reducing crash risks. These results align with previous research emphasizing the critical role of sidewalks

in enhancing pedestrian safety near bus stops. Pessaro et al. (2017) highlighted that proper pedestrian infrastructure, such as sidewalks, encourages safer walking behavior. Similarly, Sukor and Fisal (2020) found that safety is the strongest factor influencing walkability to bus stops, further reinforcing the importance of adequate sidewalks. Tiboni and Rossetti (2013) found that one-third of the bus stops they surveyed lacked sidewalks, which exposed pedestrians to greater traffic risks. Likewise, Rossetti et al. (2020) emphasized that pedestrians who walk along roadways without sidewalks face an increased risk of crashes. Overall, the findings highlight the critical importance of sidewalks in reducing pedestrian–vehicle interactions and ensuring safer access to bus stops (Yu, 2024; Samani and Amador-Jimenez, 2023; Lakhota et al., 2020; Mukherjee et al., 2023), making sidewalks a vital component of pedestrian infrastructure.

5.2.7. Marked crosswalks

The model results indicate (Table 4) that the absence of marked crosswalks near bus stops is associated with a higher risk of pedestrian crashes, as shown by approximately a 0.074 unit increase in pedestrian–vehicle crash frequency near bus stops. This finding suggests that locations without marked crosswalks are more likely to experience pedestrian crashes. Marked crosswalks provide a designated space for pedestrians to cross safely, reducing exposure to traffic and improving the predictability of pedestrian movements for drivers. Without crosswalks, pedestrians are more likely to cross at unpredictable locations, increasing crash risk due to reduced driver awareness and slower response times. Crosswalks also help organize traffic flow by signaling drivers to anticipate pedestrian activity, leading to safer road interactions. These findings align with previous research emphasizing the importance of marked pedestrian crossings near bus stops. For instance, Tiboni and Rossetti (2013) found that over 30% of inspected bus stops lacked marked crossings, contributing to unsafe pedestrian conditions. Similarly, Mukherjee et al. (2023) demonstrated that the presence of zebra markings at pedestrian crossings can reduce the likelihood of fatal crashes by 23%, particularly during nighttime, by enhancing driver awareness of potential pedestrian interactions. The literature also highlights the value of pedestrian traffic signals in conjunction with crosswalks for improving safety. Studies show that pedestrian signals help regulate both driver and pedestrian behavior, encouraging drivers to expect pedestrians in crosswalks and prompting pedestrians to use designated areas safely (Jeng et al., 2003; Samani and Amador-Jimenez, 2023). Without such infrastructure, pedestrians are more likely to take risks, particularly on wider roadways or bus stops lacking proper crossing facilities (Lakhota et al., 2020).

5.2.8. Lighting

The inadequate lighting near bus stops significantly increases the likelihood of pedestrian crashes, as reflected by the positive coefficient of 0.100 in the RPNB–L model (Table 3). Poor lighting or no light at bus stops poses a higher risk during nighttime hours since its reduced visibility makes it difficult for pedestrians and drivers to anticipate movements. These findings align with previous research highlighting the critical role of adequate lighting in enhancing safety at bus stops. Mukherjee et al. (2023) identified a direct relationship between insufficient lighting and higher pedestrian fatalities, especially in sparse land use. Similarly, Tiboni and Rossetti (2013) found that 8% of the surveyed bus stops lacked artificial lighting, contributing to unsafe conditions. Studies consistently recommend enhanced illumination at bus stops and along their approaches to reduce crash risks (Pessaro et al., 2017; Salum et al., 2024). Adequate lighting ensures that drivers and pedestrians have clear visibility, enabling them to anticipate movements and react appropriately, lowering crash frequency. Overall, the model results and supporting literature underscore that improving lighting around bus stops is a critical safety measure for mitigating pedestrian crash risks.

5.2.9. Area use

Previous research has identified bus stops in high pedestrian activity areas, mainly commercial zones with heavy traffic, as high-risk locations for pedestrian crashes. These elevated risks are often attributed to poor integration with surrounding land use and proximity to parking areas, which create additional conflict points between pedestrians and vehicles (Pessaro et al., 2017; Hess et al., 2004). The findings of this study also align with previous studies and suggest that bus stops located in mixed-use areas, combined with residential, commercial, and recreational spaces, exhibit a lower likelihood of pedestrian crashes than those in strictly commercial or residential zones. This reduction in crash risk may be attributed to balanced traffic flows and the presence of improved infrastructure, such as sidewalks, crosswalks, proper lighting, and pedestrian walking warning signs. In addition, enhanced safety measures like speed bumps and pedestrian signals in mixed-use areas further contribute to safer environments. Drivers in these zones may also exercise greater caution, given the diverse types of road users, including residents and recreational pedestrians.

