Applied Spatial Analysis and Policy
https://doi.org/10.1007/s12061-023-09519-w

Spatial Association Between Urban Neighbourhood Characteristics and Child Pedestrian–Motor Vehicle Collisions

Emad Soroori[1] · Behzad Kiani[2] · Soraya Ghasemi[1] · Alireza Mohammadi[3] · Hamidreza Shabanikiya[4] · Robert Bergquist[5] · Fatemeh Kiani[6] · Hossein Tabatabaei-Jafari[7]

Received: 9 June 2022 / Accepted: 8 May 2023
© The Author(s), under exclusive licence to Springer Nature B.V. 2023

Abstract
This study examined the relationship between environmental and socioeconomic factors and the number of motor vehicle collisions involving young pedestrians. The research encompassed urban neighbourhoods as well as an entire metropolitan area and analyzed data from 7,028 motor vehicle collisions that involved pedestrians aged 18 years or younger, occurring between 2015 and 2019 in the city of Mashhad, Iran. Thirteen indices related to socioeconomic and built environmental factors were quantified at the neighbourhood level. To model the relationship between these explanatory factors and the number of collisions investigated, Poisson and negative binomial models were developed using the geographically weighted regression (GWR) technique. The GWR was used to account for the impact of location on the association between explanatory factors and the count of collisions. The study found that the population of young people, road area ratio, main road intersection ratio, average maximum speed limit, non-motorized travels, sidewalk area ratio, sidewalk disconnections, number of schools, unemployment ratio, illiteracy rate, and open space ratio were significantly associated with child pedestrian-motor vehicle collisions. However, these associations were not uniform across the entire study area. It is possible that unknown factors or an unknown interaction of known factors in different parts of the urban area may have influenced the observed associations.

Keywords Built environment · Geographically Weighted Negative Binomial Regression · Geographically Weighted Poisson Regression · Child pedestrian · Socioeconomic · Traffic safety

✉ Behzad Kiani
Behzad.Kiani@umontreal.ca

Extended author information available on the last page of the article

Published online: 24 May 2023

Introduction

Roads are one of the primary components of the transportation system and the most common way for people to travel, providing significant economic benefits and social values such as equity for communities. However, traffic accidents, as well as air pollution and noise from motorized vehicles, are negative aspects of this transportation approach. According to the World Health Organization, 1.35 million people lose their lives in traffic collisions around the world each year (World Health Organization, 2018). Vulnerable road users such as pedestrians, motorcyclists, and cyclists are at higher risk and make up half of the fatalities (World Health Organization, 2018). Among these vulnerable road users, underage children and those with less experience, creates significant challenges for effective traffic safety management. These individuals are at higher risk of being involved in traffic collisions and are more likely to suffer severe injuries or fatalities as a result. In fact, according to available global data, traffic collisions account for approximately 22% of all children fatalities worldwide. Therefore, it is crucial to prioritize measures that enhance the safety of vulnerable road users, particularly children, in efforts to reduce the number of traffic-related fatalities and injuries (Branche et al., 2008).

Based on the aforementioned statistics, it is imperative to plan for a safer transportation system that adequately addresses the needs of vulnerable road users, especially those under the age of 18. To incorporate safety considerations at the planning level, spatial statistical models are widely recognized as useful tools that can facilitate the traffic safety management process (Part, 2010). These models enable planners to quantify the impacts of various transportation system features on traffic collision incidence counts and to predict future outcomes based on current conditions. By leveraging these methods, planners can make informed decisions about how to prioritize safety initiatives and allocate resources most effectively to reduce the frequency and severity of traffic-related injuries and fatalities.

From the perspective of factors that affect collision incidences, individual and non-individual environmental factors are two broad categories of features that have been studied in the field of vehicle collisions (Shabanikiya et al., 2020). Individual factors such as age and gender can partially explain variations in child pedestrian/motor vehicle collisions (CP-MVCs), while built environment characteristics and transportation infrastructure are non-individual factors (Kuşkapan et al., 2022; Stevenson, 2006; Wang et al., 2018). Examples of built environmental factors that potentially contribute to CP-MVCs include the municipality of the intersection, two-lane roads, high vehicle speeds, lack of clear pedestrian crossing routes, sidewalk disconnections, lack or inappropriate location of pedestrian bridges, mixed-land use, and inadequacy of open public spaces (Asare & Mensah, 2020; Briz-Redón et al., 2019; Ma et al., 2018; Muazir & Hsieh, 2019; Peera et al., 2019; Rothman et al., 2022; Rus et al., 2019; Xie et al., 2019). Socio-economic determinants can also have an impact, including the general, non-individual determinants of urban areas on one hand, and the socio-economic influence of the individual and family on the other. The study of the role of built environmental factors without considering the socio-economic status of urban neighborhoods would not provide a sufficiently deep overview of the incidence of CP-MVCs (Chakravarthy et al., 2010; Li et al., 2016; Pernica et al., 2012; Silverman et al., 2013).

In terms of data availability, researchers often face the issue of unobserved heterogeneity, which can result in missing valuable information that explains some of the variation in crash data. Unobserved data can also affect the relationship between explanatory and response variables across different spatial locations. Modelling the relationships between variables can be achieved through two general approaches: global and local models. The former assumes constant relationships, while the latter considers variations in these relationships. Geographically Weighted Regression (GWR) is a local technique that modifies Generalized Linear Models (GLM) by allowing spatial variations of parameters. Geographically Weighted Poisson Regression (GWPR) and Geographically Weighted Negative Binomial Regression (GWNBR) are two extensions of the GWR method that fit crash data well, as they can be developed based on count and skewed data (Nakaya et al., 2005). Additionally, the GWNBR form can perform better when data are over dispersed, as it allows the variance parameter to have a higher value than the mean (Da Silva & Rodrigues, 2014). However, the GWNBR is a more complex model, which may result in weaker predictive performance (Soroori et al., 2021).

Urban areas are known to have heterogeneous environmental characteristics, which can lead to spatial variation in the incidence of CP-MVCs (Curtis, 2017; Shabanikiya et al., 2020). However, it is not clear whether there is a consistent pattern of CP-MVCs across all neighborhoods or districts within a large urban area, or if geographical location has a significant impact due to spatial variations in built environment, natural landscape or socioeconomic characteristics. To address this gap in knowledge, the present study aims to investigate the association between environmental and socioeconomic characteristics and the count of CP-MVCs involving pedestrians aged 18 years or younger. Methodologically, this study uses GWR methods to account for spatial heterogeneity, in contrast to conventional models. Furthermore, it is important to determine which of the two GWR extensions, GWPR or GWNBR, performs better in predicting CP-MVCs.

