Transport Policy 156 (2024) 164–181

Contents lists available at ScienceDirect

Transport Policy

journal homepage: www.elsevier.com/locate/tranpol

Spatial analysis of geographical disparities in pedestrian safety

Jinli Liu [a,*], Subasish Das [b], F. Benjamin Zhan [a], Md Nasim Khan [b,c]

[a] Geography and Environmental Studies, Texas State University, 601 University Drive, San Marcos, TX, 78666, USA
[b] Civil Engineering, Ingram School of Engineering, Texas State University, 601 University Drive, San Marcos, TX, 78666, USA
[c] AtkinsRealis, 11801 Domain Blvd Suite 500, Austin, TX 78758, United States

ARTICLE INFO

Keywords:
Pedestrian safety
Macro-level analysis
SAR
MGWR
Pedestrian safety disparities

ABSTRACT

Investigating pedestrian safety disparities across sociodemographic groups is essential for enhancing traffic safety. This study examines the impact of sociodemographic and built environment characteristics on pedestrian crashes. It introduces a comprehensive macro spatial analysis framework that includes a global regression model, spatial autoregressive models, and a local spatial regression model. Three measures of pedestrian injury are analyzed. The findings reveal that a higher percentage of the high-income population significantly correlates with lower rates of pedestrian injuries across all three measures. Conversely, a higher percentage of the low-income population shows a significant positive correlation with the proportion of crashes involving the Black population, and with the proportion of severe pedestrian crashes involving the Black population. Pedestrian-oriented network density is negatively associated with fatal or severely injurious crashes involving the Black population. These results emphasize the need to account for spatial variations and equity when addressing pedestrian safety disparities.

1. Introduction

Crash injuries are a significant global public health concern, ranking as the eighth leading cause of death worldwide (CDC, 2023). They are particularly prevalent among young people aged fifteen to twenty-nine, where they stand as the leading cause of death. The World Health Organization (WHO) emphasizes the severity of the issue, noting that road traffic crashes annually claim approximately 1.3 million lives and result in non-fatal injuries to 20 to 50 million individuals globally. Vulnerable road users, including pedestrians, cyclists, motorcyclists, and passengers, comprise over half of all road traffic deaths and injuries (WHO, 2023). Among these, pedestrian crashes often face severe outcomes; millions sustain non-fatal injuries, with some leading to permanent disabilities, and fatalities are distressingly common. Certain countries report that up to two-thirds of road traffic deaths involve pedestrians. Focusing on the United States, the situation is similarly alarming. In 2021, there were 7388 pedestrians killed in traffic crashes, marking a 12.5-percent increase compared to 2020 and the highest number of pedestrian fatalities since 1981.

Furthermore, approximately 60,577 pedestrians sustained injuries in traffic crashes in 2021, an 11-percent rise from the previous year. The trend continued to escalate in 2022, with over 7500 pedestrians struck and killed by vehicles. Notably, pedestrian deaths accounted for 17 percent of all traffic fatalities in the U.S. (NHTSA, 2021), underlining the critical need for enhanced pedestrian safety measures.

In response to these concerning statistics, significant efforts are being made to improve pedestrian safety. The U.S. Department of Transportation actively contributes to these efforts by launching initiatives such as the $15 million Complete Streets Artificial Intelligence project for small businesses (U.S. DOT, 2024). This project aims to encourage creating a safe, comfortable, and integrated network for all users, emphasizing the importance of comprehensive street design that accommodates the needs of pedestrians and enhances overall road safety.

Furthermore, national pedestrian death rates reveal a disproportionate representation of Black and Hispanic pedestrians in fatal injury statistics compared to the rate for White pedestrians (CDC, 2020). The growing emphasis on walking and the inherent vulnerability of pedestrians have recently intensified research interest in pedestrian safety (Roll and McNeil, 2022a; Schneider, 2020). Many studies have highlighted disparities in traffic fatalities, noting a significant impact on Black, Indigenous, and People of Color (BIPOC), elderly groups (Gálvez-Pérez et al., 2022; Kim and Ulfarsson, 2019), and individuals with lower incomes (Stoker et al., 2015). Research indicates that areas characterized by lower incomes and higher poverty rates are associated

* Corresponding author.
E-mail addresses: jinli.liu@txstate.edu (J. Liu), subasish@txstate.edu (S. Das), zhan@txstate.edu (F.B. Zhan), mdnasim.khan@atkinsrealis.com (M.N. Khan).

https://doi.org/10.1016/j.tranpol.2024.06.018
Received 3 April 2024; Received in revised form 14 June 2024; Accepted 22 June 2024
Available online 25 June 2024

with an increased risk of pedestrian injuries and fatalities. For instance, a national study in 2019 demonstrated that a $1000 reduction in a census tract's median income is linked to a one percent increase in pedestrian fatalities (Mansfield et al., 2018). Addressing these disparities in pedestrian safety is crucial for cultivating inclusive and equitable communities.

In response to this need, pedestrian safety studies encompass crash count analysis, crash race prediction, crash rate analysis, and more, employing individual-level and aggregated-level analyses to capture the complexity of these crashes. Individual-level analysis examines each case crash, including pedestrian crash information and demographic characteristics of the pedestrian, alongside built environment features of the crash location, providing detailed statistics. On the other hand, aggregated analysis involves grouping pedestrian crash points into geographical units, with socio-demographic and built environment characteristics of these spatial units being collected, and the relationships were analyzed. Regardless of the approach, pedestrian crash analyses often involve spatial data. By nature, spatial data may exhibit spatial autocorrelation or spatial non-stationarity (Anselin, 2013). Spatial autocorrelation refers to the tendency of data points close to one another in space to exhibit similar values. For instance, in the context of a pedestrian crash rate study, areas close to each other may show similar crash rates due to shared traffic conditions, urban design, or pedestrian behaviors.

On the other hand, spatial non-stationarity describes the variation in statistical relationships across space. For pedestrian safety analysis, this might mean that the factors influencing pedestrian crashes have different impacts in different parts of the study area. Several studies have addressed spatial autocorrelation (Adeleke et al., 2020; Pljakić et al., 2019; Tokey et al., 2023), while others have focused on identifying spatial non-stationarity (Gomes et al., 2017; Li et al., 2013). Given these complexities, employing a systematic spatial analysis framework is essential to effectively identifying and understanding geographical disparities in pedestrian safety. New spatial models, such as Multiscale Geographically Weighted Regression (MGWR), present promising methods for such analyses but have not been extensively studied. Furthermore, there is a significant gap in research focused on comparing disparities in pedestrian crash outcomes across different racial groups.

Therefore, this study aims to investigate factors influencing pedestrian crashes and explore safety disparities at the Census Block Group (CBG) level. By employing spatial analysis techniques, this study seeks to understand how various variables interact and exhibit spatial variations. The study analyzes five years of pedestrian-related traffic crash data (2017–2021) in Louisiana, utilizing a dataset compiled from multiple sources, including socio-demographics, employment and housing diversity, land use diversity, road network characteristics, and transit accessibility. A comprehensive spatial modeling framework has been proposed, incorporating global regression, Ordinary Least Squares (OLS), Spatial Autoregressive Regression (SAR), and Multiscale Geographically Weighted Regression (MGWR). These models enable the exploration of potential spatial autocorrelation and non-stationarity, providing insights into spatial variations and local relationships between variables and pedestrian crashes. Three measures of pedestrian injury are analyzed, including pedestrian crash rates, the ratio of crashes involving the Black population in total crashes, and the ratio of crashes involving the Black population in severe pedestrian crashes. This study formulates several hypotheses to guide the investigation into the disparities in pedestrian safety, particularly focusing on BIPOC communities and individuals with lower incomes at the CBG level.

2. Literature review

The study reviewed macro-level traffic safety analysis studies, particularly on pedestrians. These studies explored four key research questions. The first issue concerns determining the appropriate spatial units for aggregated crash safety analysis. These units range from segment and intersection levels to zonal, regional, mixed, or customized levels (Ziakopoulos and Yannis, 2020). This is important because different spatial units may encounter the Modifiable Areal Unit Problem (MAUP) (Ziakopoulos and Yannis, 2020). However, a universally accepted method for the correct spatial units in pedestrian safety analysis has yet to be available. The choice depends on the specific research goals, applications, and data availability. In segment and intersection level safety analysis, researchers often aggregate crashes into segments or intersections as the primary spatial units of analysis. Many studies have adopted this approach and, for example, (Kim, 2019) used intersections as the unit of analysis for pedestrian safety by age group. Ling et al. (2023) investigated right-turn lane crash frequency at the intersection level, while (Stipancic et al., 2020) focused on pedestrian injuries at the intersection level. (Aguero-Valverde, 2014; Aguero-Valverde and Jovanis, 2008) analyzed road crash frequency by segment. Mathew et al. (2022) examined teen crash frequencies by segments, extracting road network and traffic data for these segments, along with relevant land use and demographic characteristics. Moving to zonal level analysis, researchers commonly use areal units such as census tracts (CT), census block groups (CBGs), and Traffic Analysis Zones (TAZ). CTs have been widely employed in spatial pedestrian safety analysis (Patwary et al., 2024; Roll and McNeil, 2022b; Sanders and Schneider, 2022; Tasic et al., 2017). Similarly, TAZ units have found application in several pedestrian safety studies (Merlin et al., 2020; Osama and Sayed, 2017; Rahman et al., 2019; Zafri and Khan, 2022). Additionally, CBGs have been utilized in general traffic safety analysis (Li et al., 2022; Liu et al., 2024). For regional level analysis, the spatial units of interest often encompass county (Li et al., 2013), city (Gálvez-Pérez et al., 2023b; Moeinaddini et al., 2014), municipality (Gálvez-Pérez et al., 2024), and metropolitan area scales (Lee et al., 2018, 2019). Researchers have also explored customized units, including grid cell-level pedestrian crash cost analysis and Thiessen polygons level crash frequency analysis (Wang and Kockelman, 2013). Several studies integrate two or more levels of units into their analyses, with segment-level and CT or CBG level information being the most common combination. For example, Jin et al. (2023) aggregated information at both segment and CBG levels to predict the number of E-scooter sharing link flows. In the context of pedestrian safety research, regional-level analysis is often considered too broad. Accordingly, census tracts or CBGs are more suitable alternatives because they capture detailed information. It's important to note that while TAZs are also employed in similar research, the data for TAZ are currently outdated. Consequently, this analysis used CBGs as the primary unit of study.

The second question concerns the selection of explanatory variables, which primarily centers on three approaches. The first approach concentrates exclusively on roadway characteristics, geometric design, and related influences (Guo et al., 2017; Stipancic et al., 2020). The second approach adopts a more comprehensive set of variables, including socioeconomic, demographic, racial, and built environmental factors, thus facilitating an exploration from a social equity perspective (Roll and McNeil, 2022b; Wang and Kockelman, 2013; Mathew et al., 2022; Sanders and Schneider, 2022; Wu et al., 2024; Liu et al., 2024). Meanwhile, some studies focus solely on social vulnerability and demographic characteristics (Li et al., 2022). The commonly used variables mainly include: Land use (Cai et al., 2019; Mathew et al., 2022; Pfiester et al., 2021; Wang and Kockelman, 2013; Xie et al., 2017); Sociodemographic characteristics (Mathew et al., 2022; Pfiester et al., 2021; Roll and McNeil, 2022b; Sanders and Schneider, 2022; Zafri and Khan, 2022; Wu et al., 2024; Liu et al., 2024); Roadway and traffic characteristics (Cai et al., 2019, 2016; Osama and Sayed, 2017; Rahman et al., 2019; Siddiqui and Al-Kaisy, 2017; Wang et al., 2016; Wang and Kockelman, 2013); Transit accessibility (Merlin et al., 2020; Sanders and Schneider, 2022; Jin et al., 2023; Liu et al., 2024), few included environmental variables (Cottrill and Thakuriah, 2010; Lee et al., 2018; Zhai et al., 2019). After reviewing these studies, a wide range of variables, encompassing socio-demographics, employment and housing diversity, land use diversity, road network characteristics, and transit accessibility are included in the analysis.

The third question relates to determining suitable exposure variables. In pedestrian traffic safety analysis, obtaining accurate pedestrian volume data can be challenging. As a result, some studies does not include exposure variables, researchers who use exposure variables often adopt proxy measures like nighttime population (Lee et al., 2015; Merlin et al., 2020; Wang et al., 2016), daytime population (Roll and McNeil, 2022b), or vehicle miles traveled as substitutes for pedestrian crash exposure (Li et al., 2022; Osama and Sayed, 2017). Some studies adopt a two-step modeling approach, initially predicting pedestrian volume before proceeding with safety analysis (Lee et al., 2019). In this study, total nighttime population, and total travel demand were selected as two potential exposure variables.

The fourth question regards the choice of modeling techniques, which can usually be divided into non-spatial or spatial modeling. Some studies adopted non-spatial modeling techniques, which encompass a range of methods such as global statistical regression. Examples include the Negative Binomial model (Patwary et al., 2024; Roll and McNeil, 2022b; Xu and Huang, 2015), Multinomial logit models (Sanders and Schneider, 2022), Ordinary Least Squares (OLS) (Liu et al., 2024; Pfiester et al., 2021), and machine learning models (Gálvez-Pérez et al., 2023a; Jin et al., 2023; Li et al., 2017; Mokhtarimousavi et al., 2020; Rahman et al., 2019). On the other hand, spatial data may exhibit potential spatial autocorrelation and spatial non-stationarity (Anselin, 2013). In such cases, spatial modeling is more appropriate than non-spatial methods. Some studies have typically focused on random parameter modeling (Xu and Huang, 2015), which includes both random intercept and random slope models. The term random parameter arises from the incorporation of random effects, allowing for the modeling of individual-specific variations (or, in a spatial context, location-specific variations) in the model parameters. Some spatial modeling methods require spatial weight features. These include models like the Spatial Autoregressive Model (SAR) (Quddus, 2008; Zafri and Khan, 2022), Full Bayesian (FB) models with spatial effects, Bayesian Conditional Autoregressive models (Cai et al., 2019; Wang and Kockelman, 2013), all of which fall under the umbrella of global spatial regression models. Local spatial regression models like Geographically Weighted Regression (GWR) (Gomes et al., 2017; Liu et al., 2024; Pfiester et al., 2021; Wu et al., 2024) and Multiscale Geographically Weighted Regression (MGWR) (Li et al., 2022; Liu et al., 2024) are another approaches. MGWR, an advanced form of GWR, has not received extensive attention in past research. Furthermore, most studies have not adopted a systematic spatial analysis approach to explore the influence of variables and disparities in pedestrian safety.

