IJTE, Vol. 12, No. 1, Summer 2024, pp. 1741-1758
Research Paper

The Influence of Traffic, Land Use and Context Variables on Urban Crash Types (Case study: Shiraz Metropolis)

Mohammad Sedigh Bavar[1], Ali Naderan[2,*], Mahmoud Saffarzadeh[3]

Received: 2023/02/07 *Accepted: 2023/06/06*

Abstract

Built environmental factors are one of the most important causes of urban accidents. Studies have shown that in addition to accident data, which have spatial heterogeneity, factors influencing accidents also have spatial correlation. The occurrence of urban accidents depends on many human and environmental factors, so identifying the important factors influencing accidents and their spatial effects on each other is of great importance. The main goal of this study is to evaluate the spatial effects of environmental factors on the frequency of accidents in the city of Shiraz, Iran at the TAZ level. In the first step of the study, using component analysis models, important environmental factors affecting the accident were identified and composite indicators were produced as independent variables. In the second step, in order to control the effect of correlation and heterogeneity of model variables, spatial statistical models based on Euclidean distance such as geographically weighted Poisson regression (GWPR), geographically weighted negative binomial distribution (GWNBR) as well as Poisson and distribution models Negative binomial based on neighbor distance is used in spatial Bayes method with INLA approach. The results of the study showed that models based on distance and contiguity in order to evaluate the spatial effects of accident data and the factors affecting it at the TAZ level have higher accuracy than geographic weighted regression models, as well as indicators of land use diversity and access to the system. The public transport produced in the first step is effective in increasing the frequency of accidents, and in TAZs where this index is high, there is a higher probability of an accident. The results of this study can be important for city managers and planners in order to improve inner city safety measures as well as development planning and future city measures.

Keywords: Land Use, Environmental Factors, Urban Crash, Spatial effect based on neighbourhood, Spatial effect based on Sistance Matrix

[*] Corresponding author. E-mail: naderan@srbiau.ac.ir

[1] Ph.D. Candidate, Department of Civil Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

[2] Assistant Professor, Department of Civil Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

[3] Professor, Department of Civil and Environmental Engineering, Tarbiat Modares University, Tehran, Iran

1. Introduction

Very high population density in urban areas, rapid population growth, unplanned land use, road network and transportation system, heterogeneous traffic movement are among the major challenges of cities in developing countries, which have an impact on the occurrence of urban accidents (Cervero, 2013; Zafri & Khan, 2022; Zafri et al., 2020). In order to better understand and identify EIFs of urban accidents, in addition to geometric, demographic, traffic and population characteristics (Almasi & Behnood, 2022), indicators such as land use mix (Cervero & Murakami, 2009; Fuentes, Truffello, & Flores, 2022) ; Kang, 2018; Musa & Moses, 2014; Rodríguez Vignoli, 2001; Ruiz-Tagle & López, 2014; Sung, Lee, Cheon, & Yoon, 2022; Ziccardi, 2000) access to the transportation system (Kim, Pant, & Yamashita, 2010; Lake & Ferreira, 2002; Mavoa, Witten, McCreanor, & O'sullivan, 2012; Tiwari & Jain, 2012; Zhu & Liu, 2004) access to jobs (T. Chen, Sze, Chen, Labi, & Zeng, 2021; Raicu, Costescu, Raicu, & Popa, 2016) has been given the attention of researchers separately and with a thoughtful perspective.

In order to identify the impact of each of the EIFs on urban accidents in most past studies, non-spatial global methods with the assumption of constant coefficients of independent variables such as OLS (Tarko, 2018, 2023), multivariate linear regression (Rashidi, Keshavarz, Pazari, Safahieh, & Samimi, 2022), Poisson and negative binomial distribution (Ma et al., 2017,Moomen, Rezapour, Raja, & Ksaibati, 2020; H. Tang & Donnell, 2019), neural network based models and other models predictors (Huang, Zeng, Pei, Wong, & Xu, 2016; H. Tang & Donnell, 2019; Zarei, Hellinga, & Izadpanah, 2022) have been used. It has been shown in several studies that spatial data of accidents have spatial correlation and heterogeneity and the spatial effect between model variables (spatial autocorrelation) is effective in predicting the frequency of accidents (Almasi & Behnood, 2022; C. Liu & Sharma, 2017; Soroori, Mohammadzadeh Moghaddam, & Salehi, 2021; J. Tang, Zhao, Liu, Hao, & Gao, 2022) Therefore, in order to overcome the spatial correlation between model variables, spatial statistics methods based on weighted matrix (SW[1]) and Euclidean distance such as GIS-based methods(Al-Aamri, Hornby, Zhang, Al-Maniri, & Padmadas, 2021; J. Liu, Hainen, Li, Nie, & Nambisan, 2019).

In a study to investigate the effect of human activities in the spatial body of the city, in the occurrence of urban accidents, the GWPR method was used, and the results showed that collecting data from different sources is effective in the occurrence of urban accidents (Bao et al. ., 2021) In addition, the study (Sobreira & Cunto, 2021) showed the use of the time cycle of data collection during the day, such as the speed and volume of vehicles at different hours of the day, as well as the capacity of entry and exit approaches to intersections (Intini , Berloco, Fonzone, Fountas, & Ranieri, 2020). Compared to data averaging, it is more effective in identifying the time and place of accidents in the city. In order to eliminate the spatial effect between independent and dependent variables in spatial data, the study (Zafri & Khan, 2022) has shown that the variation of spatial scale in the entire urban network compared to a fixed spatial scale can reduce the spatial effect of independent variables with higher accuracy. Adjust In the study (Almasi & Behnood, 2022) which was conducted in order to identify the variables exposed to pedestrians in Tehran city, it is shown that among the spatial statistics models based on the weight matrix and Euclidean distance, the GWZIPR and GWZINBR models are considered Zero-inflated TAZs estimate expected pedestrian

[1] Spatial Weighted

accident frequencies more reliably than other spatial statistics models. Other studies have shown that predicting the frequency of inner-city accidents using the neighborhood-based spatial statistics method with the Integrated Nested Laplace Approximation (INLA[1]) approach, according to the nature of the data and how to adjust the spatial effects, can provide acceptable results, for example In several studies predicting the frequency of inner-city accidents based on travel behavior and mobile GPS tools (Stipancic et al., 2018; Stipancic, Miranda-Moreno, Saunier, & Labbe, 2019), the volume of entry at intersections of 3 controlled branches and intersections of 4 The controlled branch (Galgamuwa et al., 2021; Wang, Ivan, Ravishanker, & Jackson, 2017) was estimated using a latent Gaussian spatial model using the INLA method and the results showed that the spatial neighborhood negative binomial distribution model (SANB[2]) And Spatial Adjacent Poisson (SAP[3]) with INLA approach has higher accuracy than other models. One of the important methods of identifying the influence of independent variables in the estimation of dependent variables in terms of classification and size of data is the GLM and Tobit methods, which have been used in past studies in order to identify influential variables in the occurrence of macro accidents (Almasi et al. , 2021; Lee, Abdel-Aty, & Shah, 2019). The aim of this study is to investigate the dimensions of EIFs of urban accidents using the spatial Bayes method based on Euclidean distance and contiguity. In order to achieve this goal, the research gaps of past studies have been considered as follows.