Materials and Methods

Study Area

The dataset used in this study comprised all pedestrian-vehicle collisions involving individuals 18 years old or younger within the metropolitan area of Mashhad from March 2015 to 2019, at the neighborhood level. The city has a population of 3,372,660 people according to the 2016 census and covers a land area of 307 km^2 across 175 neighborhoods (Fig. 1). On average, each neighborhood covers 1.7 km^2 and has a population of 17,307 people. In 2020, pedestrians in Mashhad were involved in more than 51% of traffic collisions (compared to the national average of 43%), with 12.5% of those involving people under the age of 18 (Behzadnia & Shahmohammadi, 2016; Mashhad Traffic and Transportation organisation, 2020).

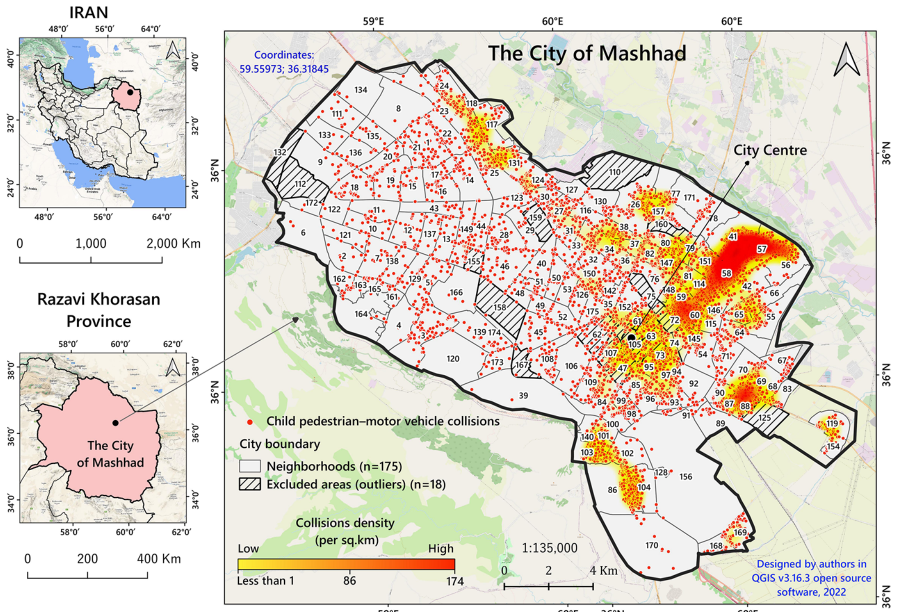

Fig. 1 Map of the study area indicating location, administration areas and the distribution of road traffic collisions with the neighbourhoods identified by number

A cross-platform and open-source desktop geographical information systems (GIS) application (QGIS 3.16.1) was used to create the map of the study area (QGIS software, 2020).

Geodatabase and Variable Selection Procedure

The dependent variable considered in this study was the count of CP-MVC in each neighbourhood between 22nd March 2015 and 20th March 2019. To scale the outcome variable, the population of people aged 18 years old or younger was used as an exposure variable. Thirteen explanatory variables were selected from the spatial data system of the Mashhad Municipality based on the literature (Table 1).

Data Analysis

Two types of analyses were conducted, conventional generalized linear models including Poisson and negative binomial regression models to investigate the association between the dependent and explanatory factors and GLMs developed based on geographically weighted regression (GWR) methods to investigate geographical impact on this association.

Table 1 Explanatory variables used to model child pedestrian-motor vehicle collisions

Variable	Definition and rationales	Measure and unit
Total number of accidents (Dependent variable)	Total number of CP-MVC in each neighbourhood	The total number of CP-MVC happened in each neighbourhood
Pop0-19 (Exposure variable)	Neighborhood total population ages 0–19	The total number of people aged 18 or younger in each neighbourhood
Road area ratio*	A road or street area is an urban surface on which motor vehicles move. With the increase in the area of the streets inside the cities, the possibility of collisions will be increased (Congiu et al., 2019; Jamshidi et al., 2017; Schuurman et al., 2020).	Road area (km^2)/neighbourhood total area (km^2) × 100.
Main road intersection ratio	A road intersection is where two or more roads meet. Some studies have shown that the probability of collisions increases with the increase in the number of intersections within cities (Congiu et al., 2019; Hwang et al., 2017; Rothman et al., 2017; Schuurman et al., 2020; Sun et al., 2018).	Number of main road intersections (includes squares)/km main road length in each neighbourhood (km), no. per km of main road length.
Average maximum speed limit*	A speed limit is the maximum speed allowed by law for road vehicles. In some studies, the role of high speed in increasing the number and severity of collisions has been proven (Bennet & Yiannakoulias, 2015; Congiu et al., 2019; Jamshidi et al., 2017; Schuurman et al., 2020).	The average speed limit specified for moving cars, km per hour for each neighbourhood (kph).
Proportion of public transit supply	Public transit supply includes facilities such as bus stations and subways. In many cases, crowding and rushing to arrive at these facilities may increase the likelihood of collisions with motor vehicles (Congiu et al., 2019; Hwang et al., 2017; Schuurman et al., 2020).	Number of bus stops/km road length in each neighbourhood) × 100 (%).
NMT proportion *	NMT includes all forms of travels that do not rely on an engine or motor for movement. With the increase in the share of this kind of transportation in urban areas, the probability of collisions will decrease (Litman, 2012; Tiwari, 1999).	Ratio of non-motorized transportation to motorized roads (%) = Total sidewalk length (km) + cycle length (km)/total street length /km) × 100 (%).
Sidewalk area ratio*	Sidewalk is a paved path for pedestrians at the side of a road; a pavement. Some studies have shown that increasing the area of sidewalks can reduce CP-MVCs (Congiu et al., 2019).	Sidewalk area (km^2) /neighbourhood total area (km^2) × 100 (%).

Table 1 (continued)