3. Methodology

3.1. Study design

This study explores the relationships between socioeconomic, demographic, and built-environment factors at the CBG level in Louisiana and their association with pedestrian crashes. As presented in Fig. 1, multi-source spatial data from the U.S. Census Bureau, OpenStreetMap, and the Smart Location Database (SLD) were compiled. Specifically, socio-demographic data were sourced from the Census, bus station distribution was extracted from OpenStreetMap, and transit, road network, employment, and land use density information were gathered from the SLD. These variables were chosen based on an extensive review of the literature. The candidate independent variables are then meticulously selected to mitigate multicollinearity. The Variance Inflation Factor (VIF) scores are computed for each chosen variable to identify multicollinearity issues. Variables with a VIF score exceeding five are considered to exhibit moderate multicollinearity and are systematically excluded one by one until all remaining variables have VIF scores below 5 (Kock and Lynn, 2012). In this study, crash rates at CBGs serve as the response variables. The pedestrian crash rate is the pedestrian crash count divided by the number of trips generated at each CBG. Additionally, we examine two other response variables: the proportion of crashes involving the Black population relative to total crashes and the proportion of crashes involving the Black population that result in KAB-level outcomes.

The study begins by fitting an OLS regression model and then evaluates the adequacy of a global model in analyzing the relationships. This evaluation involves applying Moran's I to the regression residuals. A significant result from Moran's I indicates the presence of spatial

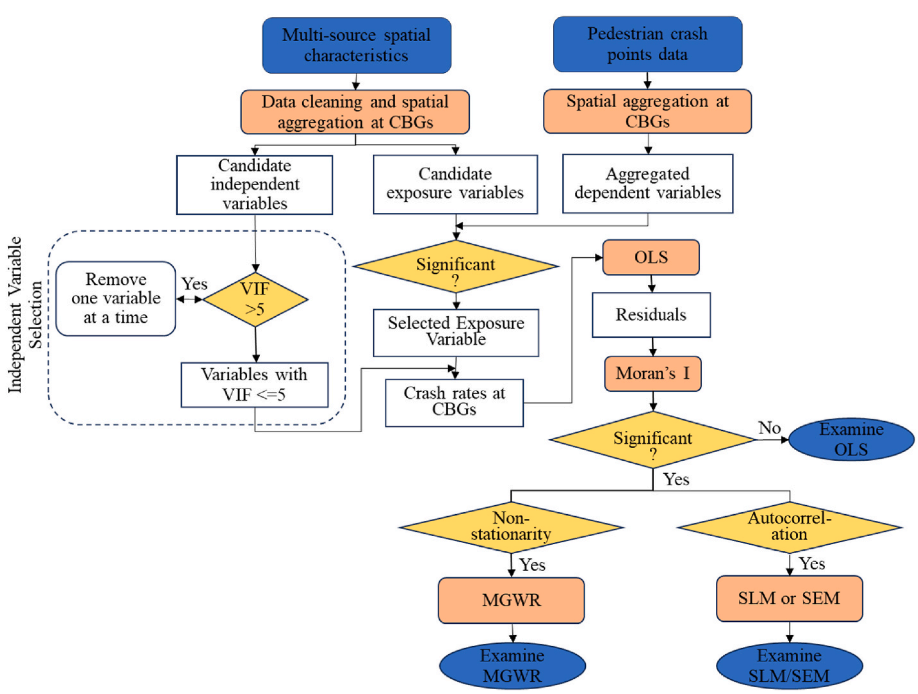


Fig. 1. Study flowchart.

correlation in the residuals, which could result from either spatial autocorrelation or non-stationarity. To address this, if spatial structure is identified in the residuals, spatially explicit models are considered. For spatial autocorrelation, the Spatial Lag Model and Spatial Error Model are employed to account for spatial dependencies in the data. To address spatial non-stationarity, a MGWR is utilized. MGWR, an advanced form of geographically weighted regression, allows for local variation in the relationships between variables across space. Subsequently, the study analyzes the relationships between the factors and pedestrian crashes. Based on the findings, the research offers relevant recommendations and insights to address pedestrian safety concerns in Louisiana. This comprehensive approach ensures a robust analysis of the factors influencing pedestrian crashes. It identifies spatial variations and potential spatially dependent relationships, enabling the development of effective strategies for mitigating pedestrian crashes and improving overall pedestrian safety in the region. The following section describes the methods used in this study.

3.2. Theory

3.2.1. Ordinary Least Squares (OLS)

Regression analysis is a widely used statistical method in the social sciences to assess relationships between two or more attributes. The most well-known technique, OLS, fits a model that minimizes the sum of squared errors between explanatory variables and a continuous or interval outcome variable (Burton, 2021). OLS serves as the initial step for spatial regression analyses, offering a global model that represents the variable or process of interest. It assumes consistent relationships between variables across space. Researchers can further evaluate model performance and refine their understanding or predictions by calculating residuals, which are the differences between observed and predicted values of the dependent variable.

3.2.2. Global Moran's I

The Global Moran's I assesses spatial autocorrelation, considering both feature locations and attribute values simultaneously (Getis and Ord, 1992). It examines whether the pattern displayed by the features is clustered, dispersed, or random. The tool available in ArcGIS Pro calculates Moran's I Index value, along with a z-score and p-value, to determine the significance of the index. The Moran's I statistic is used to quantify spatial autocorrelation:

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} z_i z_j}{S_0 \sum_{i=1}^{n} z_i^2} \tag{1}$$

Where z_i denotes the deviation of an attribute for feature $I(x_i - \widehat{X})$, $w_{i,j}$ denotes the weight between feature i and j, n denotes the total number of features, and S_0 denotes the aggregate of all the spatial weights:

$$S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} \tag{2}$$

Where z_i is calculated as:

$$z_i = \frac{I - E[I]}{\sqrt{V[I]}} \tag{3}$$

Where:

$$E[I] = -1 \Big/ (n-1) \; V[I] = E\left[I^2\right] - E[I]^2 \tag{4}$$

3.2.3. Spatial autoregressive (SAR) models

The Spatial Lag Model (SLM) and the Spatial Error Model (SEM) are two types of Spatial Autoregressive (SAR) models, which are types of regression models that explicitly incorporate spatial dependence into their structure. These models consider not only the independent variables that may influence the dependent variable but also account for how observations at nearby locations influence observations at one location. SLM focuses on the dependency in the dependent variable itself, while SEM addresses the spatial autocorrelation in the error terms of the regression model (Gaspard et al., 2019).

The SAR model can be formulated as follows:

$$Y = \rho WY + X\beta + \varepsilon \tag{5}$$

Where Y is the dependent variable vector, ρ is the spatial autoregressive parameter. W is the spatial weights matrix. X is the matrix of independent variables; β is the vector of coefficients for the independent variables; ε is the vector of error terms.

The SEM model can be formulated as follows:

$$Y = X\beta + \mu \; \mu = \lambda W\mu + \varepsilon \tag{6}$$

Where Y is the vector of the dependent variable; X is the matrix of independent variables; β is the vector of coefficients for the independent variables; μ is the vector of error terms, which itself is modeled as a spatially autoregressive process. λ is the vector of error terms. W is the spatial weights matrix. ε is the vector of independent and identically distributed error terms.

3.2.4. Multiscale geographically weighted regression (MGWR)

GWR is a spatial regression technique commonly employed in geography and other disciplines. It examines the local model of the variable or process of interest by fitting a regression equation to each feature in the dataset. These individual equations are constructed by considering the dependent and explanatory variables of features within the neighborhood of each target feature. For the GWR model, the specification is as follows (Bo, 2018):

$$y_i = \sum_{j=1}^{p} x_{ij}\beta_{ij} + \varepsilon_i, \tag{7}$$

Where y_i is the predicted value at location i. x_{ij} is the value of jth independent variable for location i, β_{ij} is a location-specific coefficient corresponding to x_{ij}, and ε_i is a random error at location i.

The MGWR is an extension of the GWR model. While GWR uses a fixed neighborhood around each feature to create a local linear regression model, MGWR allows each explanatory variable's neighborhood size to vary (ESRI, 2023). This variation is essential as different variables may operate on various spatial scales. Some variables may have gradual changes across the study area, while others may exhibit rapid changes. By matching the neighborhood of each explanatory variable to its spatial scale, MGWR can more accurately estimate the coefficients of the local regression model for interpretation and prediction. This study uses the GWR and MGWR tools available in ArcGIS Pro software to perform the modeling.

3.3. Data collection

3.3.1. Dependent variables

The crash data used in this study was obtained from the Louisiana Department of Transportation and Development (LaDOTD). Specifically, records of pedestrian crashes between 2017 and 2021 were extracted for analysis. The data includes five levels of crash severity categorized under the KABCO injury scale: K for fatal, A for incapacitating, B for non-incapacitating, C for possible, and O for no injury or property damage-only or PDO crashes. A total of 8213 records of pedestrian-related crashes were analyzed, and the distribution of the number of crashes by year is visualized in Fig. 2 a).

Pedestrian data was aggregated at the CBG level, and the total number of pedestrian crashes ranged from 0 to 157. There is a total of

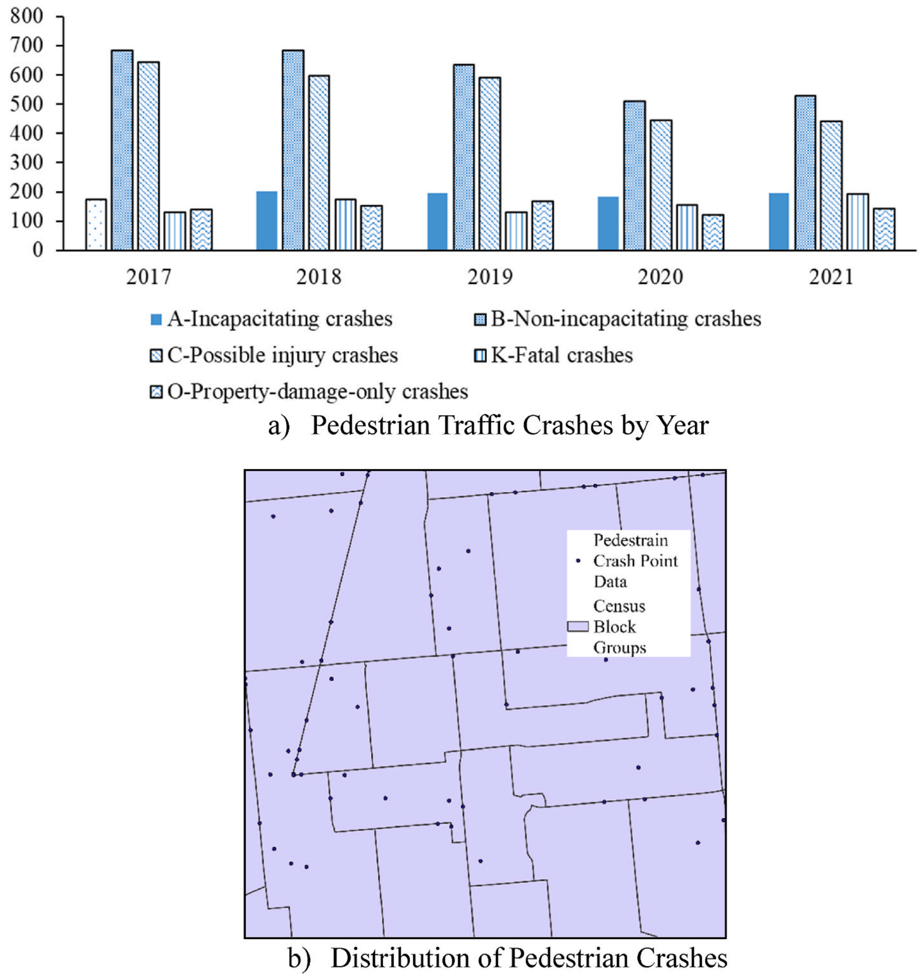


Fig. 2. Pedestrian crash point data in Louisiana.

3471 CBGs in Louisiana. Given that pedestrian crashes frequently occur on roads, which often delineate the boundary lines of CBGs as depicted in Fig. 2 b), a strategic approach was employed involving the creation of a polygon buffer. Specifically, a 10-m buffer zone was established surrounding each CBG. This methodology ensures that crashes occurring within these designated buffer zones are accounted for in both adjacent CBGs, providing a more comprehensive and accurate representation of crash data. Fig. 3 highlights that CBGs with the highest number of pedestrian crashes are predominantly located in urbanized regions such as New Orleans, Baton Rouge, Lake Charles, and Lafayette.

Upon further investigation, it was found that the race involved in pedestrian crashes was not evenly distributed; the black population is the most vulnerable. Fig. 4 presents the distribution of the ratio of the black population involved in total crashes (left) and KAB level crashes (right), excluding CBGs with zero crashes. These histograms indicate the distribution of the ratios, not the absolute number of crashes. A ratio of 1 in a specific crash means that all individuals involved were from the black population, regardless of the total number of individuals involved.

The distribution is bimodal for the ratio of crashes involving the black population in total crashes, with peaks around 0 and 1. This suggests that there are many areas where either none or all the total crashes involve the black population. There are also a substantial number of places where the ratio is around 0.5, indicating that the black population is involved in about half of the total crashes in these areas. The mean is approximately 0.501, suggesting that, on average, about 50.1% of the crashes involved the black population.