1) Examining the relationship between the frequency of EIFs accidents of the first category variables (observed environmental factors) and the second category (indices obtained by the PCA[4] method) in an integrated manner

2) Identification of the best spatial model to estimate the frequency of accidents within the city using two approaches a) Poisson regression model and geographic weighted negative binomial distribution using Euclidean distance with empirical Bayes approach b) Poisson regression model and spatial negative binomial distribution Using the neighborhood distance with the Bayesian model and the INLA approach.

In none of the previous studies, the effect of group EIFs in estimating the frequency of inner-city accidents, as well as the comparison between spatial statistics models based on geographic weighted matrix and based on neighborhood distance with Bayesian approach, so it is necessary to investigate in urban accidents. that the effect of spatial correlation between environmental factors and the frequency of accidents can be controlled by which of the approaches of spatial statistics. This study aims to use this framework to model urban macroscopic accidents in order to find a suitable modeling technique, which can produce more reliable and accurate results.

2. Scope of Study and Research Data

2.1. The Scope of the Study

The city of Shiraz is geographically located in the southwest of Iran and in the central part of Fars province. The population of Shiraz city is more than 1.5 million people according to the latest census of 1395 of Iran Statistics Center. This city is located in the geographical position (29.7323°N, 52.6351°E). Its height is 1590 meters above sea level. The area of this city is 12,990 square kilometers, the length of which is 90 kilometers and the width varies from 20 to 30 kilometers (Figure 1).

2-2- Structure of variables and its sources

[1] Integrated Nested Laplace Approximation

[2] Spatial Adjacency Negative binomial

[3] Spatial Adjacency Poison

[4] Principal Component Analysis

The data of this study has been prepared in the form of GIS-Base in the study area. In this study, EIFs consist of two categories. The first category includes environmental observational variables and the second category includes several composite indicators such as user diversity index, job access index, and public transportation system access index, which are obtained from relations 1 and 2.

In the Kallangar analysis of intra-city accident data in this study, the total frequency of accidents at the level of the traffic areas of the entire city of Shiraz has been used as a dependent variable. (Almasi & Behnood, 2022; P. Chen & Zhou, 2016; Lee et al., 2019). Geographically dependent variable information and GIS-Base have been received from Shiraz Transport and Traffic Department (TATD[1]), which is in charge of collecting urban accident data. These data are related to the period of 2018 to 2019, where a total of 34,588 accidents (fatal-injury and damage) were reported. The distribution of accidents in the studied area is shown in Figure 1.

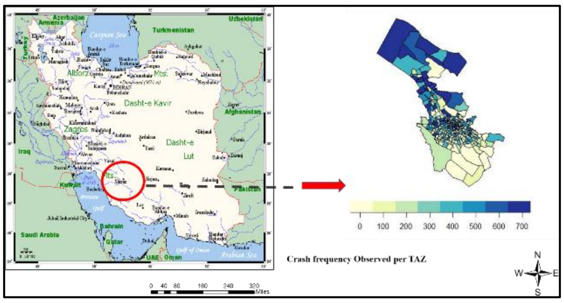

Figure 1. The study area

Based on the proposal of studies (Musa & Moses, 2014; Peera, Shekhawat, & Prasad, 2019; Sung et al., 2022; WEDAGAMA, Roger, & Dissanayake, 2008), the most important EIFs of urban accidents in the first category include land use (Sung et al. ., 2022; Umair et al., 2022) population (F. Guo & Lu, 2022) traffic characteristics such as average speed, average lane width, average number of lanes, number of bus stops (Tagar & Pulugurtha, 2021) . Combined indicators affecting urban accidents such as land use mix (Sung et al., 2022), access to public transportation and access to jobs (Su & Sze, 2022) and the frequency of urban accidents in estimating the frequency of urban accidents are less It has been studied by researchers. Calculating the indicators affecting the occurrence of an accident requires additional information such as the distance or the cost of traveling between areas, which may not be available to researchers, to see the necessary details of the indicators, refer to the relationship (1 and 2). In this study, the information layers of independent variables are based on TAZ (Bao et al., 2021), which includes urban environmental and demographic factors, including land uses and the population of Shiraz city, which was prepared from the vice city of Shiraz city planning. The descriptive information of the data of this study is shown in Table 1. The city of Shiraz has 325 traffic zones and the variety of land uses in them has been divided into 37 categories based on studies of the comprehensive transportation plan, which in order to reduce autocorrelation between variables (Yan et al., 2021) into 13 categories including: Road network, number of residential units, number of commercial units, cultural and religious, sports, educational, healthcare, green space and fields, administrative headquarters, industries and facilities and transportation, barren and dilapidated, mixed residential and non-residential, other uses were categorized and used to find the index of user diversity. In order to control the correlation between independent variables, Pearson's test was used (Rahmani et al., 2022). Therefore, in Pearson's correlation test, only the variables that are directly included in the model were used, Table 2.

3. Methodology

In this study, firstly, the variables of the second category of accident EIFs, such as the

[1] Transport and Traffic Deputy

diversity index of land use type (it is a function of 13 land use categories), the access to jobs index (it is a function of the distance between areas and the density of commercial and administrative use), and the index of access to the transportation system (it is a function of the number of bus stops and the distance between the areas) is calculated for each of the TAZs, and after the PCA review described in paragraph 3-3, they are entered into the model along with the first group of accident EIFs. In the following, by using 4 GWPR and GWNBR spatial statistics models based on the geographic distance matrix with variable coefficients for each TAZ, as well as SAP and SANB models, with the spatial Bayes model approach and the INLA method, the frequency of urban accidents is estimated.

3.1. User Diversity Index

Inferring from the study method (Delclòs-Alió & Miralles-Guasch, 2018; Fuentes et al., 2022), the land use diversity index is the total ratio of the area of each land use multiplied by its natural logarithm; It is calculated by dividing by the total area of all uses, which is called the entropy[1] index. Then the obtained value is scaled between 0 and 1, so that 1 is complete land use diversity and 0 is land use diversity or only one type of land use.

$$LUI\,(Entropy-Index)=\sum_{i=1}^{n}p_i\,\log(p_i)/\log(N) \qquad (1)$$

In relation 1. LUI, the entropy index of land use diversity; the percentage of used area i in TAZi; N is the sum of the total areas.