5.2.10. Bus stops proximity

The proximity of bus stops to intersections plays a critical role in pedestrian crash risks. Bus stops are categorized into three types based on their location: far-side, near-side, and mid-block. Far-side stops are located after intersections, near-side stops are positioned before intersections, and mid-block stops are between intersections. The model results indicate that far-side bus stops have a significantly lower likelihood of pedestrian crashes than near-side and mid-block stops, aligning with previous research (Fitzpatrick and Nowlin, 1997; Yu, 2024). Far-side stops are considered safer because they reduce pedestrian–vehicle conflicts by allowing pedestrians to cross with greater visibility and ensuring that they cross behind the bus or at the nearest marked crosswalk. In contrast, near-side stops are associated with higher crash risks, likely due to reduced driver visibility and the tendency for pedestrians to cross in front of the bus, where drivers may not anticipate their movements. Similarly, mid-block stops pose increased risks because they often lack proper pedestrian infrastructure, such as marked crosswalks, leading to unpredictable crossing behavior (Truong and Somenahalli, 2011; Pessaro et al., 2017; Rossetti et al., 2020).

5.2.11. School count

The study demonstrates that the presence of schools near bus stops significantly increases the likelihood of pedestrian crashes. The results indicate that higher pedestrian density around schools, especially during peak times, leads to more pedestrian–vehicle interactions, thereby elevating the risk of crashes. While many students may be dropped off directly at school by buses, the presence of schools increases pedestrian activity around nearby bus stops, which raises the potential for conflicts between pedestrians and vehicles. These findings align with previous research, such as Craig et al. (2019) and Geedipally (2021) found that bus stops near schools are associated with a higher incidence of collisions. Similarly, Hess et al. (2004) noted that gathering students around bus stops often leads to unpredictable pedestrian behavior, further heightening the risk of crashes.

5.3. KAB crashes

The estimation using the RPNB–L model for KAB pedestrian crashes provides several key insights when compared to the total pedestrian crash (KABCO) model, as presented in Table 4. First, most significant risk factors identified in the KABCO model remain consistent in the KAB model. However, an exception was observed with marked crosswalks and sidewalks, which did not achieve statistical significance at the 95% confidence level in the KAB model. This suggests that these features may have a greater influence on less severe crashes, primarily reflected in the KABCO and KABC models.

Table 4

Posterior model estimated parameters for pedestrian–vehicle (KAB) crashes near bus stops.

Variables	KAB Crashes		
	Mean (Std. Dev.)	95% CI	
		LL	UL
Mean of parameters			
Constant	-0.453 (0.027)	-0.643	-0.227
Exposure variable			
Ln (AADT)	0.191 (0.063)	0.025	0.441
Avg on	0.007 (0.001)	0.001	0.009
Roadway environment			
Speed limit (1 if speed limit \geq 35 mph, 0 otherwise)	0.393 (0.108)	0.087	0.661
Median type (1 if nearest median is undivided, 0 otherwise)	0.217 (0.089)	0.035	0.446
Intersection type (1 if nearest intersection is signalized, 0 otherwise)	-0.292 (0.098)	-0.372	-0.045
Bus stop design feature			
Lighting (1 if no presence of lighting, 0 otherwise)	0.231 (0.089)	0.054	0.412
Area Mix (1 if bus stop located in mixed environment, 0 otherwise)	-0.415 (0.186)	-0.917	-0.029
School count (1 if number of schools within 0.5 miles \geq 1, 0 otherwise)	0.641 (0.187)	0.210	1.108
Proximity (1 if bus stop located far side of intersection, 0 otherwise)	-0.312 (0.167)	-0.656	-0.0617
Dispersion parameter (α)	0.214 (0.118)	0.119	0.692
Lindley parameter (θ)	1.419 (0.168)	1.244	1.778
Standard deviation of random parameters			
Ln (AADT)	0.313 (0.154)	0.078	0.546
Avg on	0.043 (0.022)	0.010	0.132
Performance measure			
DIC	869.56		
Dbar	765.06		
Pd	131.5		
MAE (Train)	0.449		
MAE (Test)	0.505		
RMSE (Train)	0.577		
RMSE (Test)	0.676		