Variable	Definition and rationales	Measure and unit
Sidewalk disconnection index	It refers to the continuity and uninterrupted network of sidewalks. The unconnected sidewalks can lead to CP-MVCs (Congiu et al., 2019; Miranda-Moreno et al., 2011)	Cross intersections in sidewalks (no.)/ total length of sidewalks (km) in neighbourhoods = no. per km of sidewalk length.
Pedestrian bridges ratio	A pedestrian bridge is a structure built over a road so that people can cross from one side to the other. It is assumed that the presence of these structures helps increase the safety of children and reduces the risk of collisions (Jamshidi et al., 2017)	Number of pedestrian bridges per road length (km) in each neighbourhood × 100.
School presence ratio*	It is assumed that the concentration of schools in particular areas perhaps increases the risk of CP-MVCs (Clifton et al., 2009; Congiu et al., 2019; Elias & Shiftan, 2014; Rothman et al., 2017).	The number of schools (primary, secondary and high schools) and education centres in each neighbourhood/population $\leq 18 \times 100$ (%).
Land-use mix*	Land use mix refers to how different land uses are physically and functionally integrated into each neighbourhood. Residential, commercial, recreational (sport and public parks), and religious land-use parcels were used in this study (Congiu et al., 2019; Hagel et al., 2019; Hwang et al., 2017; Rothman et al., 2017; Schuurman et al., 2020).	Land-use mix $= \frac{\sum_1^k (p_k ln p_k}{\ln N}$ (Value ranges from 0 (homogeneity) to 1 (diversity).
Open space ratio*	Open Space is empty lands that are not intensively developed for residential, commercial, industrial, or institutional use. It serves many purposes, whether it is publicly or privately owned. It is hypothesized that open spaces can help increase child pedestrian safety and reduce the risk of collisions (Egorov et al., 2016)	Total open space (includes vacant lots, roadside parking lots and wastelands) (km^2) + total green space (includes community woodlands, landscape around buildings, neighbourhood parks and gardens) (km^2)/ total neighbourhood area (km^2) × 100 (%).
Illiteracy ratio*	The illiteracy rate is defined by the percentage of the population of a given age group that cannot read and write. Some studies have considered the high rate of illiteracy in certain urban areas as one of the risk factors in increasing the risk of CP-MVCs (Elias & Shiftan, 2014)	Ratio of illiteracy in the total population ≥ 6 years = illiterate population /population 6 years old and above × 100, (%).

Table 1 (continued)

Variable	Definition and rationales	Measure and unit
Unemployment ratio*	The unemployment rate is the percentage of the total labour force that is unemployed but actively seeking employment and willing to work. Areas with high unemployment rates are possibly related to the high rates of CP-MVCs (Currie & Hotz, 2004; Kendrick, 1993).	Ratio of unemployment in the total population = unemployed population/population 15–65 years old × 100, (%).

* Indicators marked identified by generalized liner model with the highest predictive value; NMT = non-motorized travel; CP-MVCs = child pedestrian/motor vehicle collisions

Generalized Linear Model

Poisson and negative binomial are two GLMs developed to investigate the association between CP-MVCs and explanatory factors. Poisson distribution is defined by

$$p(y_i) = \frac{\exp(-\lambda_i) \times \lambda_i^{y_i}}{y_i!}, \tag{1}$$

With mean ($E(y_i|x)$) and variance ($E(y_i|x)$)

$$E(y_i|x) = Var(y_i|x) = \lambda_i = \exp(\beta x_i). \tag{2}$$

Where y_i is dependent variable, $\boldsymbol{x}_i$ are explanatory variables with $i = 1,2,\ldots,N$ and $\boldsymbol{\beta}$ is vector of estimable parameters. Assuming equality of mean and variance in Poisson model ends with biased estimation standard errors. In negative binomial model a random term is added to mean function which allows variance to be greater than mean. With this assumption, Eq. (2) can be rewritten as

$$\lambda_i = \exp(\beta x_i + e_i), \tag{3}$$

Where follows a gamma distribution with mean 1 and variance α. With this reparameterization, the variance is

$$Var(y_i|x) = E(y_i|x) + \alpha E(y_i|x)^2 \tag{4}$$

The SAS software (version 9.4) was used to perform GLMs to investigate the association between CP-MVCs and explanatory factors. Final combination of explanatory variables was obtained using backward elimination strategy where confidence interval was equal to 95%. The Condition information matrix and Variance Inflation Factor (VIF) were checked for the obtained models to assess the multicollinearity. There was no multicollinearity according to Model multicollinearity analyses results.

Spatial Analysis

Explanatory variables associated with CP-MVCs in GLMs were included in spatial analysis using GWR methods, which is a local form of GLMs that takes the effects of non-stationary variables (unpredictable variables which cannot be modelled) into account in modelling the local associations between explanatory variables and dependent variable. Unlike GLMs, where a single coefficient is estimated for each explanatory variable, application of GWR method enables local variations (over space) in the estimation of local standardized coefficients or local t-values (Litman, 2012). It was adopted to explore the local spatial heterogeneity of the associations between CP-MVCs (as dependent variable) and explanatory variables (Table 1). This method generates a separate regression equation for each specific location. Therefore, in contrast to GLMs which shows the association between neighbourhood characteristic and CP-MVCs' in the whole data, GWR shows the impact of geographical location. In the current study, Poisson and negative binomial models were extended to geographical form. Equations 5 and 6 show the functional form of geographical models (Da Silva & Rodrigues, 2014; Oluwajana et al., 2022)

$$y_i \sim Poisson(\exp(\sum_k \beta_k(u_i, v_i)x_{ik}) \tag{5}$$

$$y_i \sim NB(\exp(\sum_k \beta_k\left(u_i, v_i\right)x_{ik}, \alpha) \tag{6}$$

In this specification, (u_i, v_i) is coordination of each unit and α is global overdispersion parameter. More specifications of geographical model are presented by other studies (Da Silva & Rodrigues, 2014; Oluwajana et al., 2022; Soroori et al., 2021).

To approve model validity, the spatial independency of GWR residuals was assessed with the Global Moran's I to ensure that the GWR residuals values are spatially random. ArcGIS 10.8 (ESRI, Redlands, CA, USA) was used for mapping the results of the GWR modelling.

Results

A total of 7,028 car collisions' with involvement of pedestrians aged 18 or younger were recorded in the study period with an average of 668 collisions' per 100,000 people in this age group. The collisions' were heterogeneously distributed, with the highest density (173.6 cases per km^2) at the eastern parts of the city and the lowest (less than 1 case per km^2) at the western parts (Fig. 1). Of the total population in the study area, 6.3% were illiterate and 4.6% unemployed, while an average of two school/education centres available per 1,000 people $\leq$ 18 years. There were 140 pedestrian bridges in the city (less than two bridges for each km road length (Table 2).

Based on the results of the study, four models were developed, including Poisson and negative binomial models and their geographical forms. The geographically weighted Poisson regression (GWPR) model was found to provide the best prediction with the highest value of Pearson's product moment correlation coefficient (r) (Tables 3 and 4). Therefore, it can be concluded that applying GWPR is an effective method for modelling Pedestrian-Motor Vehicle Collisions, compared to GWNBR, a more complex model with fewer significant variables and weaker prediction performance.

