Towards activity-based exposure measures in spatial analysis of pedestrian–motor vehicle crashes

Ni Dong[1,2,3], Fanyu Meng[4,5], Jie Zhang[1,2], S.C. Wong[6], Pengpeng Xu[6*]

[1] School of Transportation and Logistics, Southwest Jiaotong University, Chengdu, Sichuan, China

[2] National United Engineering Laboratory of Integrated and Intelligent Transportation, Southwest Jiaotong University, Chengdu, Sichuan, China

[3] Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington, US

[4] Academy for Advanced Interdisciplinary Studies, Southern University of Science and Technology, Shenzhen, China

[5] Department of Statistics and Data Science, Southern University of Science and Technology, Shenzhen, China

[6] Department of Civil Engineering, The University of Hong Kong, Hong Kong, China

* **Correspondence to** Pengpeng Xu

Room LG-208, Composite Building, The University of Hong Kong, Pokfulam Road, Hong Kong, China

Email: pengpengxu@yeah.net; Tel: 86-18570366096

Abstract

Background: Although numerous efforts have been devoted to exploring the effects of area-wide factors on the frequency of pedestrian crashes in neighborhoods over the past two decades, existing studies have largely failed to provide a full picture of the factors that contribute to the incidence of zonal pedestrian crashes, due to the unavailability of reliable exposure data and use of less sound analytical methods.

Methods: Based on a crowdsourced dataset in Hong Kong, we first proposed a procedure to extract pedestrian trajectories from travel-diary survey data. We then aggregated these data to 209 neighborhoods and developed a Bayesian spatially varying coefficients model to investigate the spatially non-stationary relationships between the number of pedestrian–motor vehicle (PMV) crashes and related risk factors. To dissect the role of pedestrian exposure, the estimated coefficients of models with population, walking trips, walking time, and walking distance as the measure of pedestrian exposure were presented and compared.

Results: Our results indicated substantial inconsistencies in the effects of several risk factors between the models of population and activity-based exposure measures. The model using walking trips as the measure of pedestrian exposure had the best goodness-of-fit. We also provided new insights that in addition to the unstructured variability, heterogeneity in the effects of explanatory variables on the frequency of PMV crashes could also arise from the spatially correlated effects. After adjusting for vehicle volume and pedestrian activity, road density, intersection density, bus stop density, and the number of parking lots were found to be positively associated with PMV crash frequency, whereas the percentage of motorways and median monthly income had negative associations with the risk of PMV crashes.

Conclusions: The use of population or population density as a surrogate for pedestrian exposure when modeling the frequency of zonal pedestrian crashes is expected to produce biased estimations and invalid inferences. Spatial heterogeneity should also not be negligible when modeling pedestrian crashes involving contiguous spatial units.

Keywords: Pedestrian safety; Crash frequency; Activity-based exposure measures; Spatial correlation; Spatial heterogeneity

1. Introduction

Of the active models of transport, walking has the advantages of reducing traffic congestion, greenhouse gas emissions, and traffic noise. Around the world, walking is also a popular physical and recreational activity, particularly among children and the elderly. Indeed, with the increasing number of short-distance trips, growing levels of traffic congestion, and higher parking costs in metropolitan areas, people are increasingly encouraged to walk more as a viable and sustainable mode of transport (Maibach et al., 2009).

Despite the well-documented benefits of walking, pedestrians are among the most vulnerable road users with substantially higher risks of fatality and injury than motorists (Retting et al., 2003; Zegeer and Bushell, 2012; Stoker et al., 2015). This is especially the case in urban areas with a dense population, where walking is indispensable in ensuring affordable and adequate mobilities for most local residents. An in-depth understanding of the factors that contribute to pedestrian crashes is therefore imperative if walking is promoted as a safe and attractive mode of transport. Improvements in safety would also encourage more people to walk on a regular basis for daily travel, thereby fostering a more livable community.