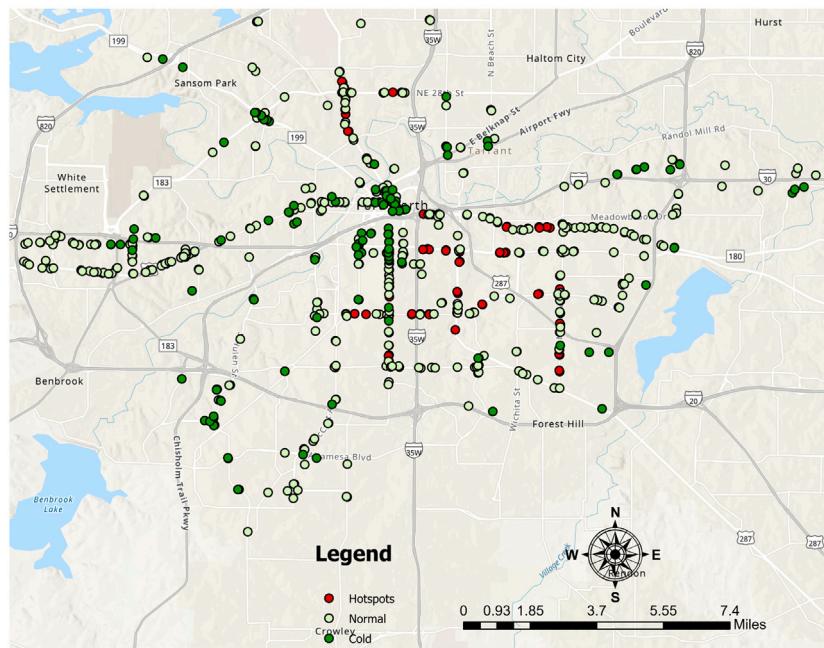
Second, similar to the KABCO model, traffic volume and average passenger boarding were found to be randomly distributed across locations in the KAB model. The posterior mean value for traffic volume (Ln AADT) was 0.191, with a standard deviation of 0.313, indicating its moderate effect on KAB crashes. Additionally, the coefficient for average passenger boarding (Avg on) was 0.007, with lower variability (standard deviation: 0.043). This suggests that while passenger boarding contributes to crash frequency, its impact on severe crashes is less pronounced. The coefficient for average passenger boarding further indicates that the relationship between traffic volume, speed limits, and pedestrian crash severity differs between severe (KAB) and less severe crashes. Specifically, lower traffic volumes on higher-speed roads tend to result in more severe crashes. The finding that higher speed limits (less or 35 mph) significantly increase crash severity reinforces this pattern, highlighting the critical role of speed in exacerbating the severity of pedestrian crashes near bus stops. The posterior parameter estimates further reveal that bus stops near signalized intersections experience fewer KAB crashes than those near non-signalized intersections. This underscores the importance of signalized intersections in reducing pedestrian crash risks, particularly for more severe incidents. Additionally, the absence of lighting at bus stops and proximity to schools were found to significantly increase crash severity. In contrast, far-side bus stops were associated with fewer severe crashes. These findings highlight the need for targeted safety interventions to address high-risk conditions and enhance pedestrian safety near bus stops.

5.4. Hazard bus stops/corridors identification

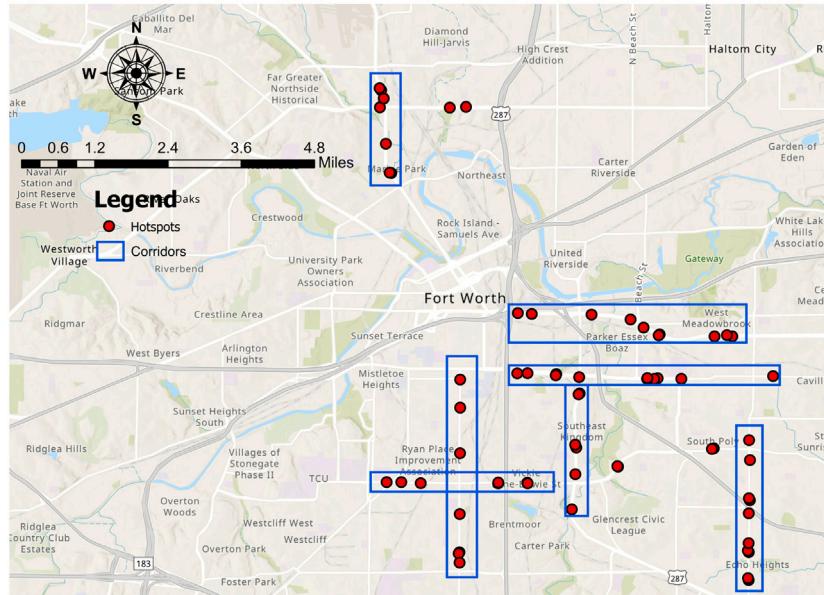
Identification of hazardous bus stops, defined as locations with a higher expected number of crashes than similar stops (Elvik, 2007), is crucial for authorities (Hauer and Persaud, 1984; Persaud et al., 1999; Miranda-Moreno et al., 2007; Montella, 2010). These high-risk locations serve as focal points for prioritizing safety investments and help delineate hazard corridors. Hotspot analysis, a widely used approach for identifying these sites, ranks locations from highest to lowest risk (Hauer et al., 2002; Hauer, 2004). However, this process

presents challenges due to random variability in crash occurrences, annual fluctuations, and the regression-to-the-mean (RTM) bias (Hauer and Persaud, 1984). The RTM bias reflects the statistical tendency for extreme crash frequencies at a site to regress toward the long-term average over time due to natural variation. Failure to account for RTM bias can lead to misclassifying sites as hazardous based on temporary crash anomalies or overlooking genuinely high-risk locations during periods of unusually low crash activity, leading to inefficient safety resource allocation. To mitigate these challenges, expected crash frequencies are commonly used, as they adjust for RTM bias and provide a more reliable measure of risk (Hauer, 1997). These frequencies can be estimated using the Full Bayes (FB) and Empirical Bayes (EB) methods (Miranda-Moreno et al., 2007; Montella, 2010; Khodadadi et al., 2022). The EB method, although widely used, has limitations in addressing unobserved heterogeneity, which can introduce bias in crash frequency predictions. Recent advancements, such as the simulation-based empirical Bayes (Sim-EB) approach, integrate random parameter models to better capture unobserved heterogeneity in crash risk estimation (Tahir et al., 2022). This refinement enhances the estimation of crash modification factors, improving the precision of safety evaluations, particularly in before-and-after treatment studies.