3.2. Index of Access to Jobs and Public Transportation System

By inference from the study method (Papa, Silva, Te Brömmelstroet, & Hull, 2016), these indices are obtained from equation (2).

$$PTI=\sum_{j=1}^{n}O_j f\,(c_{ij})/O \qquad (2)$$

In relation 2. The amount of opportunities available in the destinations (including the total area of non-residential uses in each TAZ) and the travel cost (in this study, the distance between the centers of the TAZ) between the origin and destination areas are based on the function.

3.3. Identification of Important Variables in the Model using Principal Component Analysis (PCA) Method

By extracting the most important information from the data table, principal component analysis (PCA) reduces the size of the data set, analyzes the structure of observations and variables, and then computes principal components, which are linear combinations of variables. As a result of the main variables (Lee et al., 2019), independent variables can be examined and selected, as well as the effect of indicators of user type, employment access, and transportation access, as well as other variables in the model using the GLM method and Tobit.

3.3.1. Generalized Linear Model

General linear models (GLM) are a general class of statistical models that include many common models with specific properties. The GLM equation is as follows:

$$Y=\sum_{i=1}^{m}\beta_i\chi_i+\varepsilon_i \qquad (3)$$

In the generalized linear model, the assumptions of the independent and normal distribution are given in relation to y. This distribution includes cases such as normal, Poisson, gamma, and binomial distribution (Olsson, 2002). Y is the linear prediction, and εi is the error parameter in this relation.

3.3.2. Tobit Model

In this research, the Tobit model was used to identify the confrontation criteria to remove any negative prediction of a pedestrian accident. The Tobit model is a statistical model used to describe the relationship between a non-negative dependent variable yi and an independent variable (or vector) xi.

$$y_i^*=\beta\chi_i+\varepsilon_i \quad i=1,2,......N \qquad (4)$$

[1] Entropy Index

$$y_i = \begin{cases} y_i^* & if \quad y_i^* > 0 \\ 0 & if \quad y_i^* \le 0 \end{cases} \qquad (5)$$

In this relation, yi * is a hidden variable that is observed only if it is positive. Also, N is the number of observations, yi is the dependent variable, xi is a vector of explanatory variables, β is a vector of estimable parameters, and εi is a normal and independent distribution. The error parameter also has zero mean and σ2 variance (Lord & Mannering, 2010).

3.4. Spatial Models based on Geographic Weighted Matrix

Spatial statistical models based on geographic weighted matrix; It considers the correlation between the variables of the model based on the Euclidean distance matrix between the centers of the traffic areas in such a way that the areas closer to the i area have a greater effect on the estimation of the coefficients of the independent variables in the i area and vice versa. Therefore, the coefficients of the independent variables are not the same in all TAZs. The most important models for estimating the frequency of urban accidents are the GWPR and GWNBR models (Almasi & Behnood, 2022; Lee et al., 2019; J. Liu, Khattak, & Wali, 2017).

In this study, in order to estimate the frequency of inner-city accidents based on independent environmental variables in the geographic state, the GWPR model has been used. In a GWPR model, accident frequency is predicted by a set of descriptive variables where the parameters are allowed to vary in space. This model can be written as equation (5) (J. Liu et al., 2017).

$$\ln(\lambda_i) = \beta_0(u_i, v_i) + \beta_1(u_i, v_i)\ln(E_{vi}) + \sum_{k=1}^{k} \beta_k(u_i, v_i) x_{ij} \qquad (6)$$

Relationship 4. Predicted frequency of accidents; The coordinates of the center of region j for j = 1,2,....,n, also tj is an offset variable in region j, α is the overdispersion parameter, the explanatory variable coefficient for 1,2,....K = k (J (Liu et al., 2017).

3.5. Spatial Bayes Models based on Neighbor Distance (with INLA Approach)

Bayesian inference is based on the probability distribution of EIFs coincidence parameters with respect to data in a hierarchical environment (Saha et al., 2018). The term hierarchical refers to the stepwise modeling of observed data, such that they are conditional on a set of parameters that are themselves probabilistically specified in terms of other parameters (Haining & Haining, 2003; Saha et al., 2018).

In this study, Bayesian method of integrated nested Laplace approximation (INLA) is used for inference as a well-known approach introduced by Rue et al. (Rue, Martino, & Chopin, 2009) in spatial Bayesian models with INLA approach; The estimation of the coefficients of the explanatory variables is done based on the neighborhood of area i with other areas by applying the spatial effect in such a way that if area i does not have a common border with j; Region j has no effect on the estimation of coefficients of explanatory variables of region i (Barmoudeh, Baghishani, & Martino, 2022).

Simple Poisson regression model and negative binomial distribution are both from the family of generalized linear models (GLMs). For the simple regression model, it is assumed that y=(y_1,…,y_n)^' is the response vector for n coincidence observations and x_i=(x_1i,…,x_pi) is the explanatory p-dimensional vector for each observation. Therefore, to define the GLM model with the response variable, we assume that the mean of the response variable E(y) is related to the linear predictor of the explanatory variables with the appropriate link function. One of the explanatory variables is a linear predictor with the effect of spatial contiguity in the form of equation (7).

$$\eta_i = \beta_0 + \sum_{j=1}^{p} x_{ij}\beta_j + \sum_{k=1}^{K} f(z_k) + \varepsilon_i \qquad (7)$$

In equation 5, there are model constants and coefficients that indicate the linear effect of covariates. Functions are covariates that are used to smooth these linear relationships or introduce random effects, and the error component is unstructured (i.e., lacking spatial or serial structure).

Since some spatial factors are more effective in the occurrence of accidents, also in many cases the accident data have a spatial regional dependence, which means that each region has a behavior dependent on its neighbors, so in this study, one of the assumptions of the common spatial structure for regional data, Conditional autoregression model (Besag, 1974). Intrinsic CAR (ICAR) model has been used to calculate spatial dependence (Nadifar, Baghishani, Fallah, & Rue, 2019). Now, the ICAR model assuming z=(z_1,...,z_n)' as the vector of spatial explanatory variables is summarized as equation (8).

$$f\ (z_i \left| z_{i-1}\right.), \sigma_z{}^2 \sim N\left(\frac{1}{\left|\delta_i\right|}\sum_{j\in\delta_i} z_j, \frac{\sigma_z{}^2}{\left|\delta_i\right|}\right), \qquad (8)$$

In relation 6, all members of the vector are except its number. The set of neighbors is for the neighbors of area i and the end of the set. It is also the variance parameter of the ICAR model.