The right histogram focuses on KAB-level crashes involving fatalities and severe injuries. The distribution in this histogram also shows a high

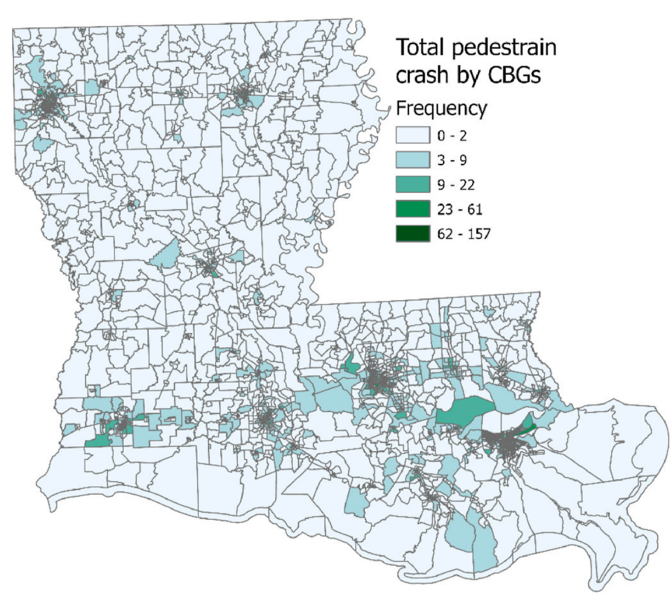


Fig. 3. Pedestrian traffic crashes aggregated at CBGs.

frequency at the extremes (0 and 1). Notably, the frequency of KAB level crashes is higher at the 0 proportion than the total crashes histogram, suggesting a lower number of severe crashes in CBGs with an entirely

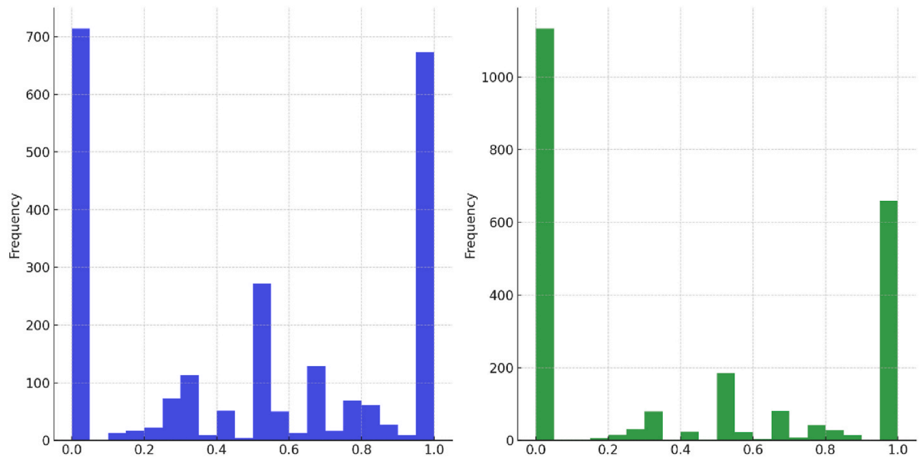


Fig. 4. Histogram of the distribution of the black population involved in total crashes (left) and KAB level crashes (right), excluding CBGs with zero crashes.

black population. The mean is approximately 0.403, suggesting that, on average, about 40.3% of the KAB-level crashes involved the black population. In summary, there are many crashes where the black population is involved in crashes. The distributions also suggest a high variability across different areas, which demographics, road conditions, and driving behaviors could influence. The distribution demonstrates the need to investigate social equity in pedestrian safety further.

3.3.2. Traffic exposure

In many prior studies, total population and population density have frequently been employed as a proxy for exposure (Merlin et al., 2020; Wang et al., 2016). In our preceding research, we discovered that the estimated trips generated for each area served as a robust measure of exposure (Liu et al., 2024). These trip estimates, encompassing both production and attraction, were derived using rates from the Institute of Transportation Engineers (ITE) Trip Generation. In the current study, we explored the correlation between both these variables and the total number of pedestrian crashes through two Negative Binomial Poisson regressions. The analysis revealed that the p-value for the relationship between total population and crashes stood at 0.05, whereas the p-value for the connection between total trips generated and crashes was less than 0.001. Consequently, the total trips generated has been adopted as a more reliable estimate for exposure in our analysis. Crash rate at *i*th CBG were derived using the formulas: $CrashRate_i = Pedestrain\ Crash\ Counts_i * 1000 / TotVT_i$.

3.3.3. Independent variables

Multi-source spatial data from the U.S. Census Bureau, OpenStreetMap, and SLD were compiled, encompassing socio-demographics, employment and housing diversity, land use diversity, road network characteristics, and transit accessibility. Specifically, socio-demographic data were sourced from the U.S. Census Bureau's 2018 records. Bus station distribution was extracted from OpenStreetMap, while information on transit, road network, employment, and land use density was gathered from the SLD. The SLD is an emerging, comprehensive resource that offers a wide array of built-environment information at the CBG level (US EPAO, 2014). These variables were selected based on previous studies on macro traffic safety studies, data availability, and the exclusion of variables with multicollinearity issues. Table 1 provides descriptive statistics for variables used in this study. The dependent variables are pedestrian crash rates, the ratio of the black population in total crashes, and the ratio of the black population in KAB-level crashes. For each variable, the table presents the following descriptive statistics: mean, standard deviation (Std), minimum (Min), median, and maximum (Max) values. These statistics summarize the data distribution and central tendencies for each variable.

4. Results and discussions

4.1. Ordinary Least Squares and global Moran's I

The dataset was analyzed using the global regression model, OLS. The results, detailed in Table 2, reveal a multiple R-squared value of 0.2557 and an adjusted R-squared value of 0.2514. These figures suggest that the model explains approximately 25.57% of the variance in the dependent variable, indicating a modest explanatory power. The Akaike's Information Criterion (AICc) is 8996, further confirming the model's moderate fit. According to Kock and Lynn (2012), all variables have VIF values below 5, indicating a moderate level of multicollinearity that is considered acceptable for this analysis. Several variables demonstrated statistically significant coefficients ($p < 0.01$) under both regular and robust standard errors, indicating a notable positive association with crash rates. These variables include PctWrkAge, PctWorker, PctZeroCar, PctHiWaWrk, PctLowWaWrk, ResD, TripEQ, AutoNetD, IntsecD, EmpAcc, and BusStopsD. On the contrary, variables such as PctHiWa, EmpD, JobPop, and PedNetD exhibited negative coefficients, indicating a significant inverse relationship with pedestrian crash rates (see Table 3).

After fitting the OLS regression model, a Moran's I test was conducted on the residuals to assess the presence of spatial structure in the data, which includes spatial autocorrelation and spatial non-stationary. The Moran's Index, at 0.059978, suggests a weak clustering tendency. The accompanying p-value of less than 0.0001 indicates this clustering is statistically significant. The presence of spatial autocorrelation implies that the OLS model may not fully capture specific spatial characteristics of the data. Significantly autocorrelated residuals could indicate spatial autocorrelation and spatial non-stationarity (Anselin, 2013; Gaspard et al., 2019), implying that the OLS model does not sufficiently capture spatially varying relationships. Adopting a spatial autoregressive model or a geographically weighted regression might be more suitable, as these models account for relationships across different spatial locations. The subsequent sections present the results of both types of models.

4.2. Spatial autoregressive modeling results

The table presents the results from two spatial regression models, the Spatial Lag Model (SLM) and the Spatial Error Model (SEM), employed in scenarios where spatial autocorrelation exists in the data. This implies that the value of crash rate at one location is influenced by the values of the same variable at neighboring locations. Starting with model performance, the Pseudo R-squared for the SLM is 0.4484, indicating that this model explains a higher proportion of variance than the OLS model. Conversely, the Pseudo R-squared for the SEM is 0.1944, lower than that

Table 1
Descriptive statistics of variables.

Explanation	Variable	Mean	Std	Min	Median	Max
*Dependent variable*rowhead						
Pedestrian crash rate (Total*1000/TotVT)	CrashRate	1.07	2.01	0.00	0.53	36.76
The ratio of crashes that involve the black population in total crashes	BLTR	0.38	0.39	0.00	0.29	1.00
The ratio of crashes that involve the black population in KAB level crashes	BLKABR	0.32	0.41	0.00	0.00	1.00
Independent variablesrowhead						
*Socio-Demographics*rowhead						
Percent of working-age population (18–64 years)	PctWrkAge	0.59	0.11	0.00	0.59	1.00
Percent of zero-car households	PctZeroCar	0.10	0.12	0.00	0.06	0.81
Percent of two-plus-car households	PctMultiCar	0.51	0.20	0.00	0.53	1.00
Percent of high income workers (home location)	PctHiWa	0.36	0.12	0.00	0.37	0.75
Percent of low income workers (home location)	PctLowWa	0.26	0.06	0.00	0.25	0.53
*Employment & Housing Diversity*rowhead						
Percent of high wage employment (work location)	PctHiWaWrk	0.31	0.19	0.00	0.28	1.00
Percent of low wage employment (work location)	PctLowWaWrk	0.29	0.16	0.00	0.29	1.00
Employment and household entropy	EmpHHEnt	0.49	0.24	0.00	0.50	0.99
Regional diversity	JobPop	0.30	0.29	0.00	0.21	1.00
Trip productions and trip attractions equilibrium index	TripEQ	0.37	0.31	0.00	0.40	1.00
*Land Use Density*rowhead						
Gross residential density (HU/acre)	ResD	2.15	3.00	0.00	1.08	31.05
Gross employment density (jobs/acre)	EmpD	1.65	5.36	0.00	0.27	145.10
*Road Network Characteristics*rowhead						
Auto-oriented network density (per square mile)	AutoNetD	0.89	2.00	0.00	0.19	25.32
Pedestrian-oriented network density (per square mile)	PedNetD	9.70	8.55	0.00	8.24	56.45
Intersection density	IntsecD	64.58	77.11	0.00	41.23	1393.37
Auto oriented regional centrality index	EmpAcc	0.49	0.30	0.00	0.53	1.00
*Transit Accessibility*rowhead						
Proportion of employment within ¼ mile of transit stop	PctEmpHalfM	0.02	0.12	0.00	0.00	1.00
Bus stop density (per acre)	BusStopsD	0.00	0.01	0.00	0.00	0.28

of the OLS model. The AICc values for both models are comparable and lower than that of the OLS model.

In the SLM, the inclusion of the spatially lagged dependent variable, W_CrashRate, with a coefficient of 0.5782, underscores strong spatial dependence, as reflected by its highly significant p-value. Variables such as PctWrkAge, PctWorker, PctZeroCar, PctHiWaWrk, TripEQ, AutoNetD, IntsecD, EmpAcc, and BusStopsD display significant positive coefficients, aligning with the direction of influence observed in the OLS model. Variables including PctHiWa, EmpD, and PedNetD show significant negative coefficients, consistent with the OLS model. However, PctLowWaWrk and ResD are not significant in this model.

The SEM addresses spatial autocorrelation in the error terms, positing that the error at one location is correlated with the error in adjacent locations. The lambda parameter, valued at 0.631, is significant, indicating spatial autocorrelation in the model's residuals. The SEM yields a lower R-squared than the OLS, suggesting inferior modeling performance in accounting for the spatial structure. Hence, further discussion on this model is omitted. In conclusion, while both models offer valuable insights, the superior fit of the SLM suggests that the spatial autocorrelation in the data might be more attributable to the lagged terms rather than the dependent variable itself. In this study, we further analyzed spatial non-stationarity by employing MGWR in subsequent analyses.

4.3. MGWR modeling results

The MGWR model diagnostics reveal an improved goodness-of-fit compared to the SLR, evidenced by an R-squared value of 0.5514. The

Table 2
Summary of OLS regression results for crash rate analysis.

Variable	Coefficient	S.T.D.	t-Statistic	p-value	VIF
Intercept	0.0000	0.0000	−1.3315	0.1831	–
PctWrkAge	0.0821	0.0164	4.9976	0.0000*	1.2496
PctWorker	0.0885	0.0154	5.7473	0.0000*	1.0987
PctZeroCar	0.1743	0.0209	8.3215	0.0000*	2.0335
PctMultiCar	0.0437	0.0230	1.8998	0.0575	2.4519
PctHiWa	−0.2533	0.0283	−8.9579	0.0000*	3.7063
PctLowWa	−0.0423	0.0262	−1.6146	0.1065	3.1803
PctHiWaWrk	0.0523	0.0197	2.6611	0.0078*	1.7923
PctLowWaWrk	0.0608	0.0191	3.1869	0.0015*	1.6891
ResD	0.1121	0.0209	5.3603	0.0000*	2.0289
EmpD	−0.1659	0.0187	−8.8928	0.0000*	1.6126
EmpHHEnt	−0.0083	0.0281	−0.2958	0.7674	3.6555
TripEQ	0.1073	0.0193	5.5465	0.0000*	1.7331
JobPop	−0.0559	0.0233	−2.3999	0.0165	2.5157
AutoNetD	0.1351	0.0154	8.7931	0.0000*	1.0941
PedNetD	−0.2376	0.0302	−7.8623	0.0000*	4.2317
IntsecD	0.2525	0.0301	8.3819	0.0000*	4.2056
PctEmpHalfM	−0.0043	0.0188	−0.2286	0.8192	1.6322
EmpAcc	0.1557	0.0199	7.8028	0.0000*	1.8445
BusStopsD	0.0733	0.0195	3.7644	0.0002*	1.7567
Number of Observations	3471				
Multiple R-Squared	0.2557				
Adjusted R-Squared	0.2514				
AICc	8996				

adjusted R-squared stands at 0.4919, and the AICc value is 7991.31, indicating a relatively better fit than global models such as the OLS and SAR. Unlike SLR, MGWR offers local coefficient estimates. Fig. 6 provides a comprehensive summary of coefficient estimates and neighborhood statistics for the MGWR model. Columns like Mean, STD, Min, Median, and Max offer descriptive statistics for the coefficient estimates of each explanatory variable. These coefficients quantify the impact of each explanatory variable on the dependent variable. The 'Significant' column indicates the percentage of features where the coefficients of explanatory variables are statistically significant. For example, PctWrkAge has a significant coefficient across all 3471 CBGs, denoting a 100% significance rate. The coefficients for this variable range from 0.0591 to 0.0718, suggesting a consistently positive effect across all states. The 'Neighbors' column stands for the count of neighboring features influenced by each explanatory variable's coefficient, offering insight into the spatial extent of the effect. A higher percentage indicates a broader regional or global impact of the variable. For instance, PctWrkAge affects all neighborhoods (100% of features), suggesting a widespread influence. Conversely, PctWorker impacts about 62.14% of features, signifying a more localized effect.