However, the FB approach remains superior to both EB and Sim-EB methods, as it fully accounts for parameter uncertainty within a hierarchical Bayesian framework rather than relying solely on prior information from reference sites. By generating posterior samples without requiring closed-form solutions, FB provides greater accuracy in parameter estimation and predictions, particularly in complex models. Given these advantages, this study employs the FB method to ensure a comprehensive and precise identification of hazardous bus stops, establishing a robust framework for safety improvements. Expected crash frequencies were estimated, and hazardous bus stops were identified using the Potential for Safety Improvement (PSI) approach. PSI quantifies the difference between expected and predicted crashes, providing a structured measure of safety risk (Hauer et al., 2002). The expected crash frequency at each bus stop was calculated following Eq. (23), and for a more detailed understanding, readers are directed to Khodadadi



(a) Bus stop categories



(b) Hazard corridors

Fig. 8. Hazard bus stops and corridor identification.

et al. (2022). Notably, the PSI method has yet to be applied using the RPNB-L model, presenting an opportunity for future research to explore this innovative approach.

$$p(y_{\text{exp}} \mid y_{\text{obs}}) = \int_{\eta, \phi} N B(e^{-\eta}, \phi) \left(\int_b p(\eta \mid b, y_{\text{obs}}) \pi_b db \right) \pi_\phi d\eta d\phi \quad (23)$$

Based on PSI scores, bus stops were classified into three categories: (1) hotspots, where PSI values are positive and within the top 10%, indicating high-priority locations for safety improvements; (2) normal zones, where PSI values are positive but not in the top 10%, signaling locations that require attention but are less critical; and (3) cold zones, where PSI values are negative, suggesting relatively safer stops (see Fig. 8a). This classification not only identifies hazardous bus stops but also provides a structured understanding of risk levels along transit corridors (see Fig. 8b). Once high-risk corridors are identified,

transportation agencies can prioritize interventions, ensuring effective resource allocation, particularly under budget constraints.

6. Summary and conclusions

This study highlights the growing concern regarding pedestrian safety near bus stops, emphasizing the necessity of identifying critical factors contributing to vehicle–pedestrian crashes in these environments. Previous research has primarily focused on pedestrian crashes at intersections and midblock locations, often treating bus stops merely as contributing factors rather than distinct safety assessment sites. Consequently, there remains a significant knowledge gap regarding the impacts of traffic exposure, roadway environment, and bus stop design on crash frequency near bus stops. Additionally, conventional statistical methods frequently overlook unobserved site-specific variations,

potentially leading to ineffective safety interventions. In addition, it provides valuable insights by identifying hazardous bus stop locations, hotspots, and high-risk corridors. Ultimately, these findings are designed to support transportation agencies in implementing data-driven safety interventions, thereby enhancing the reliability and safety of urban transportation systems.

To address these gaps, the study utilized robust statistical modeling approaches, specifically employing the Random Parameters Negative Binomial-Lindley (RPNB-L) model, to analyze pedestrian crash data from 596 bus stops in Fort Worth, Texas, covering the period from 2018 to 2022. This model effectively captured unobserved heterogeneity and addressed the challenges associated with zero-crash observations and skewed crash data distributions, outperforming competing models (NB-L and RPNB-GE).

Key factors identified as significantly influencing pedestrian crash occurrences included traffic volume (AADT), pedestrian boarding activity, intersection type, proximity to intersections, median type, presence of schools, speed limits, land use characteristics, sidewalk availability, lighting conditions, and the presence of marked crosswalks. Notably, bus stops near signalized intersections experienced fewer crashes due to controlled pedestrian crossings, whereas non-signalized intersections and near-side bus stops presented elevated crash risks. The presence of medians substantially improved pedestrian safety, facilitating safer crossings. Unexpectedly, zones with lower speed limits (35 mph or below) experienced higher crash frequencies, potentially due to increased pedestrian activity and perceived safety that encourages risky crossing behaviors. Similarly, locations near schools demonstrated elevated crash risks, while mixed-use areas were associated with fewer pedestrian incidents. Bus stop design elements, such as sidewalks, clearly marked crosswalks, and adequate lighting, significantly enhanced pedestrian safety. Proper lighting emerged as particularly critical during low-visibility conditions, while the absence of crosswalks markedly increased crash risks. Sidewalk availability reduced pedestrian exposure to vehicular traffic, thereby mitigating crash occurrences. Using the full Bays (FB) based the Potential for Safety Improvement (PSI) metric, the study systematically identified and prioritized hazardous bus stops and corridors, particularly in areas with changing roadway conditions, speed limits, or land-use patterns. The findings provide practical insights to transportation agencies for implementing targeted interventions, including infrastructure improvements, enhanced signage, speed management strategies, and public safety campaigns. Particular attention should be directed toward non-signalized intersections, areas with heavy pedestrian traffic, and school zones.