3.6. Evaluation and Validation of Models

In this study, three criteria are used as goodness of fit (GOF[1]) and prediction accuracy based on studies (Saha et al., 2018; Stipancic et al., 2018)). The DIC criterion is one of the most important criteria used to evaluate the Poisson model and the negative binomial distribution in past studies. Provides complexity and model fit. DIC can be defined as follows:

$$DIC = D(\hat{\theta}) + 2\rho D = \bar{D} + \rho D \qquad (9)$$

In relation 7, D($\hat{\theta}$) is the deviation using the posterior mean values of the desired parameter, the posterior mean of the deviations and ρD is the effective number of parameters in the model. In this study, WAIC[2], which is another Bayesian criterion to identify the most valuable model, has been used, and finally, RMSE[3] to R has been used to calculate the root mean square error of the frequency of accidents (Almasi et al., 2021; Luo & Al-Harbi, 2017).

[1] criteria as a goodness of fit

[2] Watanabe-Akaike information criterion

[3] Root Mean Square Error

Table 1. Descriptive information of study variables

	Symbol	Variables	Min	Max	Median	Interquartile Range
Traffic feature	BS	Bus stop	0	21	4	4
	IN	Number of intersections	0	13	4	3
	HD	Average headway (min)	0.05	0.8	9.15	4
	AV	average speed (km/h)	20	49.61	19.26	2.5
	ALE	average length (km)	0.128	2.5	0.261	0.107
	AG	Average green time (min)	0	2.64	0.885	0.445
Demographics	LP	population logarithm	1	4	3.5	0.625
Path geometry	AW	average width (m)	2.5	3.5	2	8.85
	AL	Average number of lines	1	4	2	1
Land use diversity index	LUI	based on relation (1)	>0.001	0.835	0.617	0.16
Access to jobs index	JI	based on relation (2)	7.77	8.25	7.82	0.12
Index of access to public transportation	PTI	based on relation (2)	6.93	8	7.05	0.189

4. Results and Discussion

4.1. Descriptive Information of the Data

Table 1 shows the descriptive information of the dependent variable (accident frequency) and the accident EIFs. From 1398 to 1400, a total of 34,588 accidents (deaths, injuries, damages) were reported in Shiraz city, which averaged 106 accidents (minimum 7 and maximum 686) in 325 traffic areas in each traffic area.

In this study, in order to show the dispersion of the study data, the interquartile range statistic has been used, which shows that the frequency of accidents parameter has a high dispersion in the study area, as well as the variable of the number of bus stops and the average width of the thoroughfare has more dispersion than are other model variables.

Table 2 shows the results of GLM and Tobit models, based on which data compression and production of diversity indicators of user type, access to jobs and public transportation system in the model leads to more suitable results, so in this study Two variable categories of EIFs of urban accidents have been used.

The first category includes variables that directly affect the estimation of the frequency of accidents within the city on the TAZ scale (Table 1: population, bus station, intersections, time interval, average speed, average width of the road, average length of the road, average number of lanes, time interval, average green time) and the second category are indicators that indirectly affect the occurrence of accidents (user diversity index, access to jobs index and access to public transportation system) these indicators are based on relationships (1) and (2) have been calculated.

The Pearson linear correlation between pairs of independent variables of the model, based on which more than 90% of the independent variables have a linear correlation of less than 0.4. Based on previous studies (Saha et al., 2018), the value of Pearson's correlation coefficient equal to 0.5 is considered as a strong correlation threshold, which was not seen in this study.

Table 2. The results of the variable selection test

Model Type	MAD	RMSE
GLM Using All Variables	45.64	77.41
GLM Using PCA Variables	25.21	61.16
Tobit Using All Variables	45.62	42.96
Tobit Using PCA Variables	17.91	25.49

4.2. Estimation of the Frequency of Inner-City Accidents using the SW Method

The parameters of both models are allowed to change in space and have different values for each TAZ, so the values shown in a range of changes in the whole city of Shiraz. The results have shown that the parameters in both GWNBR and GWPR models have passed the significance level of 10%, but in the GWNBR model, the parameters of the model with a significance level of 5% more than the GWPR model have been observed. Through the non-stationarity test, all parameters have significant spatial variations, which past studies have confirmed this spatial property of accident data (Lee et al., 2019; Li, Wang, Liu, Bigham, & Ragland, 2013; Torun, Göçer , Yeşiltepe, & Argın, 2020). In the GWPR model, the range of changes in the coefficient of the bus station variable in the entire study from the minimum (-0.294) to the maximum (0.655) was allowed to change in space, which shows that under the influence of the spatial effect in each of the TAZs, this variable How much has been effective in the occurrence of accidents (J. Liu et al., 2017) that this value is more than the amount of traffic in the GWNBR model (minimum -0.041 and maximum 0.042) so it shows that in the GWNBR model because Adjusting the effect of data overdispersion, the amount of spatial effect has been minimized (Gschlößl & Czado, 2008) Also, the maximum variation range of the coefficients of independent variables in the GWPR model is related to the AL variable (Rang: 2.124) and the minimum is related to PTI (Rang: 0.455), which is significantly less in the GWNBR model. Maximum INT Rang: 0.595 and minimum AL Rang: 0.027). Therefore, based on the results of the research in the GWPR model, the variable AL in TAZs where the spatial effect is noticeable (such as the north and south of Shiraz) has the highest impact on the occurrence of accidents, which is consistent with the study (J. Liu et al., 2017).

In TAZs where the spatial effect is less observed (Shiraz city center), the effect of green light time on the occurrence of accidents in both GWPR and GWNBR models is greater than other variables, which is consistent with the study (Paul & Ghosh, 2020). In both models, it has been shown that all the indicators created in this study have been effective in the occurrence of inner-city accidents, among which the index of user diversity has had a higher effect on the occurrence of inner-city accidents, which studies (Fuentes et al. , 2022; Musa & Moses, 2014; Su & Sze, 2022; Sung et al., 2022; Tagar & Pulugurtha, 2021).

In both models, it has been shown that the spatial effect in the TAZs of the city center is less than in other urban areas, so that in many TAZs located in the center towards the south of the city, the spatial effect is close to zero, which can be obtained in the areas Central has had less effect in estimating the coefficients of independent variables (Mutiso et al., 2022). The GWNBR model has a better goodness of fit than the GWPR model based on WAIC, DIC, and RMSE values (lower values and better fit).