A comparison of simple and geographical models indicated that geographical forms performed better in terms of both performance (based on AIC) and prediction (based on r), which highlights the importance of considering spatial heterogeneity when modeling collision data. The Poisson model revealed statistically significant associations between eleven explanatory variables and the occurrence of collisions. Among these significant variables, the population of people who are

Table 2 Descriptive statistics of the included factors across the city neighbourhoods

Variable	Mean (SD)	Range
Illiteracy ratio	6.29 (4.80)	0–28.24
Population density	13,286.7 (8,376.8)	0–42,518.5
Population between 0–19	5,503 (4192)	1–27,572
Unemployment ratio	4.58 (1.88)	0–7.88
Land-use mixed	0.002 (0.001)	0–0.006
Open space ratio	0.72 (3.03)	0–27.27
School presence ratio	0.21 (0.20)	0–1.28
Road's area ratio	31.21 (15.13)	0–82.35
Main roads intersection ratio	11.38 (15.61)	0–104.90
Average maximum vehicle speed limit	29.69 (3.93)	15.58–37.62
Public transit supply	34.49 (49.89)	0–270.30
NMT	143.13 (67.57)	0–263.86
Sidewalk area ratio	6.39 (3.50)	0.13–13.86
Sidewalk disconnection index	12.16 (7.96)	0.59–55.49
Pedestrian bridges ratio	1.87 (2.63)	0–11.86
CP-MVCs	668.67 (316.79)	0–1,753

NMT = non-motorized travel; CP-MVCs = Child Pedestrian-Motor Vehicle Collisions; $Y^{=}$ CP-MVCs used as the dependent variable

Table 3 Results obtained from global and local Poisson models

variables	Poisson	GWPR						
		Mean	STD	Min	Max	Q_1	Q_2	Q_3
Intercept	-6.911	-4.64	1.8	-8.84	-0.56	-5.99	4.45	-3.44
Ln (Pop0_19)	1.107	0.85	0.19	0.39	1.28	0.7	0.84	1
Road area ratio	0.0077	-0.00016	0.014	-0.03	0.029	-0.01	-0.0012	0.01
Main road intersection ratio	0.0029	0.0011	0.004	-0.016	0.008	-0.000035	0.0022	0.0038
Average maximum speed limit	0.21	0.01	0.019	-0.0036	0.06	-0.0015	0.005	0.02
NMT proportion	0.001	0.0011	0.003	-0.0036	0.01	-0.0008	0.00038	0.0025
Sidewalk area ratio	0.002	0.014	0.026	-0.06	0.096	0.0011	0.0065	0.017
Sidewalk disconnection index	0.012	-0.0017	0.019	-0.065	0.049	-0.011	0.001	0.011
School presence ratio	0.76	0.52	0.79	-1.22	2.30	0.12	0.49	0.95
Open space ratio	0.055	0.02	0.115	-0.39	0.2	-0.02	0.05	0.072
Illiteracy ratio	0.064	0.0108	0.078	-0.034	0.340	0.059	0.091	0.17
Unemployment ratio	-0.11	-0.30	0.096	-0.232	0.251	-0.092	-0.035	0.018
AIC	2351		1379					
AICc	2353		1511					
r	0.94		0.97					

*Confidence of all presented explanatory variables are 95 precent

Table 4 Results obtained from global and local Negative binomial models

Variables	NB	GWNBR						
		Mean	STD	Min	Max	Q_1	Q_2	Q_3
Intercept	-3.41	-3.61	1.171	-6.07	-1.44	-4.29	-3.629	-2.69
Ln (Pop0_19)	0.767	0.78	0.123	0.56	1	0.68	0.80	0.88
School presence ratio	0.561	0.71	0.297	0.11	1.29	0.45	0.66	1.02
Land-use mix	81.35	24.91	48.58	-64.63	193.87	-10.18	16.27	54.82
Open space ratio	0.04	0.05	0.065	-0.1	0.20	0.03	0.05	0.08
Illiteracy ratio	0.07	0.09	0.04	0.001	0.15	0.06	0.9	0.12
Unemployment ratio	-0.074	-0.059	0.045	-0.18	-0.001	-0.088	-0.04	-0.024
Over dispersion parameters	0.24				0.23			
AIC	1468				1412			
AICc	1469				1422			
r	0.73				0.88			

*Confidence of all presented explanatory variables are 95 precent

18 years old or younger, introduced as an exposure variable, showed the highest association with the collision counts (Table 3).

Mapping variables with significant associations revealed similarities and differences across distribution of variables (Fig. 2). Most neighbourhoods with more than average population ages 0–19 had also more than average of illiteracy ratio. In contrast, most neighbourhoods with a more than average unemployment ratio were in the parts of the city characterized by less than average population ages 0–19 and illiteracy ratio. There was also a slight similarity between the patterns of school presence and sidewalk area ratios. Open space ratio was relatively similar in most neighbourhoods. Non-motorized travel activity was more prevalent in the western part of the city, where sidewalk area ratio was higher and average maximum speed limit was lower.

The analysis of spatial variation regarding the association between CP-MVCs and the significant explanatory variables indicated notable variations in local parameters for all variables (Table 3). Nevertheless, the variations varied depending on the variable under consideration. School presence ratio and population ages 0–19 exhibited the highest spatial variations, as indicated by the highest standard deviations.

All four models were reasonably fit and the correlation between predicted and observed value (r) was high value for all models. The model performance was high in the majority of included neighbourhoods (93 out of 175), where the standardised residuals were between -0.5 and $+0.5$ (Fig. 3). The random spatial patterns of GWPR residuals indicated that the GWPR model was correctly specified, and the results of the model could be trusted for examining the spatial relationships between the dependent and explanatory variables (Moran's I= -0.02, Z-score= -0.27241 and P-value $>$ 0.05). The random residuals show that explanatory variables describe the relationship so thoroughly that only random error remains. However, the non-random patterns of the residuals signify that the explanatory variables are missing something.