Despite these valuable insights, this study has limitations. Its geographic focus on Fort Worth, Texas, may restrict the broader applicability of findings. Additionally, reliance on historical crash data introduces potential inaccuracies and underreporting biases (Arun et al., 2021). The absence of real-time pedestrian exposure data limits the precision of model predictions. Although random parameter techniques address unobserved heterogeneity, bias from omitted variables or measurement errors cannot be eliminated entirely. Additionally, this study primarily examines crashes occurring at bus stops, which could overlook risks encountered along pedestrian routes. Furthermore, variables such as weather conditions, law enforcement presence, pedestrian attitudes, and specific needs of vulnerable groups (children, older adults, and individuals with disabilities) were not considered, potentially impacting the comprehensiveness of the conclusions.

Future research directions include expanding the geographic scope to enhance model transferability by adding more urban areas and developing adjustment factors. Further studies should explore multivariate modeling structures, like the Random Parameters Poisson Lognormal-Lindley model (Tahir et al., 2024), to better capture correlations between dependent variables and unobserved heterogeneity across crash severity levels (e.g., fatal and severe injury). Integrating real-time pedestrian exposure data, such as GPS tracking or pedestrian

surveys, could provide a more dynamic understanding of behavior and risk. Future research should also focus on the safety needs of vulnerable populations, assessing the risks faced by children, older adults, and individuals with disabilities who rely heavily on public transportation. Additionally, pedestrian safety in low-density areas, where data are sparse and street vendor encroachment is a common issue, warrants further investigation. Additionally, investigating safety concerns in low-density and resource-constrained areas using innovative approaches, such as autonomous vehicle sensor data for real-time conflict detection (Li et al., 2024; Anis et al., 2024), could significantly advance proactive pedestrian safety management offers a promising alternative to traditional crash data, particularly in resource-constrained settings.

Overall, this study emphasizes the critical need to address pedestrian safety at bus stops systematically. The developed RPNB-L model provides an effective analytical framework to identify risk factors and hazardous locations, enabling targeted and data-driven safety improvements to enhance urban pedestrian safety and support reliable public transportation systems.

CRediT authorship contribution statement

Mohammad Anis: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Srinivas R. Geedipally:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Conceptualization. **Dominique Lord:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Conceptualization.

Disclaimer

The results presented in this document do not necessarily reflect those from the TxDOT.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors express their sincere gratitude to Texas Department of Transportation (TxDOT) and Trinity Metro for providing the crash details and bus stop design data required for this paper.

Data availability

Data will be made available on request.

References

- National Research Council (US). Transportation Research Board. Task Force on Development of the Highway Safety Manual and Transportation Officials. Joint Task Force on the Highway Safety Manual. 2010. Highway Safety Manual. Vol. 1, AASHTO.
- Afghari, A.P., Ismail, K., Saunier, N., Sharma, A., Miranda-Moreno, L., 2014. Pedestrian-cyclist interactions at bus stops along segregated bike paths: a case study of montreal. In: Transportation Research Board (USA) Annual Meeting. 1–1.
- Akintayo, F.O., Adibeli, S.A., 2022. Safety performance of selected bus stops in Ibadan Metropolis, Nigeria. J. Public Transp. 24, 100003.
- Amadori, M., Bonino, T., 2012. A methodology to define the level of safety of public transport bus stops, based on the concept of risk. Procedia- Soc. Behav. Sci. 48, 653–662.
- Anastasopoulos, P.C., Mannering, F.L., 2009. A note on modeling vehicle accident frequencies with random-parameters count models. Accid. Anal. Prev. 41 (1), 153–159.
- Anis, M., Li, S., Geedipally, S.R., Zhou, Y., Lord, D., 2024. Real-time risk estimation for active road safety: Leveraging Waymo AV sensor data with hierarchical Bayesian extreme value models. arXiv preprint arXiv:2407.16832.