4.3. Estimating the Frequency of Inner-City Accidents using the SA Method

Bayesian inference is based on the posterior density interval of the parameters, commonly known as the Bayesian confidence interval (CI). A 95% CI, which includes values between the 2.5th percentile and 97.5% of the posterior probability distribution, is commonly used to determine the validity of a variable. If the 95% CI of a parameter does not contain zero, that parameter is valid. A valid parameter has a positive (i.e. incremental) effect on crash frequency if the two limits of the parameter are greater than zero at the 95% CI. Similarly, a valid parameter has a negative effect (ie, reduction) on crash frequency if two bounds of the parameter are less than zero at the 95% CI (Saha et al., 2018).Considering the posterior averages of auxiliary variables, in general, the results have shown that the relationships between the selected independent variables and the frequency of accidents are well established using the SA method, which previous studies (Stipancic et al., 2018, 2019) have established. supports. In both models, a positive spatial effect (Spatial Precision) has been obtained, which shows that the spatial effect is evident in the entire study area and has been effective in estimating the coefficients of the independent variables. In both models, the posterior mean is positive for the number of bus stops. This result reflects the expectation that an increase in the number of bus stops should increase the probability of an accident. Therefore, it is logical that TAZs with more number of bus stops have more number of accidents which is consistent with the study (Deotima Mukherjee, Rao, & Tiwari, 2022). An unexpected result is the negative mean posterior PTI observed in both models, which is confirmed by the study (Su, Sze, & Bai, 2021) and with the study (Rojas-Rueda, Nieuwenhuijsen, Khreis, & Frumkin, 2020) does not match. The posterior mean of the LUI variable in this study is positive, which shows that the increase in the variety of user types in TAZ increases the frequency of accidents, which is consistent with the study (Sung et al., 2022). Therefore, the increase in the variety of users in TAZ is due to Interference of types of trips in a certain area causes accidents (Sung et al., 2022). As expected, the AS variable in the occurrence of an accident was found positive in both models, so it shows that an increase in speed in the TAZ leads to an increase in the frequency of accidents, which is consistent with previous studies (Stipancic et al., 2018, 2019). One of the most important variables that has a positive relationship with the occurrence of an accident is the AL variable, which is confirmed by studies (Álvarez, Fernández, Gordaliza, Mansilla, & Molinero, 2020; Y. Guo, Li, Liu, & Wu, 2019). It has shown the effect of the number of crossing lines on the occurrence of accidents. Also, the relationship between the frequency of accidents and IN and AG variables in both models has been positive, which, of course, have less impact than other variables such as AS and AL, so it is expected that with Increasing the number of intersections and increasing the green time in each TAZ increases the number of accidents in that TAZ, which is consistent with the study (Lee et al., 2019).

In general, out of the 12 independent variables of this study, in both the SAP and SANB models of the INLA method, 8 variables have been significant, which indicates the acceptable effect of the selected variables on the occurrence of urban accidents (Saha et al., 2018; Stipancic et al. , 2019). Comparatively, it can be said that in TAZs where the number of bus stops is more and the variety of uses is more, as well as the speed of vehicles is more, the probability of an accident is higher than in TAZs where the posterior average coefficient values of these three variables are lower than other TAZs. This type of comparison has been seen in studies (Saha et al., 2018; Yasmin & Eluru, 2016).

Figure 3. It shows the results of estimating the frequency of accidents at the TAZ level using the SA method. Also, the values of the spatial effect in each of the TAZs are shown in the

estimation of the accident frequency model. In the SAP model, it varies between (-1 to 1.5) in the entire studied range. In TAZs with green color (northern part and center of the city), more spatial effect has been observed than other TAZs, which shows that the estimation of the frequency of accidents in these areas under the spatial effect is high and positive, which means that the spatial effect causes The increase in the probability of an accident has occurred (Gschlößl & Czado, 2008). In the outskirts and south of the city, the spatial effect is close to zero and negative, which shows that the spatial effect has had a decreasing effect on the probability of accidents in these areas (Mutiso et al., 2022) in the SANB model, the results are similar to the SAP model. It has been obtained (positive spatial effect in the north and center of the city), but the range of the spatial effect is lower than that of the SAP model, which can be due to the consideration of the effect of spatial overdispersion in SANB (Almasi & Behnood, 2022).

4.4. Comparison and validation of models

Based on the evaluation criteria at the end of Table 2, the DIC values in the SAP model are lower compared to the other three models, which indicates the goodness of the model fit (Saha et al., 2018; Stipancic et al., 2019) as well as the WAIC values , R^2, RMSE shows that SA models provide better results than SW based models.