In total, 11 variables are significant across all CBGs, including PctWrkAge, PctHiWa, PctHiWaWrk, EmpD, JobPop, AutoNetD, PedNetD, IntsecD, PctEmpHalfM, EmpAcc, and BusStopsD. Variables like PctWorker, PctLowWaWrk, TripEQ, and ResD also exhibit significant coefficients in most neighborhoods (94.73%, 99.94%, 71.28%, 68.97%, respectively). However, some variables, such as PctZeroCar and PctMultiCa, show coefficients with less statistical significance across neighborhoods.

The distribution of coefficients for some of the explanatory variables and neighborhoods, is depicted in Fig. 5. A detailed analysis of the variables is categorized and described under the following headings: socio-demographics, employment & housing diversity, land use density, roadway characteristics, and transit accessibility.

4.3.1. Socio-demographics

The analysis identifies a clear association between crash rates and socio-demographic factors. A higher proportion of the working-age population (PctWrkAge) is associated with increased pedestrian crash rates. The impact of PctWrkAge is not uniform across the state, showing an intensification from west to east, as reflected in the distribution coefficients, which range from 0.0591 to 0.0718. Similarly, the percentage of the employed population (PctWorker) also positively correlates with crash rates. This may be due to areas with higher employment levels having increased pedestrian activity, potentially leading to more crashes. The coefficient pattern for PctWorker significantly differs from that of PctWrkAge; the most substantial correlation is observed in the Baton Rouge area.

Table 3
Spatial lag and spatial error model results for crash rate analysis.

	Spatial Lag Model (SLM)				Spatial Error Model (SEM)			
Variable	Coefficient	S.T.D.	z-Score	p-value	Coefficient	S.T.D.	z-Score	p-value
Intercept	−0.0128	0.0127	−1.0133	0.3109	−0.025	0.034	−0.720	0.471
PctWrkAge	0.0713	0.0142	5.0328	0.0000*	0.068	0.014	4.841	0.000*
PctWorker	0.0671	0.0133	5.0499	0.0000*	0.054	0.013	4.210	0.000*
PctZeroCar	0.0898	0.0181	4.9606	0.0000*	0.065	0.019	3.482	0.000*
PctMultiCa	0.0269	0.0198	1.3550	0.1754	0.021	0.020	1.068	0.285
PctHiWa	−0.1606	0.0245	−6.5590	0.0000*	−0.172	0.028	−6.092	0.000*
PctLowWa	−0.0368	0.0226	−1.6265	0.1038	0.012	0.024	0.480	0.631
PctHiWaWrk	0.0500	0.0170	2.9508	0.0032*	0.038	0.017	2.236	0.025
PctLowWaWrk	0.0413	0.0165	2.5088	0.0121	0.021	0.016	1.325	0.185
ResD	−0.0116	0.0181	−0.6438	0.5197	−0.086	0.022	−3.848	0.000*
EmpD	−0.1482	0.0161	−9.2055	0.0000*	−0.149	0.018	−8.525	0.000*
EmpHHEnt	−0.0062	0.0242	−0.2576	0.7967	0.013	0.026	0.499	0.618
TripEQ	0.0997	0.0167	5.9776	0.0000*	0.094	0.016	5.719	0.000*
JobPop	−0.0481	0.0201	−2.3917	0.0168	−0.065	0.022	−2.993	0.003*
AutoNetD	0.0951	0.0133	7.1575	0.0000*	0.094	0.015	6.317	0.000*
PedNetD	−0.1285	0.0262	−4.9141	0.0000*	−0.110	0.033	−3.314	0.001*
IntsecD	0.1755	0.0261	6.7338	0.0000*	0.164	0.029	5.723	0.000*
PctEmpHalfM	−0.0173	0.0162	−1.0705	0.2844	−0.055	0.023	−2.421	0.015
EmpAcc	0.0736	0.0176	4.1782	0.0000*	0.196	0.031	6.241	0.000*
BusStopsD	0.0753	0.0168	4.4856	0.0000*	0.090	0.017	5.237	0.000*
W_CrashRate	0.5782	0.0188	30.7148	0.0000*	–	–	–	–
lambda	–	–	–	–	0.631	0.018	34.221	0.000*
Number of Observations			3471			3471		
Pseudo R-squared			0.4484			0.1944		
Spatial Pseudo R-squared			0.2729			–		
Log likelihood			−4023.155			−4049.280		
AICc			8090.309			8140.559		

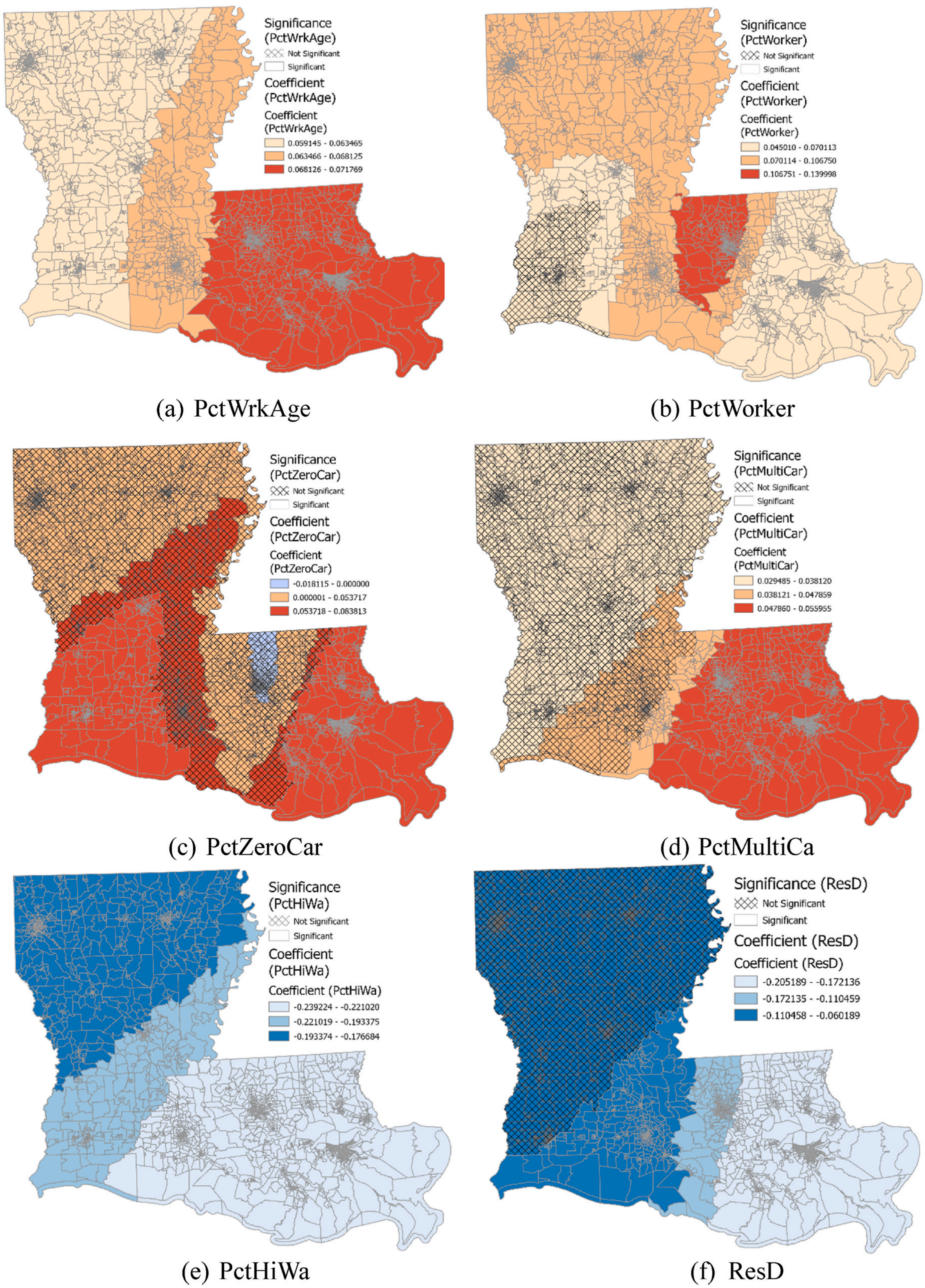


Fig. 5. Coefficient distribution of explanatory variables and neighborhoods for pedestrian crash rates.

Moreover, the percentage of the high-income population (PctHiWa) exhibits negative coefficients, suggesting the presence of better infrastructure or reduced pedestrian-vehicle interaction and, consequently, fewer crashes. This aligns with previous research, such as Mansfield et al. (2018), which reported a negative correlation between high-income CBGs and pedestrian crashes. No significant correlation is

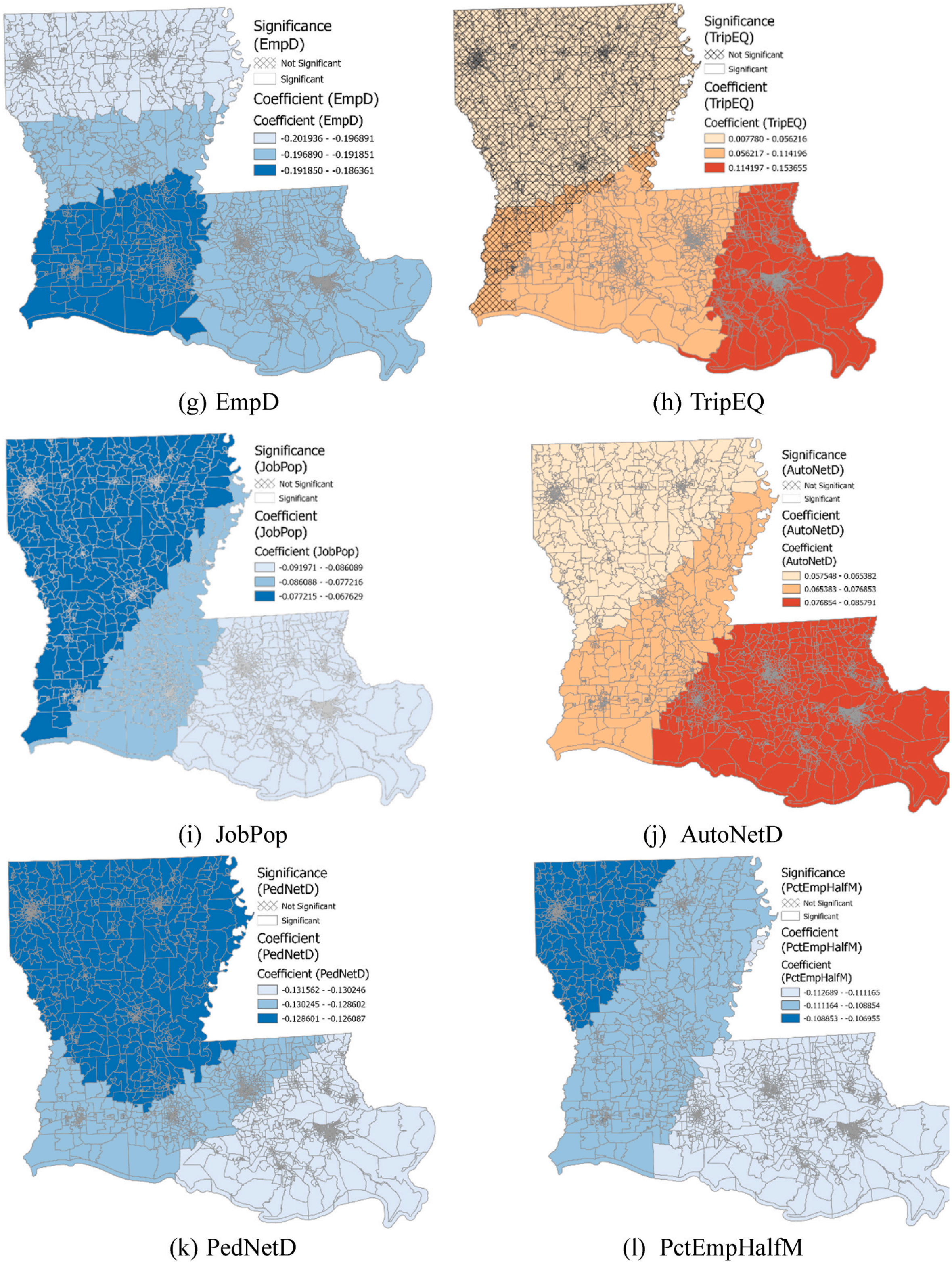


Fig. 5. (*continued*).

detected with the percentage of the low-wage population (PctLowWa). However, Tokey et al. (2023) have found that the percentage of low-wage population density has a positive association with pedestrian crashes.