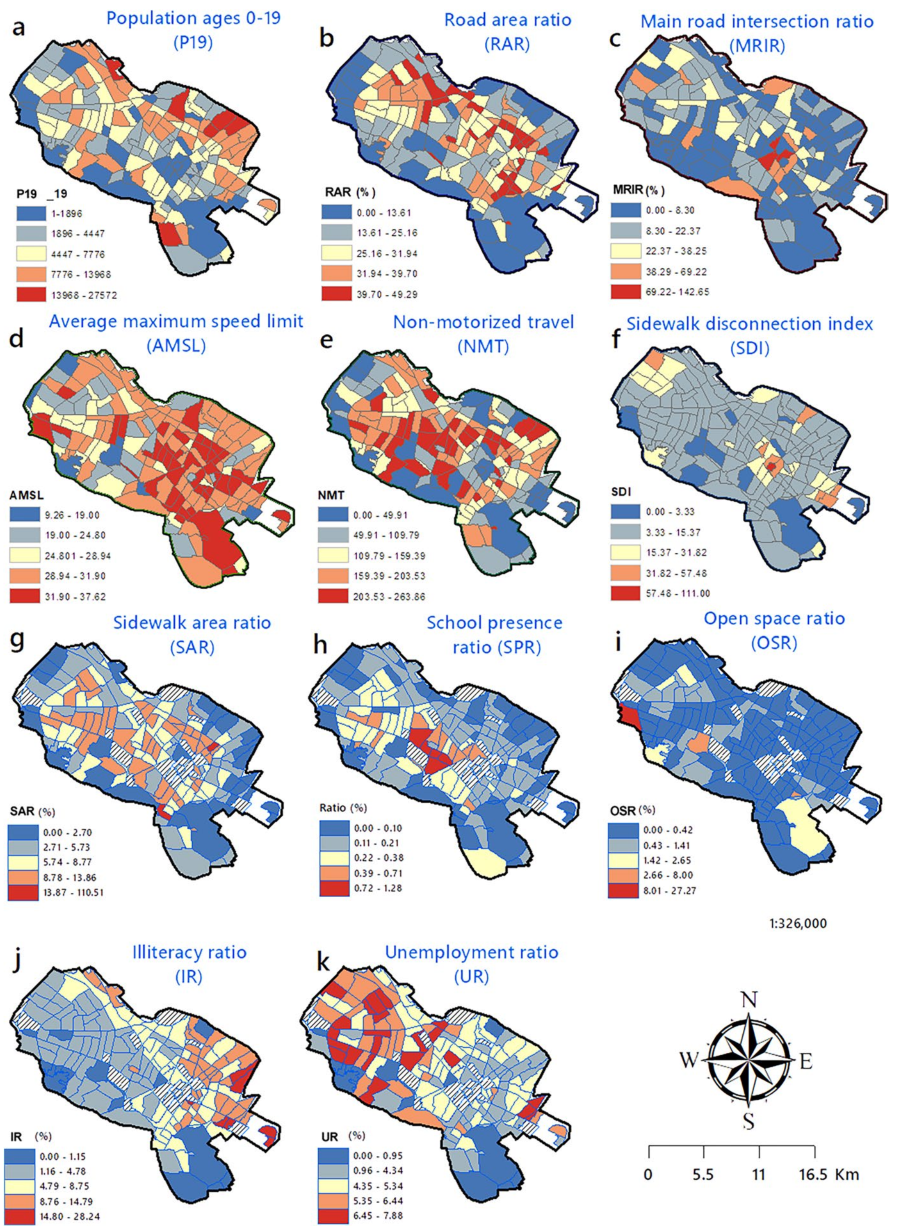

Fig. 2 Spatial distribution maps of the most important explanatory variables associated with child pedestrian-motor vehicle collisions in Mashhad neighbourhoods. Red and orange colours show areas where at least one of the significant factors are above average; Maps created by authors in ArcGIS v.10.8

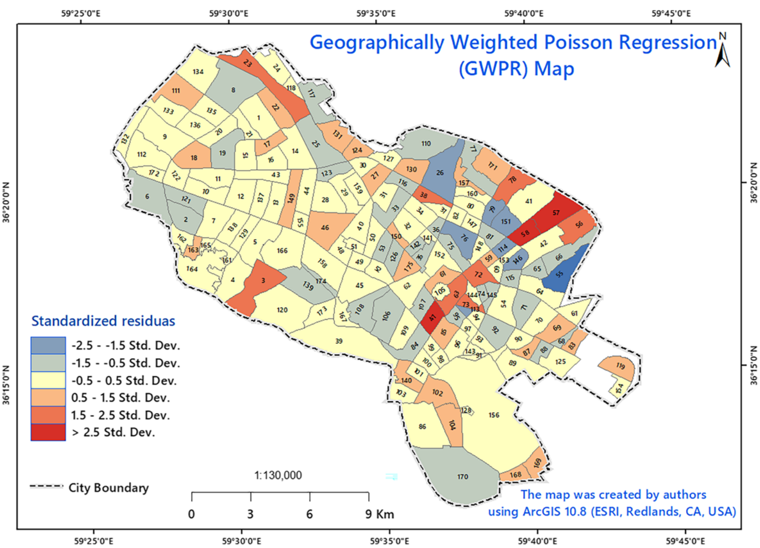

Fig. 3 Choropleth map of GWPR residuals. Yellow shade indicates neighbourhoods where the model performance was high; steel-blue colour shows low and crimson high values of the standardized residuals distributed in Mashhad neighbourhoods;

Employing local spatial analysis techniques to map the associations between CP-MVCs and explanatory variables across neighbourhoods has provided a higher level of sophistication. The findings demonstrate that the associations identified by simple GLM regressions for the entire study area are not consistent across all neighbourhoods (see Fig. 4). Each variable can act as a determinant of CP-MVC patterns in some but not all neighbourhoods. For instance, the sidewalk area ratio was significantly associated with CP-MVCs in only 40 out of 175 neighbourhoods, and some neighbourhoods did not show any association between the variables and CP-MVCs. Nevertheless, some variables exhibited similar patterns, such as the overlap between the associations of road area ratio and sidewalk area ratio, non-motorized trips, open space ratio, and school presence.

The exposure variable of population ages 0–19 had a significant effect on increasing the occurrence of crashes. Additionally, the average maximum limits, sidewalk area ratio, and school presence ratio were identified as three other variables that associate with an increase in crash counts. Furthermore, it is worth noting that although the majority of neighbourhoods had a low level of open space ratio, there was a noticeable spatial variation that transitioned from positive in the eastern neighbourhoods to negative in the central and northern regions of the city.

Upon comparing Figs. 2 and 4, a noticeable observation is that neighborhoods with significant associations were not necessarily the ones with highly variable ratios. It was found that the population of young people has a positive correlation with CP-MVC in all neighborhoods, irrespective of whether they exceed the