Over the past two decades, modeling pedestrian crashes involving contiguous spatial units, such as census tracts and traffic analysis zones, has attracted extensive research interest from traffic safety analysts (see Table A1). This allows local authorities to identify the clustering pattern of pedestrian crashes, to better determine the zonal factors that contribute to the incidence of pedestrian crashes, and to recommend area-wide countermeasures

Previous studies have suggested that the number of pedestrian crashes in a neighborhood increased significantly with the increases in traffic volume (LaScala et al., 2000; Loukaitou-Sideris et al., 2007; Wier et al., 2009; Cottrill and Thakuriah, 2010; Dumbaugh and Zhang, 2013; Wang and Kockelman, 2013; DiMaggio, 2015; Lee et al., 2015; Cai et al., 2016, 2017a; Guo et al., 2017; Osama and Sayed, 2017; Tasic et al., 2017; Xie et al., 2017; Goel et al., 2018) and pedestrian volume (Sebert Kuhlmann et al., 2009; Wang and Kockelman, 2013; Cai et al., 2016; Chen and Zhou, 2016; Guo et al., 2017; Osama and Sayed, 2017; Tasic et al., 2017; Lee et al., 2019; Sze et al., 2019). However, unlike vehicle volume which is readily obtained from counting stations, pedestrian volume is mostly surrogated as resident population (LaScala et al., 2000; Noland and Quddus, 2004; Wier et al., 2009; Chakravarthy et al., 2010; Ha and Thill, 2011; Ukkusuri et al., 2011, 2012; Siddiqui et al., 2012; Dumbaugh and Zhang, 2013; Graham et al., 2013; Noland et al., 2013; DiMaggio, 2015; Lee et al., 2015; Wang et al., 2016; Goel et al., 2018; Rothman et al., 2020) or population density (Graham and Glaister, 2003; Priyantha Wedagama et al., 2006; Loukaitou-Sideris et al., 2007; Cottrill and Thakuriah, 2010; Cai et al., 2017a) due to data unavailability. This improper representation of pedestrian exposure, however, very likely leads to biased estimations and incorrect inferences (Steinbach et al., 2014).

Although a limited number of studies have recently used walking trips (Sebert Kuhlmann et al., 2009; Delmelle et al., 2012; Cai et al., 2016; Chen and Zhou, 2016; Guo et al., 2017; Osama and Sayed, 2017; Tasic et al., 2017; Ding et al., 2018; Sze et al., 2019), walking miles (Wang and Kockelman, 2013), or walking hours (Lee et al., 2019; Sze et al., 2019) to quantify pedestrian activities, the roles played by various exposure measures in the performance of zonal pedestrian crash-frequency models have not been comprehensively investigated and thus remain largely unknown.

Road-network characteristics, such as intersection density (Graham and Glaister, 2003; Priyantha Wedagama et al., 2006; Guo et al., 2017; Osama and Sayed, 2017; Tasic et al., 2017), intersection type (Ha and Thill, 2011; Ukkusuri et al., 2011, 2012; Cai et al., 2016; Chen and Zhou, 2016; Wang et al., 2016; Tasic et al., 2017; Sze et al., 2019), road density (Graham et al., 2013; Wang et al., 2016; Sze et al., 2019), road function (Graham and Glaister, 2003; Noland and Quddus, 2004; Wier et al., 2009; Ukkusuri et al., 2011, 2012; Dumbaugh and Zhang, 2013; Noland et al., 2013; Jermprapai and Srinivasan, 2014; Cai et al., 2016, 2017a; Wang et al., 2016; Tasic et al., 2017), speed limits (Siddiqui et al., 2012; Lee et al., 2015), sidewalk density (Wang and Kockelman, 2013; Cai et al., 2016; Chen and Zhou, 2016; Cai et al., 2017a), and network topology (Guo et al., 2017; Osama and Sayed, 2017; Tasic et al., 2017) were found to be closely related to the frequency of pedestrian crashes in a neighborhood. Land use was also reported to have a significant influence on the incidence of pedestrian crashes. Specifically, higher percentages of commercial (Priyantha Wedagama et al., 2006; Loukaitou-Sideris et al., 2007; Wier et al., 2009; Ukkusuri et al., 2011; Jermprapai and Srinivasan, 2014), residential (Priyantha Wedagama et al., 2006; Loukaitou-Sideris et al., 2007; Wier et al., 2009; Wang and Kockelman, 2013), and industrial (Ukkusuri et al., 2011, 2012; Jermprapai and Srinivasan, 2014) land-use were associated with an increased likelihood of pedestrian crashes. Similar conclusions hold true for the effect of land-use intensity (Wang et al., 2016). However, inconsistent results were found for the effect of land-use mix, as Wang and Kockelman (2013) reported a significantly negative relationship between mixed land-use and the frequency of pedestrian crashes, whereas Chen and Zhou (2016), Guo et al. (2017), and Xie et al. (2017) drew the opposite conclusion.