Additionally, the results reveal a positive correlation between crash rates and the percentage of zero-car households (PctZeroCar), primarily in the southwestern and southeastern corners of Louisiana. This pattern could suggest that pedestrian traffic is higher in areas with lower car ownership, potentially raising the likelihood of pedestrian-related crashes. A positive correlation between the percentage of zero-car households and pedestrian crash rates was also observed at the census tract level in the study of Cottrill and Thakuriah (2010). Conversely, the percentage of multi-car households (PctMultiCar) shows a positive correlation with pedestrian crash rates in the southeastern region of

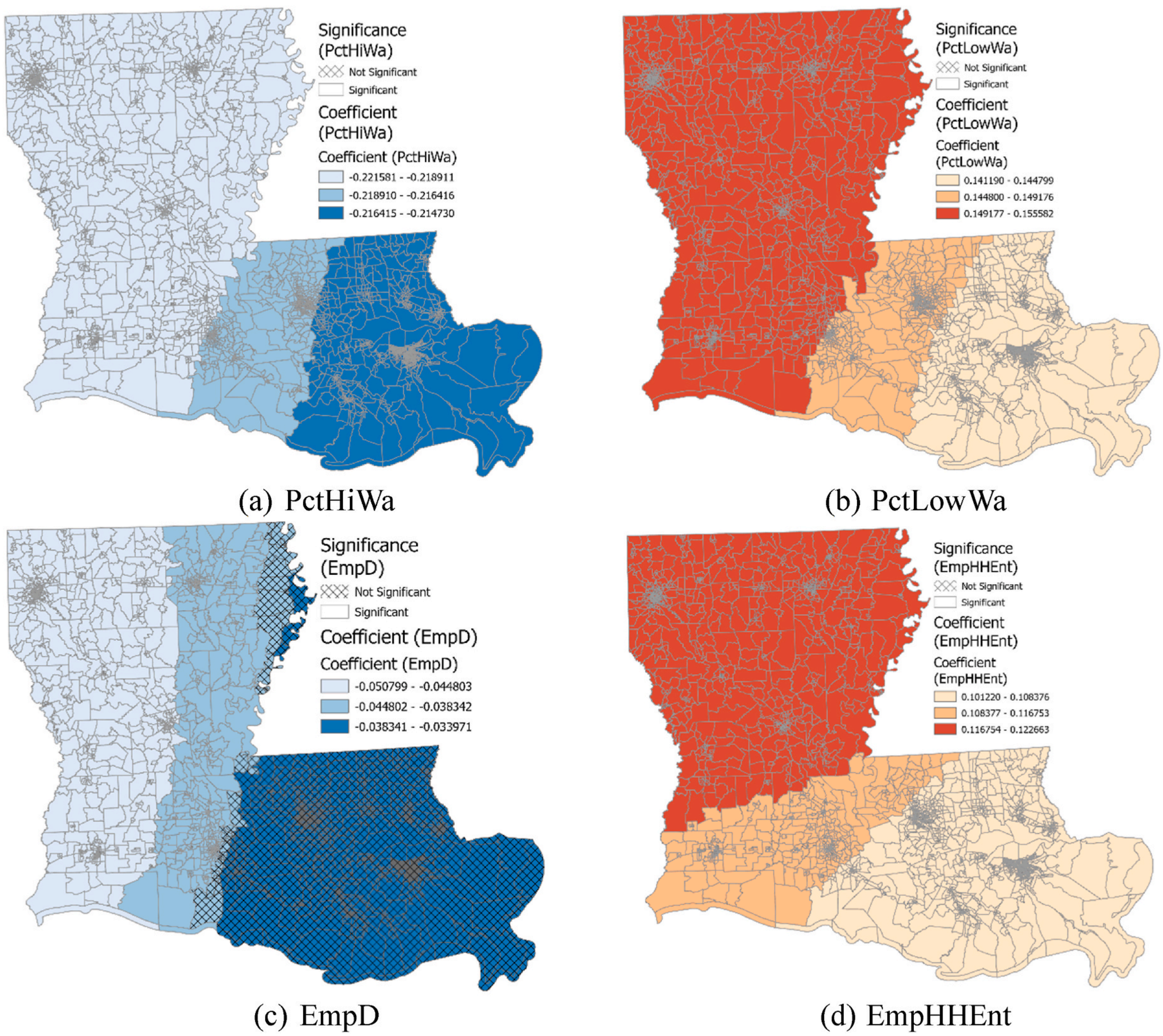


Fig. 6. Coefficient distribution of explanatory variables and neighborhoods for BLTR.

Louisiana. The increased number of cars may indicate heavier vehicular traffic, which, combined with pedestrian presence, could heighten the risk of crashes. Examining PctZeroCar and PctMultiCar together in the southeastern corner, the concurrent positive correlation of both PctZeroCar and PctMultiCar suggests that areas with a wide disparity in car ownership are associated with higher crash rates. This may also point to a more intricate interplay of pedestrian safety factors. This region's distinct socio-economic and infrastructural dynamics may influence pedestrian-vehicle interactions, indicating that a broader range of variables should be explored to thoroughly understand the factors influencing crash rates.

4.3.2. Employment & housing diversity

Regarding employment and housing diversity, regions with a higher proportion of both higher-salary (PctHiWaWrk) and lower-salary employment (PctLowWaWrk) compared to median salary employment show positive coefficients. This trend indicates that areas with a broader range of salaries among employees might see more pedestrian traffic, potentially escalating crash rates. Varied income levels could signify a mix of commercial and service activities, drawing a more diverse and extensive pedestrian crowd. Regarding land use entropy, while no studies have been found specifically linking it with employment entropy, Tokey et al. (2023) indicated that land use entropy positively associated with pedestrian crashes. On the other hand, the diversity of housing and employment (EmpHHEnt) does not show a significant correlation with crash rates.

Additionally, the balance of trip-making (TripEQ) and the regional diversity in population and total employment (JobPop) display positive and negative correlations with crash rates, respectively. These results underscore the intricacies of urban commuting patterns and their impact on pedestrian safety. A more balanced pattern of trip-making might suggest a better mix of land uses, potentially leading to more interaction points and a higher risk of crashes. Conversely, a higher regional diversity in population and employment could indicate better infrastructure and planning, promoting safer pedestrian and vehicular movements and reducing crash rates.

4.3.3. Land use density

The analysis reveals that land use characteristics, particularly residential and employment densities, have a significant association with pedestrian crash rates. Specifically, 68.97% of the CBGs demonstrate a significant association with residential density (ResD), indicating that areas with higher residential density tend to be associated with lower pedestrian crash rates. Furthermore, regions with high employment density (EmpD) also exhibit negative coefficients about pedestrian crash rates, suggesting that, compared to other types of land use, areas with high residential and employment density tend to have lower crash rates. This trend could be attributed to the concentration of workplaces in

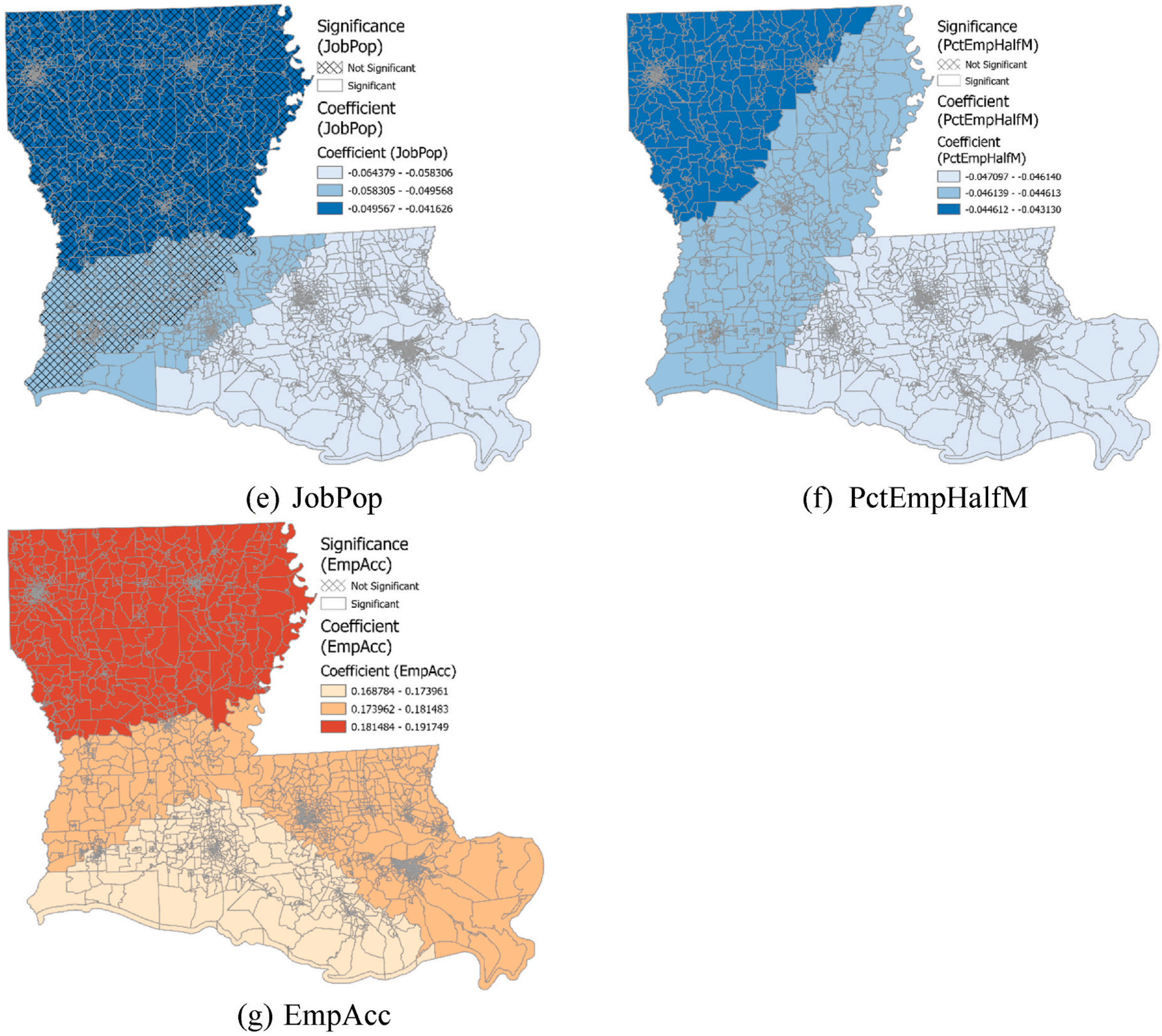


Fig. 6. (*continued*).

well-planned urban centers, where pedestrian infrastructure is more developed and traffic regulations are more strictly enforced. Areas with high employment density also promote a pedestrian-friendly culture, encouraging more individuals to walk or use public transit for their daily commute, thereby reducing the likelihood of pedestrian-vehicle conflicts. Teen crash frequency is notably higher in locations characterized by light commercial and industrial land uses (Mathew et al., 2022). Commercial land use is also significantly associated with high crash frequency (Tokey et al., 2023). According to the study by Gooch et al. (2022), higher employment density could lead to an increased risk of pedestrian crashes. However, it's also essential to consider that while high density in residential and employment areas correlates with lower pedestrian crash rates, the underlying mechanisms may involve a complex interplay of factors, including urban planning, infrastructure quality, public transportation availability, and community behavior. Further in-depth analysis would be required to disentangle these factors and fully understand the dynamics of pedestrian safety in high-density areas.

4.3.4. *Road network characteristics*

Regions with denser auto-oriented networks (AutoNetD) correlate with higher crash rates, possibly due to increased vehicular traffic and the potential for conflicts with pedestrians. High auto network density typically signifies more exposure to traffic, leading to a higher rate of crashes. These areas are often urban settings, where pedestrian traffic is more substantial. For instance, Gitelman et al. (2012) found that 75% of fatal pedestrian crashes and 95% of injury crashes occurred in urban areas. Similarly, Gårder (2004) highlighted that pedestrian risk significantly increases when crossing from high-speed to low-speed environments. Wang and Kockelman (2013) observed a positive correlation between the density of freeways and arterials and pedestrian crash frequency, a finding echoed by Wang et al. (2016) concerning the relationship between road density and pedestrian crashes.

Conversely, areas with denser pedestrian-oriented networks (PedNetD) show negative coefficients, implying these regions are safer for pedestrians, possibly due to superior infrastructure or urban planning that prioritizes pedestrian pathways. Wang and Kockelman (2013) also noted a negative correlation between sidewalk density and pedestrian crash occurrence. Moreover, a higher density of intersections (IntsecD) tends to increase crash rates, as more intersections introduce potential conflict points between pedestrians and vehicles, despite potentially enhancing pedestrian movement. Long and Ferenchak (2021) demonstrated that pedestrian fatality and severe injury hotspots correlate with lower sidewalk coverage. Likewise, Yin and Zhang (2021) reported a significant negative correlation between walkability and pedestrian safety.

4.3.5. *Transit accessibility*

Transit-related characteristics also exhibit a significant correlation with pedestrian crash rates. The higher proportion of employment within ¼ mile of a transit stop (PctEmpHalfM) is associated with lower crash rates, implying that transit-oriented developments may promote safer pedestrian environments or diminish reliance on vehicles. Conversely, regions with higher auto-oriented regional centrality (EmpAcc) and bus stop density (BusStopsD) correlate with increased crash rates. This observation aligns with findings by Tokey et al. (2023) and Wang and Kockelman (2013), identifying a positive relationship between bus stop density and pedestrian crash counts. This correlation could stem from elevated vehicular traffic and intensified interactions between pedestrians and vehicles in areas with dense public transit infrastructure, potentially heightening the risk of pedestrian-vehicle conflicts. Areas with higher auto regional centrality typically experience more vehicular traffic, increasing pedestrian exposure to crashes.

4.4. Exploring pedestrian safety disparity

4.4.1. MGWR modeling results for BLTR

To gain further insights into the relationship between these variables and the ratio of crashes involving the black population in total crashes, the MGWR model was employed (see Table 4). As shown in Table 5, the R-squared is 0.3808, and the AICc value is 8668.32, suggesting that the model's performance is not as robust as that of the crash rate model. Compared to the model analyzing total crash rates, there has been a decrease in the number of significant variables. Specifically, variables including PctWrkAge, PctWorker, PctZeroCar, PctMultiCa, PctHiWaWrk, PctLowWaWr, ResD, TripEQ, AutoNetD, PedNetD, IntsecD, and BusStopsD, which were significant in the previous model, do not maintain their significance at any CBGs in this model. This suggests that there is no significant correlation between these variables and the proportion of crashes involving the black population in total crashes. PctLowWa and EmpHHEnt emerge as significant variables, both displaying a positive association with the proportion of crashes involving the black population in total crashes (see Table 6).