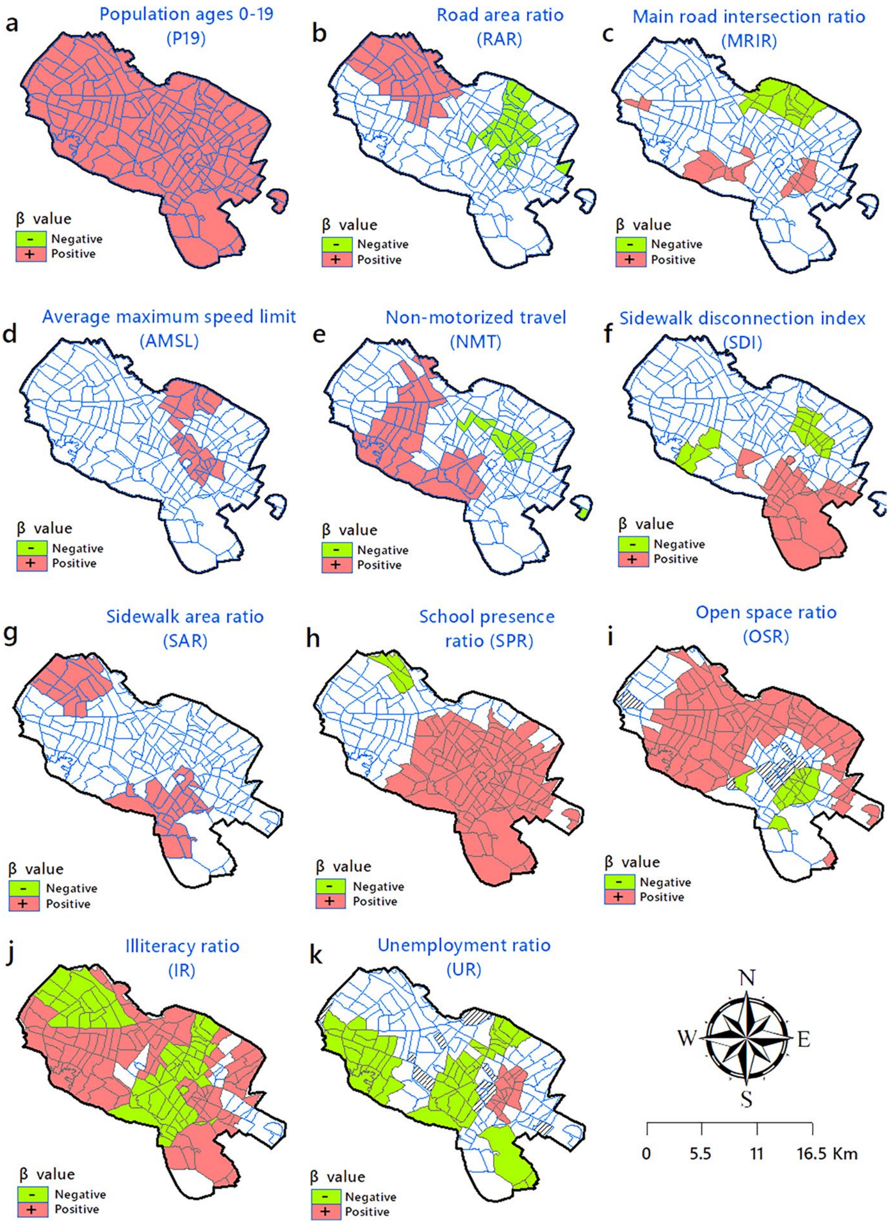

Fig. 4 Choropleth maps of the distribution of the local coefficient (β) based on GWPR modelling collisions calculated for the pedestrian age group of 0–18 years old in Mashhad

average or not. The results for other variables do not show exact similar patterns between neighborhoods with values exceeding the average and spatially significant associations.

Discussion

This study aimed to explore the relationship between neighbourhood characteristics and CP-MVCs at the urban neighbourhood level as well as at the level of the entire metropolitan area. The results indicated that although certain socio-economic and built environmental factors exhibited a significant association with the CP-MVCs in the overall study area, associations at the local geographical level often differed among neighbourhoods. Therefore, in some neighbourhoods, there was no local association observed between the significant factors that had a significant association with CP-MVCs. Among the different models, the GWPR obtained the best results for modelling Pedestrian-Motor Vehicle Collisions. The same result was obtained by another study (Soroori et al., 2021) that illustrated among GWR forms of Poisson and negative binomial model, the model developed using Poisson distribution outperform the other models.

Our results are of significant importance as they suggest that the relationship between neighbourhood characteristics and CP-MVCs rate in models utilizing a large study area may not be as robust, which limits their ability to investigate local associations across smaller areas. As numerous variables may be associated with the incidence rate of car collisions, modeling some of these variables may not be straightforward due to their non-stationary nature. Nonetheless, approaches that take the effect of these variables into account can increase the robustness of the results. GWR is a technique that enables regression coefficients to vary across individual locations, capturing the effects of non-stationarity and revealing variations in the importance of variables across the study area (Song et al., 2017). Consistent with the findings of the present study, previous research has demonstrated that GWR exhibits higher performance compared to non-spatial models such as linear regression when examining the effect of socio-economic and built-environmental factors on CP-MVC incidence (Almasi et al., 2021; Xiao et al., 2019).

Taking into account the effect of local factors in GWR highlights that the most significant determinants of CP-MVCs in a neighborhood may differ from those of another neighborhood. Therefore, implementing a general policy to address collisions across the entire area may not be as effective as taking into consideration the local determinants. Additionally, when the performance of the GWR model varies across neighborhoods, it is recommended that future studies focus on the neighborhoods with lower performance to identify the significant explanatory variables. These neighborhoods may have urban structures or characteristics that are distinct from the rest of the city, requiring a unique investigative approach, such as conducting case studies.

The importance of local characteristics in different geographical locations can help to explain the current inconsistencies in the literature about the role of neighbourhood characteristics in the global incidence of CP-MVCs. The positive and significant impact of the population ages 0–19 on crashes across all neighbourhoods suggests that this age group is associated with an increased risk of collisions, likely due to longer trips and a greater likelihood of choosing non-motorized modes of transportation such as biking. This is further supported by significant parameters estimated for non-motorized travel, which also contribute to an increase in collisions in most neighbourhoods. Another variable that can contribute to an increase in crash occurrence is a disconnected sidewalk path, as it provides space for more traffic conflicts. Graham et al. (2013) showed a positive association between CP-MVCs and population density in various areas of seven major British cities. They also examined the association between CP-MVCs and the deprivation levels of areas and found that the collision rate was typically greater than 10 times higher in the most deprived areas than in the least deprived areas. However, deprivation was more strongly associated with collisions than population density (Graham et al., 2013).

Our study findings are consistent with previous research demonstrating an association between CP-MVCs and certain neighbourhood characteristics such as illiteracy, school presence, sidewalk area, and opening space ratios (Clifton et al., 2009; Cottrill & Thakuriah, 2010; Kim, 2019; Tiwari, 1999). Low-income areas have been found to have fewer traffic safety features, such as traffic calming measures, and a higher incidence of CP-MVCs (Rothman et al., 2020).

Walking to school is an important factor contributing to CP-MVCs for school children, and the number of schools in an area may increase the risk of collisions due to a higher rate of walking to school (Rothman et al., 2017; Tetali et al., 2016). This can help explain the positive relationship between CP-MVCs and sidewalk area ratio. However, the quality of walking pathways in each neighbourhood can also be a significant determinant of the collision counts, potentially contributing to the discrepancy in the local association of school presence ratio and CP-MVCs across neighbourhoods (Litman, 2012).

Open spaces and parks in an area can reduce the incidence of CP-MVCs through several mechanisms (Egorov et al., 2016). Pedestrians, such as students walking to school, often use these spaces as safe routes. Additionally, open spaces and parks can reduce traffic speeds, which is important in decreasing vehicle collisions (Egorov et al., 2016).