In addition, the prevalence of specific facilities, such as bus stops (Ukkusuri et al., 2011; Lee et al., 2015; Xie et al., 2017; Goel et al., 2018), metro stations (Ukkusuri et al., 2011, 2012; Jermprapai and Srinivasan, 2014; Lee et al., 2015), schools (Cottrill and Thakuriah, 2010; Ukkusuri et al., 2011, 2012; Jermprapai and Srinivasan, 2014; Lee et al., 2015), hotels (Lee et al., 2015; Cai et al., 2016), and licensed liquor outlets (Sebert Kuhlmann et al., 2009), was also found to significantly increase the likelihood of pedestrian crashes in neighborhoods.

With respect to socio-economic characteristics, neighborhoods with a denser population (LaScala et al., 2000; Graham and Glaister, 2003; Loukaitou-Sideris et al., 2007; Sebert Kuhlmann et al., 2009; Chakravarthy et al., 2010; Cottrill and Thakuriah, 2010;

Ha and Thill, 2011; Siddiqui et al., 2012; Graham et al., 2013; Wang and Kockelman, 2013; Jermprapai and Srinivasan, 2014; Cai et al., 2016, 2017a), higher proportions of ethnic minorities (Loukaitou-Sideris et al., 2007; Chakravarthy et al., 2010; Ukkusuri et al., 2011; Lee et al., 2019) and less-educated population (LaScala et al., 2000; Chakravarthy et al., 2010; Ukkusuri et al., 2011; Goel et al., 2018), more children (Graham et al., 2013) and elderly people (Ukkusuri et al., 2012; Dumbaugh and Zhang, 2013; Xie et al., 2017; Lee et al., 2019; Sze et al., 2019), lower vehicle ownership per capita (Cottrill and Thakuriah 2010; Noland et al. 2013), a higher unemployment rate (LaScala et al., 2000; Cai et al., 2016; Xie et al., 2017), a higher poverty level (Wier et al., 2009; Chakravarthy et al., 2010; Ha and Thill, 2011; Jermprapai and Srinivasan, 2014; Lee et al., 2015), and a lower household median income (Siddiqui et al., 2012; Noland et al., 2013; Jermprapai and Srinivasan, 2014; Cai et al., 2017a; Rothman et al., 2020) were also associated with more pedestrian crashes.

The relationship between the aforementioned explanatory variables and the number of pedestrian crashes can be established using crash prediction models. Traditional Poisson and negative-binomial models have a strong assumption that their observations should be mutually independent (Lord and Mannering, 2010). This fundamental hypothesis is almost always violated (Mannering and Bhat, 2014), particularly because pedestrian crashes collected in contiguous spatial units usually display spatial correlation (Ziakopoulos and Yannis, 2020). A range of spatial statistical techniques have therefore been used to incorporate this spatial dependence into pedestrian crash-frequency modeling. The Bayesian hierarchical models are most prevalent, in which the spatial correlation is typically modeled via the intrinsic conditional autoregressive (CAR) prior proposed by Besag et al. (1991) at the second level of hierarchy (Sebert Kuhlmann et al., 2009; Siddiqui et al., 2012; Graham et al., 2013; Noland et al., 2013; Wang and Kockelman, 2013; DiMaggio, 2015; Lee et al., 2015; Chen and Zhou, 2016; Wang et al., 2016; Guo et al., 2017; Osama and Sayed, 2017; Zeng et al., 2017, 2019, 2020; Goel et al., 2018; Lee et al., 2019; Wen et al., 2019). Alternative CAR specifications were also introduced by Richardson et al. (1992), Cressie (1993), and Leroux et al. (1999). Lee (2011) made a comprehensive comparison and concluded that the model of Leroux et al. (1999) was the most appealing, because it consistently performed well in the presence of spatial independence and strong spatial correlation.

Although most safety analysts have attempted to tackle the spatial correlation in model residuals, only a relatively limited number of studies have focused specifically on spatial heterogeneity or spatial non-stationarity. Variables do not usually vary constantly across space, and the relationship between pedestrian crashes and related risk factors may not necessarily be fixed across the study area. The capability of accounting for this spatial heterogeneity by allowing parameters to vary spatially holds considerable promise. Although a few studies have used the random-parameters count-data models to account for the heterogeneous effects in pedestrian crash frequency (Ukkusuri et al., 2012; Sze et al., 2019), the regression coefficients in these random-parameters models typically arise independently from some univariate distributions, and no attention is paid to the locations to which the parameters refer. This hypothesis may be inappropriate, particularly in cases where the unobserved factors are correlated through space (Xu and Huang, 2015; Xu et al., 2017). It is thus not surprising that Sze et al. (2019) reported a significantly negative relationship between vehicle volume and the number of pedestrian crashes. This counter-intuitive finding is very likely attributed to the neglect of spatial correlated effects in their random-parameters models. Xu and Huang (2015) therefore advocated the development of a model based on the principle that the estimated parameters on a geographical surface are related to each other with closer values more similar than distant ones.