The PctHiWa maintains a negative coefficient across the entire study area, aligning with the findings of the total crash rate model. Nevertheless, the distribution of coefficients contrasts with that in the total crash rate model. Locations that exhibited larger negative coefficients in the crash rate model present smaller negative coefficients in this model. Conversely, the PctLowWa coefficient, ranging from 0.1412 to 0.1556, is significant across all CBGs. This reflects a positive correlation between the percentage of the low-income population and the rate of pedestrian crashes involving the black population. Notably, PctLowWa does not demonstrate significance in the total crash rate model, highlighting a distinct trend. Low-income areas are disproportionately affected by pedestrian crashes involving the black population, a pattern not mirrored in the broader scope of pedestrian crash rates.

Employment density (EmpD), with values between −0.0508 and −0.034, is significant for 41.11% of the area. This contrasts with its universal significance in the total crash model, and the distribution of coefficients for this variable varies significantly. Employment and household entropy (EmpHHEnt), a variable not present in the total crash model, ranges from 0.1012 to 0.1227. This suggests a correlation between pedestrian crashes involving the black population and areas with higher levels of employment and household entropy. JobPop, with a range from −0.0644 to −0.0416, is significant for 64.07% of the area. The ranges for PctEmpHalfM (from −0.0471 to −0.0431) and EmpAcc (from 0.1688 to 0.1917) maintain a negative coefficient alignment with the total crash rate model, reinforcing the relationship observed in the broader dataset. Transit accessibility (PctEmpHalfM) is negatively associated with crash rates, whereas auto-oriented regional centrality (EmpAcc) correlates with increased crash rates. These observations highlight spatial variations in the interaction between certain factors and racial disparities in pedestrian crashes. Thorough analysis of these patterns is essential for devising targeted interventions to improve pedestrian safety and effectively address the identified racial disparities.

4.4.2. MGWR modeling results for BLKABR

The study also examined the correlation between various factors and the ratio of crashes involving the Black population in KAB-level crashes. The model's performance is comparable to that of the BLTR model. Variables such as PctHiWa, EmpD, and PedNetD exhibit negative coefficients, while PctLowWa, EmpHHEnt, EmpAcc, and BusStopsD show positive coefficients. EmpHHEnt is significant in 2.19% of the CBGs, and PedNetD is significant in 42.12% of the CBGs. The rest of the variables have significant coefficients across the entire study area. Additionally, only PedNetD demonstrates a local bandwidth (affecting 76.61% of

Table 4
Summary statistics for coefficients estimates and neighborhoods for pedestrian crashes rates.

Explanatory Variables	Mean	STD	Min	Median	Max	Neighbors (% of Features)[a]	Significant (% of Features)[b]
Intercept	0.0453	0.5997	−0.6223	−0.1065	5.3064	30 (0.86)	180 (5.19)
PctWrkAge	0.067	0.0042	0.0591	0.0695	0.0718	3471 (100.00)	3471 (100.00)
PctWorker	0.0746	0.0256	0.045	0.07	0.14	2157 (62.14)	3288 (94.73)
PctZeroCar	0.0535	0.0257	−0.0181	0.0669	0.0838	1965 (56.61)	1666 (48.00)
PctMultiCar	0.045	0.0092	0.0295	0.0495	0.056	3471 (100.00)	1991 (57.36)
PctHiWa	−0.2201	0.0237	−0.2392	−0.2357	−0.1767	3279 (94.47)	3471 (100.00)
PctLowWa	−0.0208	0.0073	−0.0357	−0.0162	−0.0149	3471 (100.00)	0 (0.00)
PctHiWaWrk	0.0516	0.0024	0.0451	0.0526	0.0549	3471 (100.00)	3471 (100.00)
PctLowWaWrk	0.048	0.0033	0.0399	0.0498	0.0511	3471 (100.00)	3469 (99.94)
ResD	−0.138	0.0583	−0.2052	−0.1364	−0.0602	2157 (62.14)	2394 (68.97)
EmpD	−0.1943	0.0032	−0.2019	−0.1944	−0.1864	3471 (100.00)	3471 (100.00)
EmpHHEnt	0.0055	0.0399	−0.0492	0.0271	0.0547	2157 (62.14)	0 (0.00)
TripEQ	0.0939	0.0478	0.0078	0.0881	0.1537	1846 (53.18)	2474 (71.28)
JobPop	−0.0833	0.0082	−0.092	−0.0879	−0.0676	3471 (100.00)	3471 (100.00)
AutoNetD	0.0765	0.0099	0.0575	0.0823	0.0858	3471 (100.00)	3471 (100.00)
PedNetD	−0.1295	0.0015	−0.1316	−0.1299	−0.1261	3471 (100.00)	3471 (100.00)
IntsecD	0.1833	0.0178	0.1579	0.1759	0.2256	2659 (76.61)	3471 (100.00)
PctEmpHalfM	−0.1112	0.0017	−0.1127	−0.1122	−0.107	3471 (100.00)	3471 (100.00)
EmpAcc	0.2091	0.0535	0.1435	0.1947	0.2835	1846 (53.18)	3471 (100.00)
BusStopsD	0.0744	0.0009	0.0735	0.074	0.0763	3471 (100.00)	3471 (100.00)
R-Squared		0.5514					
Adjusted R-Squared		0.4919					
AICc		7991.3068					

[a] : This number in the parenthesis ranges from 0 to 100%, and can be interpreted as a local, regional, and global scale based on the geographical context from low to high.

[b] : In the parentheses, the percentage of features that have significant coefficients of an explanatory variable.

Table 5
Summary statistics for coefficients estimates and neighborhoods for BLTR crashes.

Explanatory Variables	Mean	STD	Min	Median	Max	Neighbors (% of CBGs)	Significant (% of CBGs)
Intercept	0.0283	0.3069	−0.843	0.0414	0.9609	55 (1.58)	53 (1.53)
PctWrkAge	−0.0067	0.0065	−0.0183	−0.0034	0	3471 (100.00)	0 (0.00)
PctWorker	0.0097	0.0206	−0.0183	−0.0002	0.0512	2659 (76.61)	0 (0.00)
PctZeroCar	0.0335	0.0017	0.0289	0.0345	0.0349	3471 (100.00)	0 (0.00)
PctMultiCa	−0.0123	0.0025	−0.0206	−0.0119	−0.0064	3471 (100.00)	0 (0.00)
PctHiWa	−0.2178	0.0023	−0.2216	−0.2168	−0.2147	3471 (100.00)	3471 (100.00)
PctLowWa	0.1472	0.004	0.1412	0.1459	0.1556	3471 (100.00)	3471 (100.00)
PctHiWaWrk	−0.0323	0.0041	−0.0378	−0.0343	−0.0231	3471 (100.00)	0 (0.00)
PctLowWaWrk	0.0093	0.004	0.0038	0.0077	0.0161	3471 (100.00)	0 (0.00)
ResD	0.0172	0.0561	−0.0818	0.0024	0.1283	2038 (58.72)	0 (0.00)
EmpD	−0.0395	0.0056	−0.0508	−0.0359	−0.034	3471 (100.00)	1427 (41.11)
EmpHHEnt	0.1105	0.0076	0.1012	0.1074	0.1227	3471 (100.00)	3471 (100.00)
TripEQ	0.0064	0.0015	0.0015	0.0071	0.0106	3471 (100.00)	0 (0.00)
JobPop	−0.056	0.0077	−0.0644	−0.0596	−0.0416	3471 (100.00)	2224 (64.07)
AutoNetD	0.0082	0.0029	0.0012	0.008	0.0153	3471 (100.00)	0 (0.00)
PedNetD	−0.0459	0.0273	−0.079	−0.0562	0.0157	2540 (73.18)	0 (0.00)
IntsecD	0.0441	0.0211	0.0192	0.0456	0.0919	2467 (71.07)	0 (0.00)
PctEmpHalfM	−0.0461	0.0012	−0.0471	−0.0469	−0.0431	3471 (100.00)	3471 (100.00)
EmpAcc	0.1776	0.0058	0.1688	0.1749	0.1917	3471 (100.00)	3471 (100.00)
BusStopsD	0.0313	0.0021	0.0264	0.0326	0.0336	3471 (100.00)	0 (0.00)
R-Squared		0.3808					
Adjusted R-Squared		0.338					
AICc		8668.3153					

Table 6
Summary statistics for coefficients estimates for BLKABR crashes.

Explanatory Variables	Mean	STD	Min	Median	Max	Neighbors (% of Features)	Significant (% of Features)
Intercept	0.0218	0.3217	−0.8561	0.012	1.15	49 (1.41)	38 (1.09)
PctWrkAge	0.0097	0.004	0.0011	0.0119	0.0152	3471 (100.00)	0 (0.00)
PctWorker	0.017	0.0408	−0.0821	0.028	0.1072	1153 (33.22)	0 (0.00)
PctZeroCar	0.0135	0.0028	0.0102	0.0121	0.0181	3471 (100.00)	0 (0.00)
PctMultiCa	−0.0273	0.0031	−0.0332	−0.0264	−0.023	3471 (100.00)	0 (0.00)
PctHiWa	−0.2041	0.0026	−0.2085	−0.2024	−0.2013	3471 (100.00)	3471 (100.00)
PctLowWa	0.1327	0.0034	0.1281	0.1315	0.139	3471 (100.00)	3471 (100.00)
PctHiWaWrk	−0.0098	0.0026	−0.0158	−0.0093	−0.0062	3471 (100.00)	0 (0.00)
PctLowWaWrk	0.0085	0.0065	0.0016	0.005	0.0194	3471 (100.00)	0 (0.00)
ResD	−0.0322	0.0316	−0.0911	−0.0326	0.0252	2157 (62.14)	18 (0.52)
EmpD	−0.0481	0.0039	−0.057	−0.0455	−0.0445	3471 (100.00)	3471 (100.00)
EmpHHEnt	0.0616	0.0007	0.0599	0.0613	0.0649	3471 (100.00)	76 (2.19)
TripEQ	−0.0121	0.0057	−0.0219	−0.0091	−0.0048	3471 (100.00)	0 (0.00)
JobPop	−0.0251	0.0013	−0.0291	−0.0251	−0.0211	3471 (100.00)	0 (0.00)
AutoNetD	0.0303	0.0041	0.0231	0.029	0.0402	3471 (100.00)	0 (0.00)
PedNetD	−0.063	0.0334	−0.0989	−0.0809	−0.0059	2659 (76.61)	1462 (42.12)
IntsecD	0.0433	0.0032	0.0389	0.0423	0.0484	3471 (100.00)	0 (0.00)
PctEmpHalfM	−0.0305	0.0013	−0.0317	−0.0314	−0.0277	3471 (100.00)	0 (0.00)
EmpAcc	0.2133	0.0046	0.2025	0.2125	0.2254	3471 (100.00)	3471 (100.00)
BusStopsD	0.0515	0.0006	0.0499	0.0518	0.0527	3471 (100.00)	3471 (100.00)
R-Squared		0.3762					
Adjusted R-Squared		0.3268					
AICc		8763.6786					

features), indicating a more localized influence, whereas the other variables exhibit global bandwidth, signifying a broader, more uniform impact across the study area.

Fig. 7 presents the distribution of coefficients for explanatory variables and neighborhoods concerning the Black population in KAB-level crashes. The percentage of high-wage income workers (PctHiWa) in CBGs exhibits a negative coefficient, aligning with the findings of the previous crash rate model and model involving the black population. Moreover, the coefficient distribution mirrors that of the BLTR model. This indicates that areas with a higher percentage of high-income populations tend to experience fewer pedestrian crashes, fewer black-involved crashes, and fewer black-involved KAB-level pedestrian crashes. Conversely, the percentage of low-income workers (PctLowWa) consistently shows a significant positive coefficient across all CBGs. This relationship suggests that an increase in PctLowWa correlates with a higher frequency of KAB-level injuries among the black population, a finding also observed in the BLTR model. The association between lower incomes, higher poverty rates, and increased fatal injuries can be attributed to several interrelated factors.

Lower-income areas often have less investment in infrastructure, such as appropriate pedestrian crossings, street lighting, and traffic calming measures, which can contribute to higher rates of crashes and fatalities. Studies have shown that poor infrastructure and a lack of safety measures are linked to higher crash rates and fatalities (Morency et al., 2012; Noland and Quddus, 2004). Additionally, residents in higher poverty areas may have limited access to healthcare and emergency services, leading to a higher likelihood of fatal outcomes when injuries do occur. Furthermore, these areas may have higher levels of pedestrian and vehicle traffic due to a reliance on walking and public transportation, increasing exposure to potential crashes. Toran Pour et al. (2017) found that increases in the percentage of public transport usage in pedestrians' residential neighborhoods were linked to higher probabilities of fatal and severe injury crashes. Employment density (EmpD) exhibits a significant negative coefficient across CBGs,