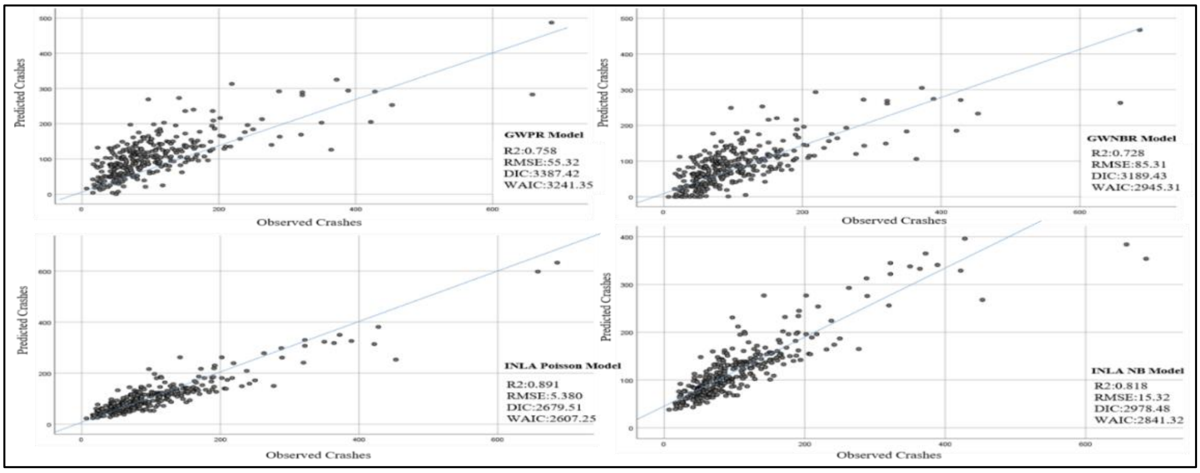

Figure 2. Comparison between predicted values using SA and SW models

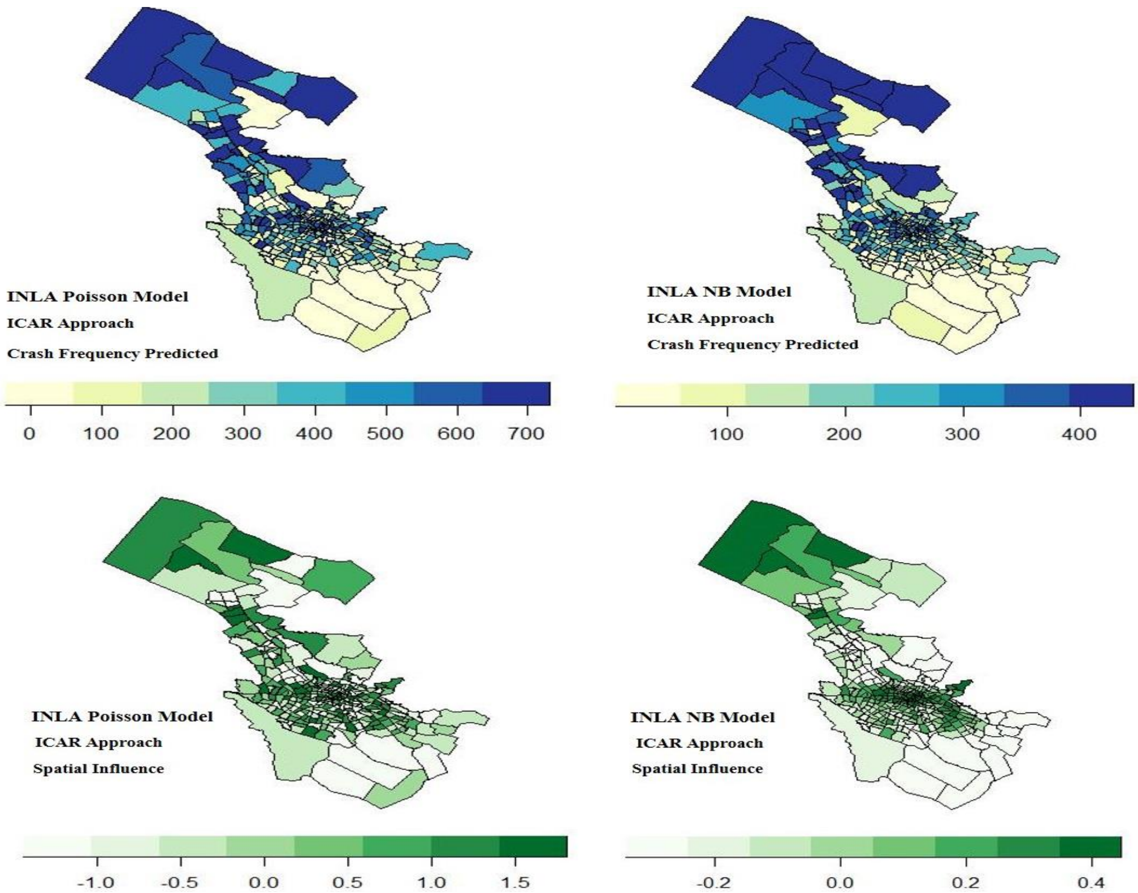

Figure 3. Estimation of the frequency of inner-city accidents and the spatial effect of independent variables using the SA method

Table 3. Results of Poisson and NB modeling by spatial INLA method

Variables	**PO**				**NB**			
Statistic	Mean	SD	95% CI		Mean	SD	95% CI	
BS	0.0442	0.00941	0.025885	0.062827	0.0442	0.009	0.0258	0.0628
AH	-0.010	0.00701	-0.02379	0.003742	-0.010	0.0070	-0.023	0.0037
AS	0.1412	0.03376	0.074868	0.207458	0.1412	0.0337	0.0748	0.2074
AL	0.22486	0.05816	0.110609	0.339025	0.22486	0.0581	0.1100	0.33902
ALE	0.00892	0.00650	-0.00388	0.021682	0.00892	0.0065	-0.003	0.02168
AG	0.0906	0.03535	0.05978	0.12097	0.0901	0.0354	0.019	0.1202
IN	0.07958	0.01436	0.051427	0.107853	0.07958	0.0143	0.0514	0.10785
PTI	-0.0499	0.12275	-0.2829	-0.009356	-0.0499	0.1227	0.082	0.1993
INT	0.0814	0.03099	0.02249	0.149252	0.0614	0.0309	0.012	0.1492
JI	0.10890	0.11606	-0.12603	0.330095	0.10890	0.116	-0.126	0.33009
LP	0.05835	0.04601	-0.03188	0.148882	0.05835	0.0460	-0.031	0.14888
Summary Statistics								
DIC		2679.51			2978.48			
WAIC		2607.25			2841.32			
Pseudo R2		0.891			0.818			
RMSE		5.380			15.32			

5. Conclusion

This study was conducted in order to investigate the spatial effect of EIFs of urban accidents using spatial statistics methods based on the Euclidean distance of the geographic weight matrix (SW) and methods based on distance and adjacency (SA). In the first step of the study, all the available variables were collected and according to the previous studies, data compression was done in order to reduce the effect of collinearity and autocorrelation between the independent variables using the PCA method. In the PCA method, indicators of user diversity, access to the public transportation system, and access to jobs were included as categories of indirect environmental factors in the modeling. In all models, the values of the number of bus stops, average speed, average green light time, as well as the diversity index of the type of use have been identified as the most important accident EIFs. It is different, so by controlling these factors according to its value in each TAZ, the spatial effect of accident EIFs can be reduced in that TAZ. In this study, 4 spatial statistics models GWPR, GWNBR based on Euclidean distance and SAP and SANB models based on neighborhood distance with INLA approach have been used to estimate the frequency of accidents in the city by considering the spatial effects of accident EIFs. Based on the results of the study, the spatial estimation of EIFs of inner-city accidents in the neighborhood approach (SA) using the INLA method provided better results, in other words, the estimation of the coefficients of EIFs in the neighborhood method gave better results than the same estimation in the geographic weighted matrix (SW) method. The comparison between SA and SW spatial statistics models has not been observed in previous studies. This study showed that in the inner city areas at the TAZ level, due to the contiguity and common border in all TAZs, the SA method estimates the coefficients of EIFs better. Therefore, in this study, it was shown that the neighborhood effect can have a negative effect on the estimation of EIFs coefficients, in other words, the spatial effect may reduce the severity of accidents in the spatial model compared to the classical statistical model.

In this study, in both SA and SW methods, the reduction of spatial effects in the estimation of EIFs coefficients in SANB models was quite noticeable, which shows that considering the effect of over dispersion in the estimation of SANB models has reduced the range of spatial effects.

The identification of accident EIFs after the multi-stage investigation in this study at the district level can be a useful element for formulating policies in support of urban accident reduction measures. Modifying the characteristics of bus stops in inappropriate places that increase the tension (conflict) between vehicles, The rescheduling of traffic lights and proper control over land use changes in the city are among the measures that can reduce the probability of accidents in urban areas.

6. References

- Al-Aamri, A. K., Hornby, G., Zhang, L.-C., Al-Maniri, A. A., & Padmadas, S. S. (2021). Mapping road traffic crash hotspots using GIS-based methods: A case study of Muscat Governorate in the Sultanate of Oman. Spatial Statistics, 42, 100458.

- Almasi, S. A., & Behnood, H. R. (2022). Exposure based geographic analysis mode for estimating the expected pedestrian crash frequency in urban traffic zones; case study of Tehran. Accident Analysis & Prevention, 168, 106576.

- Almasi, S. A., Behnood, H. R., & Arvin, R. (2021). Pedestrian crash exposure analysis using alternative geographically weighted regression models. Journal of advanced transportation, 2021.

- Álvarez, P., Fernández, M. A., Gordaliza, A., Mansilla, A., & Molinero, A. (2020). Geometric road design factors affecting the risk of urban run-off crashes. A case-control study. PLoS one, 15(6), e0234564.