This study is subject to some limitations. A key limitation is the use of census statistics to calculate population figures, which only account for the resident's population of each area. To improve the accuracy of population estimates, it would be ideal to use data on non-fixed populations (including tourists), but such information is currently unavailable at both the total and neighbourhood level. Additionally, the total population of children was used to scale the dependent variable. However, it would be more appropriate to use the number of children who walk to school or use non-motorized modes of transportation as the denominator, as this provides a more accurate measure of exposure. Unfortunately, such data are often difficult to obtain, making it challenging to accurately assess the risk of CP-MVCs in relation to active transportation among children.

Conclusions

This study shows that socio-demographic and built-environment factors play a significant role in the incidence of child pedestrian-motor vehicle collision. However, the association is not consistent across the whole area, and unknown factors may also have an impact. To generate specific knowledge for each uniform area, geographical regression is necessary to find the local determinants at smaller area levels such as the neighbourhood. This approach allows policymakers and planners to tailor effective health promotion strategies for each specific area, which can ultimately lead to a reduction in child pedestrian-motor vehicle collision incidence. Further research is needed to identify these local determinants and to develop targeted interventions.

Acknowledgements We express our gratitude to Mashhad Municipality for their provision of spatial data. Additionally, we extend our appreciation to Mashhad University of Medical Sciences for their provision of data on Children Pedestrian Road Traffic Collisions.

Data Availability The data is available by request from the corresponding author.

Declarations

Conflicting Interests The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Almasi, S. A., Behnood, H. R., & Arvin, R. (2021). Pedestrian crash exposure analysis using alternative geographically weighted regression models. *Journal of Advanced Transportation*, *2021*, 1–13.

Asare, I. O., & Mensah, A. C. (2020). Crash severity modelling using ordinal logistic regression approach. *International Journal of Injury Control and Safety Promotion, 27*(4), 412–419.

Behzadnia, S., & Shahmohammadi, S. (2016). Road traffic injuries among iranian children and adolescents: An epidemiological review. *Journal of Pediatrics Review, 4*(1), 0.

Bennet, S. A., & Yiannakoulias, N. (2015). Motor-vehicle collisions involving child pedestrians at intersection and mid-block locations. *Accident Analysis & Prevention, 78*, 94–103.

Peden, M., Oyegbite, K., Ozanne-Smith, J., Hyder, A. A., Branche, C., Rahman, A. K. M. F., Rivara, F., Bartolomeos, K., editors. World Report on Child Injury Prevention. Geneva: World Health Organization; 2008. PMID: 26269872.

Briz-Redón, Á, Martínez-Ruiz, F., & Montes, F. (2019). Estimating the occurrence of traffic accidents near school locations: A case study from Valencia (Spain) including several approaches. *Accident Analysis & Prevention, 132*, 105237.

Chakravarthy, B., Anderson, C. L., Ludlow, J., Lotfipour, S., & Vaca, F. E. (2010). The relationship of pedestrian injuries to socioeconomic characteristics in a large Southern California County. *Traffic Injury Prevention, 11*(5), 508–513.

Clifton, K. J., Burnier, C. V., & Akar, G. (2009). Severity of injury resulting from pedestrian–vehicle crashes: What can we learn from examining the built environment? *Transportation Research Part D: Transport and Environment, 14*(6), 425–436.

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.

Cottrill, C. D., & Thakuriah, P. V. (2010). Evaluating pedestrian crashes in areas with high low-income or minority populations. *Accident Analysis & Prevention, 42*(6), 1718–1728.

Currie, J., & Hotz, V. J. (2004). Accidents will happen?: Unintentional childhood injuries and the effects of child care regulations. *Journal of Health Economics, 23*(1), 25–59.

Curtis, J. W. (2017). Spatial distribution of child pedestrian injuries along census tract boundaries: Implications for identifying area-based correlates. *PLoS One, 12*(6), e0179331.

Da Silva, A. R., & Rodrigues, T. C. V. (2014). Geographically weighted negative binomial regression—incorporating overdispersion. *Statistics and Computing, 24*, 769–783.

Egorov, A. I., Mudu, P., Braubach, M., & Martuzzi, M. (2016). *Urban green spaces and health-A review of evidence*. WHO Regional Office for Europe.

Elias, W., & Shiftan, Y. (2014). Analyzing and modeling risk exposure of pedestrian children to involvement in car crashes. *Accident Analysis & Prevention, 62*, 397–405.

Graham, D. J., McCoy, E. J., & Stephens, D. A. (2013). Quantifying the effect of area deprivation on child pedestrian casualties by using longitudinal mixed models to adjust for confounding, interference and spatial dependence. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 176*(4), 931–950.

Hagel, B. E., Macpherson, A., Howard, A., Fuselli, P., Cloutier, M. S., Winters, M., Richmond, S. A., Rothman, L., Belton, K., & Buliung, R. (2019). The built environment and active transportation safety in children and youth: A study protocol. *BMC Public Health, 19*, 1–13.

Hwang, J., Joh, K., & Woo, A. (2017). Social inequalities in child pedestrian traffic injuries: Differences in neighborhood built environments near schools in Austin, TX, USA. *Journal of Transport & Health, 6*, 40–49.

Jamshidi, E., Moradi, A., & Majdzadeh, R. (2017). Environmental risk factors contributing to traffic accidents in children: A case-control study. *International Journal of Injury Control and Safety Promotion, 24*(3), 338–344.

Kendrick, D. (1993). Prevention of pedestrian accidents. *Archives of Disease in Childhood, 68*(5), 669–672.

Kim, D. (2019). The transportation safety of elderly pedestrians: Modeling contributing factors to elderly pedestrian collisions. *Accident Analysis & Prevention, 131*, 268–274.

Kuşkapan, E., Sahraei, M. A., Çodur, M. K., & Çodur, M. Y. (2022). Pedestrian safety at signalized intersections: Spatial and machine learning approaches. *Journal of Transport & Health, 24*, 101322.

Li, Q., Alonge, O., & Hyder, A. A. (2016). Children and road traffic injuries: Can't the world do better? *Archives of Disease in Childhood, 101*(11), 1063–1070.

Litman, T. (2012). *Evaluating non-motorized transportation benefits and costs*. Victoria Transport Policy Institute Victoria.

Ma, C., Hao, W., Xiang, W., & Yan, W. (2018). The impact of aggressive driving behavior on driver-injury severity at highway-rail grade crossings accidents. *Journal of Advanced Transportation*, *2018*. https://doi.org/10.1155/2018/9841498

Mashhad Traffic and Transportation organisation (2020). *Annual Reports of Transportation- Annual Report 2020. Retrieved from Mashha*. https://traffic.mashhad.ir/. Accessed on 1 March 2023.

Miranda-Moreno, L. F., Morency, P., & El-Geneidy, A. M. (2011). The link between built environment, pedestrian activity and pedestrian–vehicle collision occurrence at signalized intersections. *Accident Analysis & Prevention, 43*(5), 1624–1634.