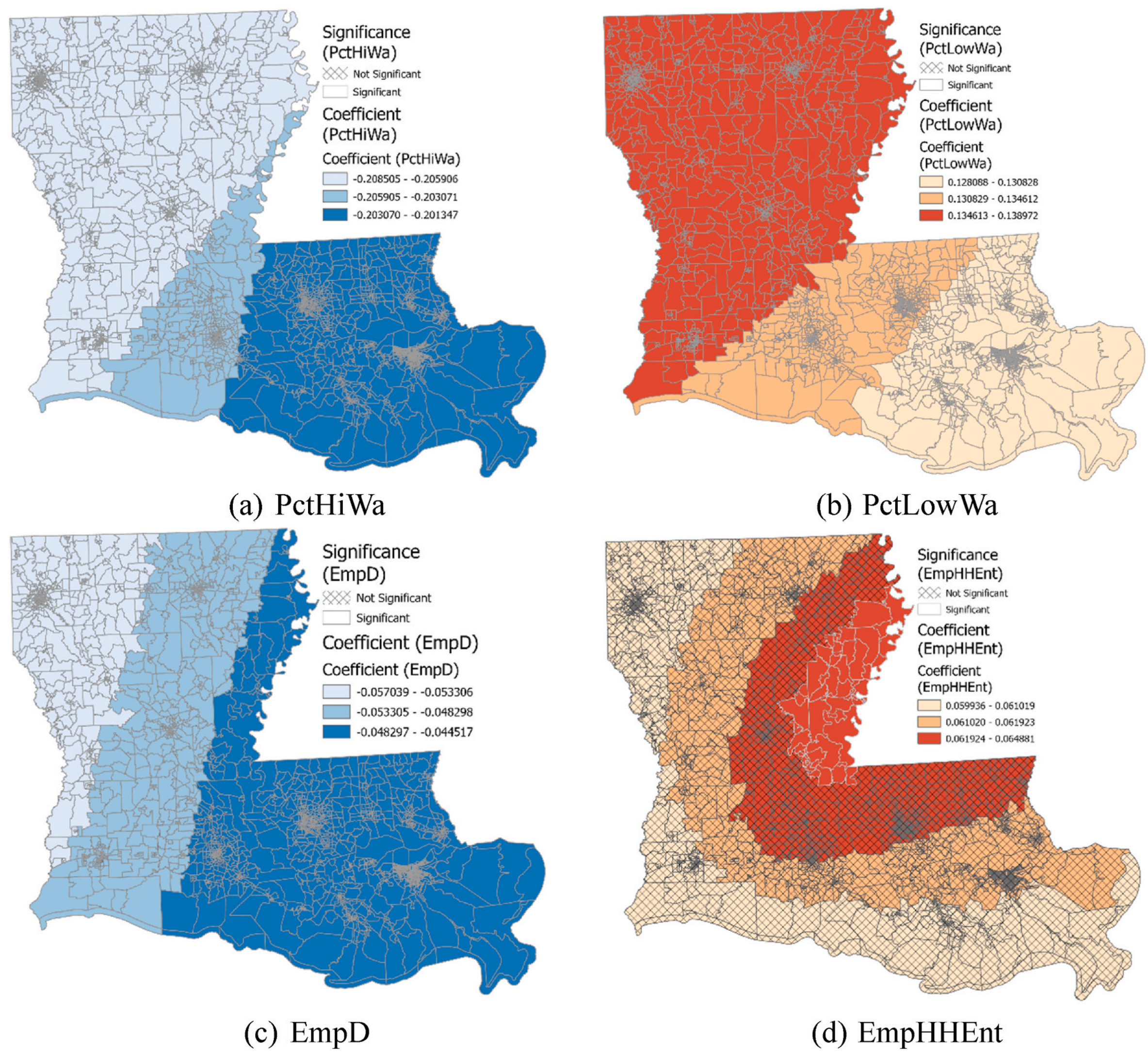


Fig. 7. Coefficient distribution of explanatory variables and neighborhoods for BLKABR.

indicating that areas with higher employment density tend to have lower KAB-level injuries among the black population. The coefficient for pedestrian-oriented network density (PedNetD) is significant for 42.12% of CBGs. While not significant in the BLTR model, it was significant in the crash rate model. In the state's northeast corner, the negative influence suggests that a denser network of pedestrian-oriented intersections correlates with a reduced likelihood of KAB-level injuries among the black population. The variations in associations likely stem from local-specific factors, such as intersection design, traffic patterns, or pedestrian behaviors.

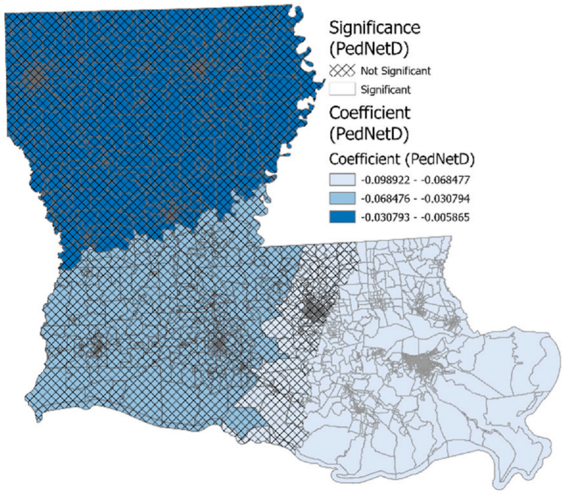
Regional centrality (EmpAcc) demonstrates a significant positive association across all three models, with consistent distribution patterns. Meanwhile, bus stop density (BusStopsD) shows a significant positive association in both the BLKABR model and the total pedestrian crash rate model but not in the BLTR model. This implies a correlation between denser bus stops and an increased total number of pedestrian crashes, as well as a higher ratio of KAB-level crashes involving the black population.

5. Conclusions

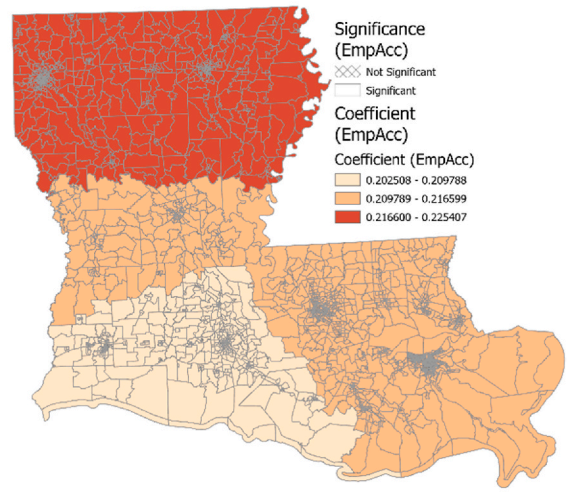
This study conducted a macro-level pedestrian safety analysis using data collected from multiple sources. The study adopted a spatial modeling framework. The analysis began with a global regression model, OLS, which provided initial insights into the relationships between variables and pedestrian crashes. However, spatial autocorrelation in the residuals indicated the need to consider the spatial structure of the data. Hence, spatial autoregressive and multiscale geographically weighted regression models were employed to capture the spatial variations in these relationships better. Moreover, Multiscale Geographically Weighted Regression was proved to deliver superior performance compared to other models and be able to provide local coefficients.

Exposure estimation is a critical question in pedestrian crash frequency studies. Population and total travel demand were collected and calculated as potential exposure variables. The results indicate that the relationship between total population and crash count is less significant than the connection between total travel demand and crash count. Therefore, total travel demand is adopted as a more reliable estimate for exposure.

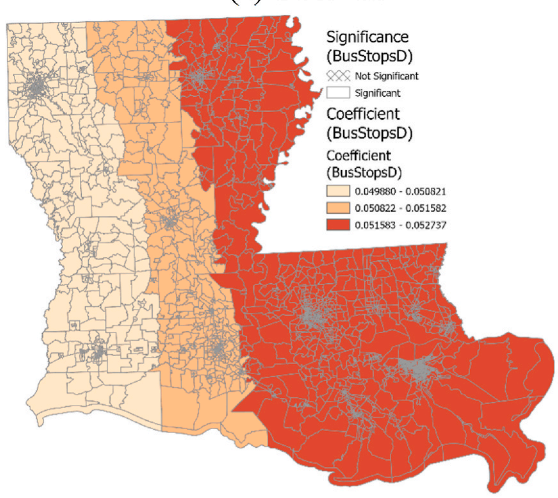
This study investigated the associations between socio-demographics, employment, housing diversity, land use diversity, road network characteristics, and transit accessibility at the CBG level in Louisiana with pedestrian crash rates. A higher proportion of the working-age population was found to be associated with increased pedestrian crash rates. In contrast, the percentage of the high-income population exhibits a negative association. No significant correlation



(e) PedNetD



(f) EmpAcc



(g) BusStopsD

Fig. 7. (*continued*).

was detected with the percentage of the low-wage population. Regarding employment and housing diversity, regions with a higher proportion of higher-salary and lower-salary jobs show positive coefficients compared to median-salary employment. The balance of trip-making and regional diversity in population and total employment display positive and negative correlations with crash rates, respectively.

Additionally, regions with denser auto-oriented networks correlate with higher crash rates, whereas denser pedestrian-oriented networks show negative associations. Higher transit accessibility for the employment population is associated with lower crash rates, and regions with higher auto-oriented regional centrality correlate with increased crash rates. Crash rates tend to be associated with higher bus stop density.

The study further examined the impact of various factors on the ratio of crashes involving the black population in total crashes and the ratio of crashes involving the black population in killed and severe injury crashes. The ratio of crashes involving the black population is negatively associated with the percentage of the high-income population across the entire area. It positively associates with the percentage of the low-income population, which shows no significant association with the total crash rate. Job and population equilibrium, and transit accessibility for the employment population, also showed negative associations. Employment and household entropy and regional centrality show positive associations with the black population in total crashes and the ratio of crashes involving the black population in severe crashes. Comparing the results to the black population in the total crashes model, the percentage of high-income and low-income populations maintained consistent negative and positive associations, respectively. However, pedestrian-oriented network density showed a negative association. Additionally, the density of bus stops showed positive associations, indicating that transit stops were associated with a higher likelihood of severe injuries for the black population.

Overall, the study highlights the importance of considering spatial variations and equity when analyzing pedestrian safety. This research contributes to a better understanding of the factors influencing pedestrian crashes and their associations with the black population's involvement in crashes. It underscores the significance of localized analyses to address spatial heterogeneity and promote equitable pedestrian safety measures across different regions in Louisiana. Key recommendations include focusing on developing and maintaining pedestrian-oriented infrastructure to reduce crash rates. This involves increasing the density of pedestrian networks and improving transit accessibility, which has been shown to correlate with lower crash rates.

Additionally, addressing socio-demographic disparities is essential. Efforts should be made to ensure equitable access to safe, pedestrian-friendly environments across all income levels, with particular attention to protecting vulnerable populations, such as the Black community, who are disproportionately affected by pedestrian crashes. Implementing targeted safety measures in regions with dense auto-oriented networks and high bus stop density could also mitigate the risk of crashes and severe injuries. Fostering a balanced regional development that promotes trip-making diversity and supports a mixed-use urban fabric can significantly reduce pedestrian crash rates and improve overall community well-being. The study has limitations. Future studies could explore additional demographics and more detailed infrastructure variables to provide a more comprehensive understanding of the factors influencing pedestrian crashes. Specifically, variables such as sidewalk width, signalized intersections, pedestrian crossing density, and the distribution of points of interest within census tracts can be included. Moreover, as aging is a significant trend, understanding how an aging population interacts with various infrastructural elements can help design safer environments for all age groups. Furthermore, future research could consider the Spatial Durbin Model to capture the spatial spillover effects more effectively and assess its potential advantages in explaining the spatial dynamics of pedestrian crashes.

CRediT authorship contribution statement

Jinli Liu: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation. **Subasish Das:** Writing – review & editing, Validation, Supervision, Resources, Data curation, Conceptualization. **F. Benjamin Zhan:** Writing – review & editing, Validation, Resources, Conceptualization. **Md Nasim Khan:** Writing – review & editing, Validation, Methodology, Formal analysis.

Data availability

Data will be made available on request.

References

Adeleke, R., Osayomi, T., Iyanda, A.E., 2020. Geographical patterns and effects of human and mechanical factors on road traffic crashes in Nigeria. Int. J. Inj. Control Saf. Promot. 28, 3–15. https://doi.org/10.1080/17457300.2020.1823996.

Aguero-Valverde, J., 2014. Direct spatial correlation in crash frequency models: estimation of the effective range. J. Transport. Saf. Secur. 6, 21–33. https://doi.org/10.1080/19439962.2013.799108.

Aguero-Valverde, J., Jovanis, P., 2008. Analysis of road crash frequency with spatial models. Transport. Res. Record 2061 55–63. https://doi.org/10.3141/2061-07.

Anselin, L., 2013. Spatial Econometrics: Methods and Models. Springer Science & Business Media.

Bo, H., 2018. Comprehensive Geographic Information Systems. Elsevier.

Burton, A.L., 2021. OLS (linear) regression. In: The Encyclopedia of Research Methods in Criminology and Criminal Justice. John Wiley & Sons, Ltd, pp. 509–514. https://doi.org/10.1002/9781119111931.ch104.

Cai, Q., Abdel-Aty, M., Lee, J., Huang, H., 2019. Integrating macro- and micro-level safety analyses: a Bayesian approach incorporating spatial interaction. Transportmetrica: Transport. Sci. 15, 285–306. https://doi.org/10.1080/23249935.2018.1471752.

Cai, Q., Lee, J., Eluru, N., Abdel-Aty, M., 2016. Macro-level pedestrian and bicycle crash analysis: incorporating spatial spillover effects in dual state count models. Accid. Anal. Prev. 93, 14–22. https://doi.org/10.1016/j.aap.2016.04.018.

CDC, 2023. Road traffic injuries and deaths—a global Problem [WWW Document]. Centers for Disease Control and Prevention. URL. https://www.cdc.gov/injury/features/global-road-safety/index.html, 2.7.24.

CDC, 2020. QuickStats. Morbidity and Mortality Weekly Report, Vol. 69, pp. 1434. [WWW Document] URL https://www.cdc.gov/mmwr/volumes/69/wr/mm6939a7.htm (Accessed 24 July 2023).

Cottrill, C.D., Thakuriah, P.Vonu, 2010. Evaluating pedestrian crashes in areas with high low-income or minority populations. Accid. Anal. Prev. 42, 1718–1728. https://doi.org/10.1016/j.aap.2010.04.012.

ESRI, 2023. Multiscale geographically weighted regression (MGWR) [WWW Document]. URL. https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/multiscale-geographically-weighted-regression.htm. (Accessed 23 July 2023).

Gálvez-Pérez, D., Guirao, B., Ortuño, A., 2024. Analysis of the elderly pedestrian traffic accidents in urban scenarios: the case of the Spanish municipalities. Int. J. Inj. Control Saf. Promot. 0, 1–20. https://doi.org/10.1080/17457300.2024.2335482.

Gálvez-Pérez, D., Guirao, B., Ortuño, A., 2023a. Analysis of the elderly pedestrian injury severity in urban traffic accidents in Spain using machine learning techniques. Transport. Res. Procedia, XV Conf. Transport Eng., CIT2023 71, 6–13. https://doi.org/10.1016/j.trpro.2023.11.051.

Gálvez-Pérez, D., Guirao, B., Ortuño, A., Picado-Santos, L., 2022. The influence of built environment factors on elderly pedestrian road safety in cities: the experience of madrid. Int. J. Environ. Res. Publ. Health 19, 2280. https://doi.org/10.3390/ijerph19042280.