- Bao, J., Liu, P., & Ukkusuri, S. V. (2019). A spatiotemporal deep learning approach for citywide short-term crash risk prediction with multi-source data. Accident Analysis & Prevention, 122, 239-254.

- Bao, J., Yang, Z., Zeng, W., & Shi, X. (2021). Exploring the spatial impacts of human activities on urban traffic crashes using multi-source big data. Journal of transport geography, 94, 103118.

- Barmoudeh, L., Baghishani, H., & Martino, S. (2022). Bayesian spatial analysis of crash severity data with the INLA approach: Assessment of different identification constraints. Accident Analysis & Prevention, 167, 106570.

- Chen, P., & Zhou, J. (2016). Effects of the built environment on automobile-involved pedestrian crash frequency and risk. Journal of Transport & Health, 3(4), 448-456.

- Chen, T., Sze, N., Chen, S., Labi, S., & Zeng, Q. (2021). Analysing the main and interaction effects of commercial vehicle mix and roadway attributes on crash rates using a Bayesian random-parameter Tobit model. Accident Analysis & Prevention, 154, 106089.

- Congiu, T., Sotgiu, G., Castiglia, P., Azara, A., Piana, A., Saderi, L., & Dettori, M. (2019). Built environment features and pedestrian accidents: An Italian retrospective study. Sustainability, 11(4), 1064.

- Da Silva, A. R., & Fotheringham, A. S. (2016). The multiple testing issue in geographically weighted regression. Geographical Analysis, 48(3), 233-247.

- Delclòs-Alió, X., & Miralles-Guasch, C. (2018). Looking at Barcelona through Jane Jacobs's eyes: Mapping the basic conditions for urban vitality in a Mediterranean conurbation. Land Use Policy, 75, 505-517.

- Ewing, R., & Dumbaugh, E. (2009). The built environment and traffic safety: a review of empirical evidence. Journal of Planning Literature, 23(4), 347-367.

- Fuentes, L., Truffello, R., & Flores, M. (2022). Impact of Land Use Diversity on Daytime Social Segregation Patterns in Santiago de Chile. Buildings, 12(2), 149.

- Galgamuwa, U., Du, J., & Dissanayake, S. (2021). Bayesian spatial modeling to incorporate unmeasured information at road segment levels with the INLA approach: A methodological advancement of estimating crash modification factors. Journal of traffic and transportation engineering (English edition), 8(1), 95-106.

- Gschlößl, S., & Czado, C. (2008). Modelling count data with overdispersion and spatial effects. Statistical papers, 49(3), 531-552.

- Guo, F., & Lu, D. (2022). How many crashes does cellphone use contribute to? Population attributable risk of cellphone use while driving. Journal of safety research, 82, 385-391.

- Guo, Y., Li, Z., Liu, P., & Wu, Y. (2019). Exploring risk factors with crashes by collision type at freeway diverge areas: accounting for unobserved heterogeneity. IEEE Access, 7, 11809-11819.

- Hezaveh, A. M., Arvin, R., & Cherry, C. R. (2019). A geographically weighted regression to estimate the comprehensive cost of traffic crashes at a zonal level. Accident Analysis & Prevention, 131, 15-24.

- Huang, H., Zeng, Q., Pei, X., Wong, S., & Xu, P. (2016). Predicting crash frequency using an optimised radial basis function neural network model. Transportmetrica A: transport science, 12(4), 330-345.

- Intini, P., Berloco, N., Fonzone, A., Fountas, G., & Ranieri, V. (2020). The influence of traffic, geometric and context variables on urban crash types: A grouped random parameter multinomial logit approach. Analytic methods in accident research, 28, 100141.

- Iyanda, A. E., Adeleke, R., Lu, Y., Osayomi, T., Adaralegbe, A., Lasode, M., Osundina, A. M. (2020). A retrospective cross-national examination of COVID-19 outbreak in 175 countries: a multiscale geographically weighted regression analysis (January 11-June 28, 2020). Journal of infection and public health, 13(10), 1438-1445.

- Kang, C.-D. (2018). The S+ 5Ds: Spatial access to pedestrian environments and walking in Seoul, Korea. Cities, 77, 130-141.

- Lee, J., Abdel-Aty, M., & Shah, I. (2019). Evaluation of surrogate measures for pedestrian trips at intersections and crash modeling. Accident Analysis & Prevention, 130, 91-98.

- Liu, C., & Sharma, A. (2017). Exploring spatio-temporal effects in traffic crash trend analysis. Analytic methods in accident research, 16, 104-116.

- Liu, J., Hainen, A., Li, X., Nie, Q., & Nambisan, S. (2019). Pedestrian injury severity in motor vehicle crashes: an integrated spatio-temporal modeling approach. Accident Analysis & Prevention, 132, 105272.

- Moomen, M., Rezapour, M., Raja, M., & Ksaibati, K. (2020). Predicting downgrade crash frequency with the random-parameters negative binomial model: Insights into the impacts of geometric variables on downgrade crashes in Wyoming. IATSS Research, 44(2), 94-102.

- Mukherjee, D., & Mitra, S. (2021). Investigating the fatal pedestrian crash occurrence in urban setup in a developing country using multiple-risk source model. Accident Analysis & Prevention, 163, 106469.

- Mukherjee, D., Rao, K. R., & Tiwari, G. (2022). Built-environment risk assessment for pedestrians near bus-stops: a case study in Delhi. International journal of injury control and safety promotion, 1-10.

- Munira, S., Sener, I. N., & Dai, B. (2020). A Bayesian spatial Poisson-lognormal model to examine pedestrian crash severity at signalized intersections. Accident Analysis & Prevention, 144, 105679.

- Musa, I. J., & Moses, A. O. (2014). An analysis of the effect of land use on road traffic accidents in Zaria. International Journal of Development and Sustainability, 3(3), 520-529.

- Mutiso, F., Pearce, J. L., Benjamin-Neelon, S. E., Mueller, N. T., Li, H., & Neelon, B. (2022). Bayesian negative binomial regression with spatially varying dispersion: Modeling COVID-19 incidence in Georgia. Spatial Statistics, 100703.

- Nadifar, M., Baghishani, H., Fallah, A., & Rue, H. (2019). Statistical modeling of groundwater quality assessment in Iran using a flexible Poisson likelihood. arXiv preprint arXiv:1908.02344.

- Obelheiro, M. R., da Silva, A. R., Nodari, C. T., Cybis, H. B. B., & Lindau, L. A. (2020). A new zone system to analyze the spatial relationships between the built environment and traffic safety. Journal of transport geography, 84, 102699.

- Olsson, U. (2002). Generalized linear models: an applied approach: Lund, Studentlitteratur.