- Arun, A., Haque, M.M., Bhaskar, A., Washington, S., Sayed, T., 2021. A systematic mapping review of surrogate safety assessment using traffic conflict techniques. *Accid. Anal. Prev.* 153, 106016.
- Aryuyuen, S., Bodhisuwan, W., 2013. The negative binomial-generalized exponential (NB-GE) distribution. *Appl. Math. Sci.* 7 (22), 1093–1105.
- Barua, S., El-Basyouny, K., Islam, M.T., 2016. Multivariate random parameters collision count data models with spatial heterogeneity. *Anal. Methods Accid. Res.* 9, 1–15.
- Beck, L.F., Dellinger, A.M., O'Neil, M.E., 2007. Motor vehicle crash injury rates by mode of travel, United States: using exposure-based methods to quantify differences. *Am. J. Epidemiol.* 166 (2), 212–218.
- Campbell, B.J., Zegeer, C.V., Huang, H.H., Cynecki, M.J., et al., 2003. A review of pedestrian safety research in the United States and abroad.
- Chand, S., Chandra, S., 2017. Improper stopping of buses at curbside bus stops: Reasons and implications. *Transp. Dev. Econ.* 3, 1–9.
- Clifton, K.J., Burnier, C.V., Akar, G., 2009. Severity of injury resulting from pedestrian-vehicle crashes: What can we learn from examining the built environment? *Transp. Res. Part D: Transp. Environ.* 14 (6), 425–436.
- Craig, C.M., Morris, N.L., Van Houten, R., Mayou, D., 2019. Pedestrian safety and driver yielding near public transit stops. *Transp. Res. Rec.* 2673 (1), 514–523.
- Das, S., Dutta, A., Avelar, R., Dixon, K., Sun, X., Jalayer, M., 2019. Supervised association rules mining on pedestrian crashes in urban areas: identifying patterns for appropriate countermeasures. *Int. J. Urban Sci.* 23 (1), 30–48.
- El-Basyouny, K., Sayed, T., 2009. Urban arterial accident prediction models with spatial effects. *Transp. Res.* 2102 (1), 27–33.
- Elvik, R., 2007. State-Of-The-Art Approaches to Road Accident Black Spot Management and Safety Analysis of Road Networks. Transportøkonomisk institutt Oslo, Norway.
- Fitzpatrick, K., Nowlin, R.L., 1997. Effects of bus stop design on suburban arterial operations. *Transp. Res. Rec.* 1571 (1), 31–41.
- Geedipally, S., 2021. Effects of bus stops on pedestrian safety at signalized intersections. In: Conference of Transportation Research Group of India. Springer, pp. 131–144.
- Geedipally, S.R., Lord, D., Dhavala, S.S., 2012. The negative binomial-lindley generalized linear model: Characteristics and application using crash data. *Accid. Anal. Prev.* 45, 258–265.
- Gelfand, A.E., Sahu, S.K., Carlin, B.P., 1995. Efficient parametrizations for normal linear mixed models. *Biometrika* 82 (3), 479–488.
- Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. *Statist. Sci.* 7 (4), 457–472.
- Haleem, K., Alluri, P., Gan, A., 2015. Analyzing pedestrian crash injury severity at signalized and non-signalized locations. *Accid. Anal. Prev.* 81, 14–23.
- Harwood, D.W., Bauer, K.M., Richard, K.R., Gilmore, D.K., Graham, J.L., Potts, I.B., Torbic, D.J., Hauer, E., 2008. Pedestrian Safety Prediction Methodology. Technical Report..
- Hasan, T., Abdel-Aty, M., 2024. Short-term safety performance functions by random parameters negative binomial-lindley model for part-time shoulder use. *Accid. Anal. Prev.* 199, 107498.
- Hauer, E., 1997. Observational Before/after Studies in Road Safety. Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety.
- Hauer, E., 2004. Statistical road safety modeling. *Transp. Res. Rec.* 1897 (1), 81–87.
- Hauer, E., Kononov, J., Allery, B., Griffith, M.S., 2002. Screening the road network for sites with promise. *Transp. Res.* 1784 (1), 27–32.
- Hauer, E., Persaud, B.N., 1984. Problem of identifying hazardous locations using accident data. *Transp. Res.* 975, 36–43.
- Hess, P.M., Moudon, A.V., Matlick, J.M., 2004. Pedestrian safety and transit corridors. *J. Public Transp.* 7 (2), 73–93.
- Hu, W., Cicchino, J.B., 2018. An examination of the increases in pedestrian motor-vehicle crash fatalities during 2009–2016. *J. Saf. Res.* 67, 37–44.
- Institute, T.T., Foundation, T.A.&M.R., 1996. Guidelines for the Location and Design of Bus Stops. Vol. 19, National Academy Press.
- Jeng, O.-J., Fallat, G., et al., 2003. Pedestrian Safety and Mobility Aids for Crossings at Bus Stops. Technical Report, New Jersey. Department of Transportation. Bureau of Research.
- Khan, S.A., Afghari, A.P., Yasin, S., Haque, M.M., 2023. Effects of design consistency on run-off-road crashes: an application of a Random Parameters Negative Binomial Lindley model. *Accid. Anal. Prev.* 186, 107042.
- Khodadadi, A., Shirazi, M., Geedipally, S., Lord, D., 2023. Evaluating alternative variations of Negative Binomial-Lindley distribution for modelling crash data. *Transp. A: Transp. Sci.* 19 (3), 2062480.
- Khodadadi, A., Tsapakis, I., Shirazi, M., Das, S., Lord, D., 2022. Derivation of the Empirical Bayesian method for the Negative Binomial-Lindley generalized linear model with application in traffic safety. *Accid. Anal. Prev.* 170, 106638.
- Kuşkapan, E., Sahraei, M.A., Çodur, M.K., Çodur, M.Y., 2022. Pedestrian safety at signalized intersections: Spatial and machine learning approaches. *J. Transp. Heal.* 24, 101322.
- Lakhotia, S., Lassarre, S., Rao, K.R., Tiwari, G., 2020. Pedestrian accessibility and safety around bus stops in Delhi. *IATSS Res.* 44 (1), 55–66.
- Lee, J., Abdel-Aty, M., Choi, K., Huang, H., 2015. Multi-level hot zone identification for pedestrian safety. *Accid. Anal. Prev.* 76, 64–73.
- Li, S., Anis, M., Lord, D., Zhang, H., Zhou, Y., Ye, X., 2024. Beyond 1D and oversimplified kinematics: A generic analytical framework for surrogate safety measures. *Accid. Anal. Prev.* 204, 107649.