Gálvez-Pérez, D., Guirao, B., Picado-Santos, L., Casado-Sanz, N., 2023b. The effect of built environment features on vehicle-pedestrian collisions in big cities: a comparison between Lisbon and Madrid. In: Transportation Research Procedia, TRA Lisbon 2022 Conference Proceedings Transport Research Arena (TRA Lisbon 2022), pp. 588–595. https://doi.org/10.1016/j.trpro.2023.11.443, 14th-17th November 2022, Lisboa, Portugal 72.

Gårder, P.E., 2004. The impact of speed and other variables on pedestrian safety in Maine. Accid. Anal. Prev. 36, 533–542.

Gaspard, G., Kim, D., Chun, Y., 2019. Residual spatial autocorrelation in macroecological and biogeographical modeling: a review. J. Ecol. Environ. 43, 19. https://doi.org/10.1186/s41610-019-0118-3.

Getis, A., Ord, J.K., 1992. The analysis of spatial association by use of distance statistics. Geogr. Anal. 24, 189–206.

Gitelman, V., Balasha, D., Carmel, R., Hendel, L., Pesahov, F., 2012. Characterization of pedestrian accidents and an examination of infrastructure measures to improve pedestrian safety in Israel. Accid. Anal. Prev. 44, 63–73.

Gomes, M.J.T.L., Cunto, F., da Silva, A.R., 2017. Geographically weighted negative binomial regression applied to zonal level safety performance models. Accid. Anal. Prev. 106, 254–261. https://doi.org/10.1016/j.aap.2017.06.011.

Gooch, J., Hamilton, I., Polin, B., Tanzen, R., Cohen, T., 2022. Systemic safety analysis of midblock pedestrian crashes in Massachusetts. Transport. Res. Rec. 2676, 722–730.

Guo, Q., Xu, P., Pei, X., Wong, S.C., Yao, D., 2017. The effect of road network patterns on pedestrian safety: a zone-based Bayesian spatial modeling approach. Accid. Anal. Prev. 99, 114–124. https://doi.org/10.1016/j.aap.2016.11.002.

Jin, S.T., Wang, L., Sui, D., 2023. How the built environment affects E-scooter sharing link flows: a machine learning approach. J. Transport Geogr. 112, 103687 https://doi.org/10.1016/j.jtrangeo.2023.103687.

Kim, D., 2019. The transportation safety of elderly pedestrians: modeling contributing factors to elderly pedestrian collisions. Accid. Anal. Prev. 131, 268–274. https://doi.org/10.1016/j.aap.2019.07.009.

Kim, S., Ulfarsson, G.F., 2019. Traffic safety in an aging society: analysis of older pedestrian crashes. J. Transport. Saf. Secur. 11, 323–332. https://doi.org/10.1080/19439962.2018.1430087.

Kock, N., Lynn, G., 2012. Lateral collinearity and misleading results in variance-based SEM: an illustration and recommendations. J. Assoc. Inf. Syst. Online 13. https://doi.org/10.17705/1jais.00302.

Lee, J., Abdel-Aty, M., Cai, Q., Wang, L., Huang, H., 2018. Integrated modeling approach for non-motorized mode trips and fatal crashes in the framework of transportation safety planning. Transport. Res. Rec. 2672, 49–60. https://doi.org/10.1177/0361198118772704.

Lee, J., Abdel-Aty, M., Huang, H., Cai, Q., 2019. Transportation safety planning approach for pedestrians: an integrated framework of modeling walking duration and pedestrian fatalities. Transport. Res. Rec. 2673, 898–906. https://doi.org/10.1177/0361198119837962.

Lee, J., Abdel-Aty, M., Jiang, X., 2015. Multivariate crash modeling for motor vehicle and non-motorized modes at the macroscopic level. Accid. Anal. Prev. 78, 146–154. https://doi.org/10.1016/j.aap.2015.03.003.

Li, D., Ranjitkar, P., Zhao, Y., Yi, H., Rashidi, S., 2017. Analyzing pedestrian crash injury severity under different weather conditions. Traffic Inj. Prev. 18, 427–430. https://doi.org/10.1080/15389588.2016.1207762.

Li, X., Yu, S., Huang, X., Dadashova, B., Cui, W., Zhang, Z., 2022. Do underserved and socially vulnerable communities observe more crashes? A spatial examination of social vulnerability and crash risks in Texas. Accid. Anal. Prev. 173, 106721 https://doi.org/10.1016/j.aap.2022.106721.

Li, Z., Wang, W., Liu, P., Bigham, J.M., Ragland, D.R., 2013. Using geographically weighted Poisson regression for county-level crash modeling in California. Saf. Sci. 58, 89–97. https://doi.org/10.1016/j.ssci.2013.04.005.

Ling, L., Zhang, W., Bao, J., Ukkusuri, S.V., 2023. Influencing factors for right turn lane crash frequency based on geographically and temporally weighted regression models. J. Saf. Res. 86, 191–208. https://doi.org/10.1016/j.jsr.2023.05.010.

Liu, J., Das, S., Khan, M.N., 2024. Decoding the impacts of contributory factors and addressing social disparities in crash frequency analysis. Accid. Anal. Prev. 194, 107375 https://doi.org/10.1016/j.aap.2023.107375.

Long, Jr.B., Ferenchak, N.N., 2021. Spatial equity analysis of nighttime pedestrian safety: role of land use and alcohol establishments in Albuquerque, NM. Transport. Res. Rec. 2675, 622–634.

Mansfield, T.J., Peck, D., Morgan, D., McCann, B., Teicher, P., 2018. The effects of roadway and built environment characteristics on pedestrian fatality risk: a national assessment at the neighborhood scale. Accid. Anal. Prev. 121, 166–176. https://doi.org/10.1016/j.aap.2018.06.018.

Mathew, S., Pulugurtha, S.S., Duvvuri, S., 2022. Exploring the effect of road network, demographic, and land use characteristics on teen crash frequency using geographically weighted negative binomial regression. Accid. Anal. Prev. 168, 106615 https://doi.org/10.1016/j.aap.2022.106615.

Merlin, L.A., Cherry, C.R., Mohamadi-Hezaveh, A., Dumbaugh, E., 2020. Residential accessibility's relationships with crash rates per capita. J. Transport and Land Use 13, 113–128.

Moeinaddini, M., Asadi-Shekari, Z., Zaly Shah, M., 2014. The relationship between urban street networks and the number of transport fatalities at the city level. Saf. Sci. 62, 114–120. https://doi.org/10.1016/j.ssci.2013.08.015.

Mokhtarimousavi, S., Anderson, J.C., Azizinamini, A., Hadi, M., 2020. Factors affecting injury severity in vehicle-pedestrian crashes: a day-of-week analysis using random parameter ordered response models and artificial neural networks. Int. J. Transport. Sci. Technol. 9, 100–115. https://doi.org/10.1016/j.ijtst.2020.01.001.

Morency, P., Gauvin, L., Plante, C., Fournier, M., Morency, C., 2012. Neighborhood social inequalities in road traffic injuries: the influence of traffic volume and road design. Am. J. Publ. Health 102, 1112–1119. https://doi.org/10.2105/AJPH.2011.300528.

NHTSA, 2021. Traffic safety fact (2021 data): pedestrian (No. DOT HS 813 458). National Highway Traffic Safety Admin. [WWW Document]. Washington, DC. URL https://crashstats.nhtsa.dot.gov/#/ (Accessed 24 July 2023).

Noland, R.B., Quddus, M.A., 2004. A spatially disaggregate analysis of road casualties in England. Accid. Anal. Prev. 36, 973–984. https://doi.org/10.1016/j.aap.2003.11.001.

Osama, A., Sayed, T., 2017. Evaluating the impact of connectivity, continuity, and topography of sidewalk network on pedestrian safety. Accid. Anal. Prev. 107, 117–125. https://doi.org/10.1016/j.aap.2017.08.001.

Patwary, A.L., Haque, A.M., Mahdinia, I., Khattak, A.J., 2024. Investigating transportation safety in disadvantaged communities by integrating crash and Environmental Justice data. Accid. Anal. Prev. 194, 107366 https://doi.org/10.1016/j.aap.2023.107366.

Pfiester, L.M., Thompson, R.G., Zhang, L., 2021. Spatiotemporal exploration of Melbourne pedestrian demand. J. Transport Geogr. 95, 103151 https://doi.org/10.1016/j.jtrangeo.2021.103151.

Pljakić, M., Jovanović, D., Matović, B., Mićić, S., 2019. Macro-level accident modeling in Novi Sad: a spatial regression approach. Accid. Anal. Prev. 132, 105259 https://doi.org/10.1016/j.aap.2019.105259.

Quddus, M.A., 2008. Modelling area-wide count outcomes with spatial correlation and heterogeneity: an analysis of London crash data. Accid. Anal. Prev. 40, 1486–1497. https://doi.org/10.1016/j.aap.2008.03.009.

Rahman, M.S., Abdel-Aty, M., Hasan, S., Cai, Q., 2019. Applying machine learning approaches to analyze the vulnerable road-users' crashes at statewide traffic analysis zones. J. Saf. Res. 70, 275–288. https://doi.org/10.1016/j.jsr.2019.04.008.

Roll, J., McNeil, N., 2022a. Race and income disparities in pedestrian injuries: factors influencing pedestrian safety inequity. Transport. Res. Transport Environ. 107, 103294.

Roll, J., McNeil, N., 2022b. Race and income disparities in pedestrian injuries: factors influencing pedestrian safety inequity. Transport. Res. Transport Environ. 107, 103294.

Sanders, R.L., Schneider, R.J., 2022. An exploration of pedestrian fatalities by race in the United States. Transport. Res. Transport Environ. 107, 103298 https://doi.org/10.1016/j.trd.2022.103298.

Schneider, R.J., 2020. United States pedestrian fatality trends, 1977 to 2016. Transport. Res. Rec. 2674, 1069–1083. https://doi.org/10.1177/0361198120933636.

Siddiqui, S., Al-Kaisy, A., 2017. Transportation research board. Effect of Advisory Variable Speed Limit on the Fundamental Flow Diagrams along Urban Freeway, p. 20p.

Stipancic, J., Miranda-Moreno, L., Strauss, J., Labbe, A., 2020. Pedestrian safety at signalized intersections: modelling spatial effects of exposure, geometry and signalization on a large urban network. Accid. Anal. Prev. 134, 105265 https://doi.org/10.1016/j.aap.2019.105265.

Stoker, P., Garfinkel-Castro, A., Khayesi, M., Odero, W., Mwangi, M.N., Peden, M., Ewing, R., 2015. Pedestrian safety and the built environment: a review of the risk factors. J. Plann. Lit. 30, 377–392.

Tasic, I., Elvik, R., Brewer, S., 2017. Exploring the safety in numbers effect for vulnerable road users on a macroscopic scale. Accid. Anal. Prev. 109, 36–46. https://doi.org/10.1016/j.aap.2017.07.029.

Tokey, A.I., Shioma, S.A., Uddin, M.S., 2023. Assessing the effectiveness of built environment-based safety measures in urban and rural areas for reducing the non-motorist crashes. Heliyon 9, e14076. https://doi.org/10.1016/j.heliyon.2023.e14076.

Toran Pour, A., Moridpour, S., Tay, R., Rajabifard, A., 2017. Neighborhood influences on vehicle-pedestrian crash severity. J. Urban Health 94, 855–868. https://doi.org/10.1007/s11524-017-0200-z.

U.S. DOT, 2024. U.S. DOT Launches $15 Million Complete Streets Artificial Intelligence Initiative for Small Businesses. US Department of Transportation [WWW Document]. URL. https://www.transportation.gov/briefing-room/us-dot-launches-15-million-complete-streets-artificial-intelligence-initiative-small, 2.6.24.

US EPA, O, 2014. Smart location mapping [WWW Document]. URL. https://www.epa.gov/smartgrowth/smart-location-mapping. (Accessed 23 July 2023).

Wang, X., Yang, J., Lee, C., Ji, Z., You, S., 2016. Macro-level safety analysis of pedestrian crashes in Shanghai, China. Accid. Anal. Prev. 96, 12–21. https://doi.org/10.1016/j.aap.2016.07.028.

Wang, Y., Kockelman, K.M., 2013. A Poisson-lognormal conditional-autoregressive model for multivariate spatial analysis of pedestrian crash counts across neighborhoods. Accid. Anal. Prev. 60, 71–84. https://doi.org/10.1016/j.aap.2013.07.030.

WHO, 2023. Road safety. Road Traffic injuries [WWW Document]. https://www.who.int/health-topics/road-safety. (Accessed 24 July 2023).

Wu, D., Zhang, Y., Xiang, Q., 2024. Geographically weighted random forests for macro-level crash frequency prediction. Accid. Anal. Prev. 194, 107370 https://doi.org/10.1016/j.aap.2023.107370.

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, 1459–1476. https://doi.org/10.1111/risa.12785.

Xu, P., Huang, H., 2015. Modeling crash spatial heterogeneity: random parameter versus geographically weighting. Accid. Anal. Prev. 75, 16–25. https://doi.org/10.1016/j.aap.2014.10.020.

Yin, L., Zhang, H., 2021. Building walkable and safe neighborhoods: assessing the built environment characteristics for pedestrian safety in Buffalo, NY. J. Transport Health 22, 101129.

Zafri, N.M., Khan, A., 2022. A spatial regression modeling framework for examining relationships between the built environment and pedestrian crash occurrences at macroscopic level: a study in a developing country context. Geogr. Sustain. 3, 312–324.

Zhai, X., Huang, H., Sze, N.N., Song, Z., Hon, K.K., 2019. Diagnostic analysis of the effects of weather condition on pedestrian crash severity. Accid. Anal. Prev. 122, 318–324. https://doi.org/10.1016/j.aap.2018.10.017.

Ziakopoulos, A., Yannis, G., 2020. A review of spatial approaches in road safety. Accid. Anal. Prev. 135, 105323 https://doi.org/10.1016/j.aap.2019.105323.