Muazir, S., & Hsieh, H. C. (2019). Urban network in strategic areas in Indonesia case study: Sambas Regency, West Kalimantan. *Journal of Design and Built Environment, 19*(2), 14–29.

Nakaya, T., Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2005). Geographically weighted poisson regression for disease association mapping. *Statistics in Medicine, 24*(17), 2695–2717.

Oluwajana, S. D., Park, P. Y., & Cavalho, T. (2022). Macro-level collision prediction using geographically weighted negative binomial regression. *Journal of Transportation Safety & Security, 14*(7), 1085–1120.

Part, D. (2010). *Highway safety manual* (p. 19192). American Association of State Highway and Transportation Officials.

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.

Pernica, J. M., LeBlanc, J. C., Soto-Castellares, G., Donroe, J., Carhuancho-Meza, B. A., Rainham, D. G. C., & Gilman, R. H. (2012). Risk factors predisposing to pedestrian road traffic injury in children living in Lima, Peru: A case–control study. *Archives of Disease in Childhood, 97*(8), 709–713.

QGIS software (2020). *QGIS: A Free and Open Source Geographic Information System*. https://qgis.org/en/site/. Accessed on 10 December 2022.

Rothman, L., Howard, A., Buliung, R., Macarthur, C., Richmond, S. A., & Macpherson, A. (2017). School environments and social risk factors for child pedestrian-motor vehicle collisions: A case-control study. *Accident Analysis & Prevention, 98*, 252–258.

Rothman, L., Cloutier, M. S., Manaugh, K., Howard, A. W., Macpherson, A. K., & Macarthur, C. (2020). Spatial distribution of roadway environment features related to child pedestrian safety by census tract income in Toronto, Canada. *Injury Prevention, 26*(3), 229–233.

Rothman, L., Schwartz, N., Cloutier, M. S., Winters, M., Macarthur, C., Hagel, B. E., Macpherson, A. K., El Amiri, N., Fuselli, P., & Howard, A. W. (2022). Child pedestrian and cyclist injuries, and the built and social environment across canadian cities: The child active transportation safety and the Environment Study (CHASE). *Injury Prevention, 28*(4), 311–317.

Rus, M., Tasențe, T., & Rus, C. V. (2020). Preventive Behaviors and Attitudes of the Children as Active Participants in Road Traffic: Array. Acta Universitatis Danubius. Administratio, 11(1). Retrieved from https://dj.univ-danubius.ro/index.php/AUDA/article/view/138

Schuurman, N., Walker, B. B., Swanlund, D., Amram, O., & Yanchar, N. L. (2020). Qualitative field observation of pedestrian injury hotspots: a mixed-methods approach for developing built-and socioeconomic-environmental risk signatures. *International Journal of Environmental Research and Public Health*, *17*(6), 2066.

Shabanikiya, H., Hashtarkhani, S., Bergquist, R., Bagheri, N., VafaeiNejad, R., Amiri-Gholanlou, M., Akbari, T., & Kiani, B. (2020). Multiple-scale spatial analysis of paediatric, pedestrian road traffic injuries in a major city in North-Eastern Iran 2015–2019. *BMC Public Health, 20*(1), 1–11.

Silverman, J. D., Hutchison, M. G., & Cusimano, M. D. (2013). Association between neighbourhood marginalization and pedestrian and cyclist collisions in Toronto intersections. *Canadian Journal of Public Health, 104*, e405–e409.

Song, C., Kwan, M. P., & Zhu, J. (2017). Modeling fire occurrence at the city scale: A comparison between geographically weighted regression and global linear regression. *International Journal of Environmental Research and Public Health, 14*(4), 396.

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 and Security, 13*(9), 1000–1024. https://doi.org/10.1080/19439962.2020.1712671

Stevenson, M. (2006). Building safer environments: Injury, safety, and our surroundings. *Injury Prevention, 12*(1), 1–3.

Sun, Z., Wang, J., Chen, Y., & Lu, H. (2018). Influence factors on injury severity of traffic accidents and differences in urban functional zones: The empirical analysis of Beijing. *International Journal of Environmental Research and Public Health, 15*(12), 2722.

Tetali, S., Edwards, P., Murthy, G. V. S., & Roberts, I. (2016). Road traffic injuries to children during the school commute in Hyderabad, India: Cross-sectional survey. *Injury Prevention, 22*(3), 171–175.

Tiwari, G. (1999). Towards a sustainable urban transport system: Planning for non-motorized vehicles in cities. *Transport and Communications Bulletin for Asia and the Pacific, 68*, 49–66.

Wang, H., Schwebel, D. C., Tan, D., Shi, L., & Miao, L. (2018). Gender differences in children's pedestrian behaviors: Developmental effects. *Journal of Safety Research, 67*, 127–133.

World Health Organization. (2018). *Global status report on road safety 2018*. available at https://www.who.int/publications/i/item/9789241565684. Accessed on 1 March 2023.

Xiao, D., Xu, X., & Duan, L. (2019). Spatial-temporal analysis of injury severity with geographically weighted panel logistic regression model. *Journal of Advanced Transportation*, *2019*. https://doi.org/10.1155/2019/8521649

Xie, B., An, Z., Zheng, Y., & Li, Z. (2019). Incorporating transportation safety into land use planning: Pre-assessment of land use conversion effects on severe crashes in urban China. *Applied Geography, 103*, 1–11.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Emad Soroori[1] · **Behzad Kiani**[2] · **Soraya Ghasemi**[1] · **Alireza Mohammadi**[3] · **Hamidreza Shabanikiya**[4] · **Robert Bergquist**[5] · **Fatemeh Kiani**[6] · **Hossein Tabatabaei-Jafari**[7]

Emad Soroori
emad.soroori@mail.um.ac.ir

Soraya Ghasemi
sori.qasemi@gmail.com

Alireza Mohammadi
a.mohammadi@uma.ac.ir

Hamidreza Shabanikiya
ShabaniKiaHR@mums.ac.ir

Robert Bergquist
Robert.Bergquist@outlook.com

Fatemeh Kiani
KianiF981@mums.ac.ir

Hossein Tabatabaei-Jafari
Hossein.Tabatabaeijafari@canberra.edu.au

1 Department of Civil Engineering, Ferdowsi University of Mashhad, Mashhad, Iran

2 School of Public Health, University of Montreal, Montreal, Canada

3 Department of Geography and Urban Planning, Faculty of Social Sciences, University of Mohaghegh Ardabili, Ardabil, Iran

4 Social Determinants of Health Research Centre, Mashhad University of Medical Sciences, Mashhad, Iran

5 Ingerod, Brastad, Sweden (formerly with the UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases, World Health Organization), Geneva, Switzerland

6 Department of Health Information Technology and Management, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

7 Mental Health Policy Unit, Health Research Institute, The University of Canberra, Canberra, Australia