- Li, D., Ranjitkar, P., Zhao, Y., Yi, H., Rashidi, S., 2017. Analyzing pedestrian crash injury severity under different weather conditions. *Traffic Inj. Prev.* 18 (4), 427–430.
- Lord, D., Geedipally, S.R., 2011. The negative binomial-Lindley distribution as a tool for analyzing crash data characterized by a large amount of zeros. *Accid. Anal. Prev.* 43 (5), 1738–1742.
- Lord, D., Qin, X., Geedipally, S.R., 2021. Highway Safety Analytics and Modeling. Elsevier.
- Lord, D., Washington, S.P., Ivan, J.N., 2005. Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory. *Accid. Anal. Prev.* 37 (1), 35–46.
- Lord, D., Washington, S., Ivan, J.N., 2007. Further notes on the application of zero-inflated models in highway safety. *Accid. Anal. Prev.* 39 (1), 53–57.
- Mannering, F.L., Shankar, V., Bhat, C.R., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. *Anal. Methods Accid. Res.* 11, 1–16.
- Miranda-Moreno, L.F., Labbe, A., Fu, L., 2007. Bayesian multiple testing procedures for hotspot identification. *Accid. Anal. Prev.* 39 (6), 1192–1201.
- Mitra, S., Washington, S., 2007. On the nature of over-dispersion in motor vehicle crash prediction models. *Accid. Anal. Prev.* 39 (3), 459–468.
- Montella, A., 2010. A comparative analysis of hotspot identification methods. *Accid. Anal. Prev.* 42 (2), 571–581.
- Mukherjee, D., Rao, K.R., Tiwari, G., 2023. Built-environment risk assessment for pedestrians near bus-stops: a case study in Delhi. *Int. J. Inj. Control. Saf. Promot.* 30 (2), 185–194.
- Persaud, B., Lyon, C., Nguyen, T., 1999. Empirical Bayes procedure for ranking sites for safety investigation by potential for safety improvement. *Transp. Res. Rec.* 1665 (1), 7–12.
- Pessaro, B., Catalá, M., Wang, Z., Spicer, M., et al., 2017. Impact of transit stop location on pedestrian safety.
- Pulugurtha, S.S., Sambhara, V.R., 2011. Pedestrian crash estimation models for signalized intersections. *Accid. Anal. Prev.* 43 (1), 439–446.
- Pulugurtha, S.S., Vanapalli, V.K., 2008. Hazardous bus stops identification: An illustration using GIS. *J. Public Transp.* 11 (2), 65–83.
- Quistberg, D.A., Howard, E.J., Ebel, B.E., Moudon, A.V., Saelens, B.E., Hurvitz, P.M., Curtin, J.E., Rivara, F.P., 2015a. Multilevel models for evaluating the risk of pedestrian–motor vehicle collisions at intersections and mid-blocks. *Accid. Anal. Prev.* 84, 99–111.
- Quistberg, D.A., Koepsell, T.D., Johnston, B.D., Boyle, L.N., Miranda, J.J., Ebel, B.E., 2015b. Bus stops and pedestrian–motor vehicle collisions in Lima, Peru: a matched case-control study. *Inj. Prev.* 21 (e1), e15–e22.
- Rahman Shaon, M.R., Qin, X., 2016. Use of mixed distribution generalized linear models to quantify safety effects of rural roadway features. *Transp. Res. Rec.* 2583 (1), 134–141.
- Rossetti, S., Tiboni, M., et al., 2020. In field assessment of safety, security, comfort and accessibility of bus stops: A planning perspective. *European Transport/Trasporti Europei* (80).
- Rusli, R., Haque, M.M., Afghari, A.P., King, M., 2018. Applying a random parameters Negative Binomial Lindley model to examine multi-vehicle crashes along rural mountainous highways in Malaysia. *Accid. Anal. Prev.* 119, 80–90.
- Salum, J.H., Kisabanzirwa, N., Alluri, P., Sando, T., 2024. Toolkit for the assessment of safety at bus stops. *Transp. Res. Rec.* 2678 (2), 279–292.
- Samani, R.R., Amador-Jimenez, L., 2023. Exploring road safety of pedestrians in proximity to public transit access points (bus stops and metro stations), a case study of montreal, Canada. *Can. J. Civ. Eng.* 50 (6), 536–547.
- Sanchez Rodriguez, O., Ferenczak, N.N., 2024. Longitudinal spatial trends in US pedestrian fatalities, 1999–2020. *Transp. Res. Rec.* 2678 (5), 904–916.
- Santner, T.J., Duffy, D.E., 2012. The Statistical Analysis of Discrete Data. Springer Science & Business Media.
- Shaon, M.R.R., Qin, X., Shirazi, M., Lord, D., Geedipally, S.R., 2018. Developing a random parameters negative binomial-lindley model to analyze highly over-dispersed crash count data. *Anal. Methods Accid. Res.* 18, 33–44.
- Shinar, D., 2012. Safety and mobility of vulnerable road users: pedestrians, bicyclists, and motorcyclists. pp. 1–2.
- Stewart, T., 2023. Overview of Motor Vehicle Traffic Crashes in 2021. Technical Report.
- Sukor, N.S.A., Fisal, S.F.M., 2020. Safety, connectivity, and comfortability as improvement indicators of walkability to the bus stops in Penang Island. *Eng. Technol. Appl. Sci. Res.* 10 (6), 6450–6455.
- Tahir, H.B., Haque, M.M., Yasin, S., King, M., 2022. A simulation-based empirical bayes approach: Incorporating unobserved heterogeneity in the before-after evaluation of engineering treatments. *Accid. Anal. Prev.* 165, 106527.
- Tahir, H.B., Yasin, S., Haque, M.M., 2024. A poisson lognormal-lindley model for simultaneous estimation of multiple crash-types: Application of multivariate and pooled univariate models. *Anal. Methods Accid. Res.* 41, 100315.
- Thakali, L., Kwon, T.J., Fu, L., 2015. Identification of crash hotspots using kernel density estimation and kriging methods: a comparison. *J. Mod. Transp.* 23, 93–106.
- Tiboni, M., Rossetti, S., 2013. Implementing a road safety review approach for existing bus stops. *WIT Trans. the Built Environ.* 130, 699–709.
- Toran Pour, A., Moridpour, S., Tay, R., Rajabifard, A., 2017. Modelling pedestrian crash severity at mid-blocks. *Transp. A: Transp. Sci.* 13 (3), 273–297.
- Torbic, D.J., Harwood, D.W., Bokenkroger, C.D., Srinivasan, R., Carter, D., Zegeer, C.V., Lyon, C., 2010. Pedestrian safety prediction methodology for urban signalized intersections. *Transp. Res. Rec.* 2198 (1), 65–74.