- Pal, C., Hirayama, S., Narahari, S., Jeyabharath, M., Prakash, G., & Kulothungan, V. (2018). An insight of World Health Organization (WHO) accident database by cluster analysis with self-organizing map (SOM). Traffic injury prevention, 19(sup1), S15-S20.

- Papa, E., Silva, C., Te Brömmelstroet, M., & Hull, A. (2016). Accessibility instruments for planning practice: A review of European experiences. Journal of transport and land use, 9(3), 57-75.

- Paul, M., & Ghosh, I. (2020). Influence of green signal countdown timer on severe crash types at signalized intersections due to red light violations. Transportation letters, 12(8), 528-539.

- Peera, K. M., Shekhawat, R. S., & Prasad, C. (2019). Traffic analysis zone level road traffic accident prediction models based on land use characteristics. International journal for traffic and transport engineering (Belgrade), 9(4), 376-386.

- Rahmani, M., Lotfata, A., Zebardast, E., Rastegar, S., Sanchez, T. W., Goharrizi, B. A., & Landi, S. (2022). Land use suitability assessment for economic development at the provincial level: The case study of Yazd Province, Iran. Sustainable Cities and Society, 87, 104163.

- Rashidi, M. H., Keshavarz, S., Pazari, P., Safahieh, N., & Samimi, A. (2022). Modeling the accuracy of traffic crash prediction models. IATSS Research.

- Saha, D., Alluri, P., Gan, A., & Wu, W. (2018). Spatial analysis of macro-level bicycle crashes using the class of conditional autoregressive models. Accident Analysis & Prevention, 118, 166-177.

- Satria, R., Aguero-Valverde, J., & Castro, M. (2021). Spatial analysis of road crash frequency using Bayesian models with Integrated Nested Laplace Approximation (INLA). Journal of Transportation Safety & Security, 13(11), 1240-1262.

- Se, C., Champahom, T., Jomnonkwao, S., Chaimuang, P., & Ratanavaraha, V. (2021). Empirical comparison of the effects of urban and rural crashes on motorcyclist injury severities: A correlated random parameters ordered probit approach with heterogeneity in means. Accident Analysis & Prevention, 161, 106352.

- Sobreira, L. T. P., & Cunto, F. (2021). Disaggregated traffic conditions and road crashes in urban signalized intersections. Journal of safety research, 77, 202-211.

- Soroori, E., Mohammadzadeh Moghaddam, A., & Salehi, M. (2021). Modeling spatial nonstationary and overdispersed crash data: Development and comparative analysis of global and geographically weighted regression models applied to macrolevel injury crash data. Journal of Transportation Safety & Security, 13(9), 1000-1024.

- Stipancic, J., Miranda-Moreno, L., Saunier, N., & Labbe, A. (2018). Surrogate safety and network screening: Modelling crash frequency using GPS travel data and latent Gaussian Spatial Models. Accident Analysis & Prevention, 120, 174-187.

- Stipancic, J., Miranda-Moreno, L., Saunier, N., & Labbe, A. (2019). Network screening for large urban road networks: using GPS data and surrogate measures to model crash frequency and severity. Accident Analysis & Prevention, 125, 290-301.

- Su, J., & Sze, N. (2022). Safety of walking trips accessing to public transportation: a Bayesian spatial model in Hong Kong. Travel behaviour and society, 29, 125-135.

- Su, J., Sze, N., & Bai, L. (2021). A joint probability model for pedestrian crashes at macroscopic level: Roles of environment, traffic, and population characteristics. Accident Analysis & Prevention, 150, 105898.

- Sung, H., Lee, S., Cheon, S., & Yoon, J. (2022). Pedestrian Safety in Compact and Mixed-Use Urban Environments: Evaluation of 5D Measures on Pedestrian Crashes. Sustainability, 14(2), 646.

- Tagar, S., & Pulugurtha, S. S. (2021). Predictor variables influencing merging speed change lane crash risk by interchange type in urban areas. Transportation research interdisciplinary perspectives, 10, 100375.

- Tang, H., & Donnell, E. T. (2019). Application of a model-based recursive partitioning algorithm to predict crash frequency. Accident Analysis & Prevention, 132, 105274.

- Tang, J., Zhao, C., Liu, F., Hao, W., & Gao, F. (2022). Analyzing travel destinations distribution using large-scaled GPS trajectories: A spatio-temporal Log-Gaussian Cox process. Physica A: Statistical Mechanics and its Applications, 599, 127305.

- Tarko, A. P. (2018). Estimating the expected number of crashes with traffic conflicts and the Lomax Distribution–A theoretical and numerical exploration. Accident Analysis & Prevention, 113, 63-73.

- Tarko, A. P. (2023). Maximum likelihood method of estimating the conflict-crash relationship. Accident Analysis & Prevention, 179, 106875.

- Tiwari, G., & Jain, D. (2012). Accessibility and safety indicators for all road users: case study Delhi BRT. Journal of transport geography, 22, 87-95.

- Umair, M., Rana, I. A., & Lodhi, R. H. (2022). The impact of urban design and the built environment on road traffic crashes: a case study of Rawalpindi, Pakistan. Case studies on transport policy, 10(1), 417-426.

- Yan, X., Feng, Y., Tong, X., Li, P., Zhou, Y., Wu, P.,. Liu, S. (2021). Reducing spatial autocorrelation in the dynamic simulation of urban growth using eigenvector spatial filtering. International Journal of Applied Earth Observation and Geoinformation, 102, 102434.

- Yasmin, S., & Eluru, N. (2016). Latent segmentation based count models: analysis of bicycle safety in Montreal and Toronto. Accident Analysis & Prevention, 95, 157-171.

- Yoon, S. H., Kim, Y. W., & Ji, Y. G. (2019). The effects of takeover request modalities on highly automated car control transitions. Accident Analysis & Prevention, 123, 150-158.

- 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. Geography and sustainability, 3(4), 312-324.

- Zafri, N. M., Prithul, A. A., Baral, I., & Rahman, M. (2020). Exploring the factors influencing pedestrian-vehicle crash severity in Dhaka, Bangladesh. International journal of injury control and safety promotion, 27(3), 300-307.

- Zarei, M., Hellinga, B., & Izadpanah, P. (2022). CGAN-EB: A non-parametric empirical Bayes method for crash frequency modeling using conditional generative adversarial networks as safety performance functions. International Journal of Transportation Science and Technology.

- Zhai, X., Huang, H., Gao, M., Dong, N., & Sze, N. (2018). Boundary crash data assignment in zonal safety analysis: an iterative approach based on data augmentation and Bayesian spatial model. Accident Analysis & Prevention, 121, 231-237.

- Zhou, Y., Jiang, X., Fu, C., Liu, H., & Zhang, G. (2022). Bayesian spatial correlation, heterogeneity and spillover effect modeling for speed mean and variance on urban road networks. Accident Analysis & Prevention, 174, 106756.