- Truong, L.T., Somenahalli, S.V., 2011. Using GIS to identify pedestrian-vehicle crash hot spots and unsafe bus stops. *J. Public Transp.* 14 (1), 99–114.
- Tubis, A.A., Skupień, E.T., Rydleński, M., 2021. Method of assessing bus stops safety based on three groups of criteria. *Sustainability* 13 (15), 8275.
- Tyndall, J., 2021. Pedestrian deaths and large vehicles. *Econ. Transp.* 26, 100219.
- Ulak, M.B., Kocatepe, A., Yazici, A., Ozguven, E.E., Kumar, A., 2021. A stop safety index to address pedestrian safety around bus stops. *Saf. Sci.* 133, 105017.
- Vangala, P., Lord, D., Geedipally, S.R., 2015. Exploring the application of the negative binomial-generalized exponential model for analyzing traffic crash data with excess zeros. *Anal. Methods Accid. Res.* 7, 29–36.
- Walgren, S., 1998. Using geographic information system (GIS) to analyze pedestrian accidents. In: 68th Annual Meeting of the Institute of Transportation EngineersInstitute of Transportation Engineers. ITE.
- Washington, S., Karlaftis, M.G., Mannering, F., Anastasopoulos, P., 2020. Statistical and econometric methods for transportation data analysis. Chapman and Hall/CRC.
- Williams, R., 2012. Using the margins command to estimate and interpret adjusted predictions and marginal effects. *Stata J.* 12 (2), 308–331.
- Wretstrand, A., Holmberg, B., Berntman, M., 2014. Safety as a key performance indicator: Creating a safety culture for enhanced passenger safety, comfort, and accessibility. *Res. Transp. Econ.* 48, 109–115.
- Xie, K., Ozbay, K., Kurkcu, A., Yang, H., 2017. Analysis of traffic crashes involving pedestrians using big data: Investigation of contributing factors and identification of hotspots. *Risk Anal.* 37 (8), 1459–1476.
- Yendra, D., Haworth, N., Watson-Brown, N., 2024. A comparison of factors influencing the safety of pedestrians accessing bus stops in countries of differing income levels. *Accid. Anal. Prev.* 207, 107725.
- Yu, C.-Y., 2024. The impact of built environments on pedestrian safety around bus stops. *J. Urban Aff.* 46 (2), 267–280.
- Zahabi, S.A.H., Strauss, J., Manaugh, K., Miranda-Moreno, L.F., 2011. Estimating potential effect of speed limits, built environment, and other factors on severity of pedestrian and cyclist injuries in crashes. *Transp. Res. Rec.* 2247 (1), 81–90.
- Zamani, H., Ismail, N., 2010. Negative binomial-lindley distribution and its application. *J. Math. Stat.* 6 (1), 4–9.
- Zegeer, C.V., Bushell, M., 2012. Pedestrian crash trends and potential countermeasures from around the world. *Accid. Anal. Prev.* 44 (1), 3–11.
- Zhang, C., Du, B., Shen, J., Ma, C., 2023. Revealing safety impact of bus stops on passenger-cyclist interactions—Evidence from nanjing, China. *Travel. Behav. Soc.* 32, 100578.