To address the spatially correlated effects in varying coefficients, one promising approach is the geographically weighted regression model (Fotheringham et al., 2002; Yang et al., 2020; Zhao et al., 2020). This method is similar to local linear models, depending on the calibration of multiple regression models for different geographical entities. Recent studies have empirically demonstrated the superiority of the method, with a substantially improvement in goodness-of-fit and the ability to explore the spatially varying relationships between crash counts and predictive factors (Hadayeghi et al., 2010; Li et al., 2013; Pirdavani et al., 2014; Shariat-Mohaymany et al., 2015; Xu and Huang, 2015; Yao et al., 2015; Bao et al., 2017; Huang et al., 2018; Gomes et al., 2019; Hezaveh et al., 2019; Ariannezhad et al., 2020). An alternatively potential method is the Bayesian spatially varying coefficients (BSVC) models (Xu et al., 2017), which has long been used in statistics to examine the non-constant relationships between variables (Congdon, 1997). Such an approach fits naturally into the Bayesian paradigm, where all parameters are treated as stochastic. Obviously, the BSVC model differs from the geographically weighted regression model in that the former is a single statistical model specified in a hierarchical manner, whereas the latter is an assembly of local spatial regression models. Wheeler and Calder (2007) conducted a series of simulation studies to evaluate the accuracy of regression coefficients in these two types of models. Their evidence suggested that the BSVC model produced more reliable and easily interpreted inferences, thereby providing more flexibility. However, to assume that the regression coefficients are spatially clustered solely is a strong prior belief. In reality, spatial pooling with smoothly varying coefficients over contiguous areas may exhibit over-smoothness (Xu et al., 2017), particularly in the presence of clear discontinuities (Congdon, 2014). In this vein, a robust model with a mechanism to collectively accommodate the global and local smoothing would be favorable.

To summarize, despite that numerous research efforts have been devoted to the development of various predictive models to explore the effects of area-wide factors on the frequency of pedestrian crashes within the past two decades, existing studies have largely failed to provide a full picture of the factors that contribute to the incidence of zonal pedestrian crashes, mainly due to the unavailability of reliable exposure data and use of less sound analytical methods. Based on a comprehensive dataset of 7,103 pedestrian–

motor vehicle (PMV) crashes aggregated in 209 tertiary planning units (TPUs) over a 3-year period in Hong Kong, our study aims to assess the geographical variations of PMV crashes with respect to land-use, road-network attributes, traffic characteristics, the presence of public facilities, and socio-demographic characteristics. Specifically, the objectives of our study are: 1) to propose a procedure to extract pedestrian trajectories from household travel-diary survey data and to dissect the role of various measures of pedestrian exposure (i.e., population, walking trips, walking time, and walking distance) in the performance of zonal PMV crash-frequency models; 2) to extend the fixed-coefficients approach that is commonly used to model spatially correlated error-terms to estimate the spatially non-stationary relationships within a full Bayesian context; and 3) to identify the factors that contribute to the frequency of PMV crashes in neighborhoods in an urban city. Based on our study, the spatially heterogeneous effects of various factors on the frequency of PMV crashes are expected to be quantified. Such information is of paramount importance for policymakers in the formulation and implementation of multifaceted interventions to reduce environmental hazards and to remove barriers to walking in targeted neighborhoods. The proposed procedure to extract pedestrian trajectories from household travel-diary survey data and the developed BSVC model can also be directly generalized to other regions when modeling the frequency of PMV crashes to obtain more accurate and reliable estimations.

2. Methods

2.1 Data preparation

Our crash data were obtained from the Traffic Road Accident Database System, which is maintained by the Hong Kong Police Force and the Hong Kong Transport Department (Xu et al., 2019a). These data are collected by the police officers at the crash scenes (Meng et al., 2017; Xie et al., 2018 ; Zhou et al., 2020). Only crashes that result in injuries are recorded in the database. With available information on geographical coordinates, the crashes were first mapped onto an ArcGIS map and geo-validated using a procedure developed by Loo (2006). A total of 7,381 PMV crashes were reported by the police on normal weekdays during 2010–2012. Of these, 96.23% were successfully geo-coded. These crashes were then immediately aggregated at the TPU level, which is the smallest unit for planning purpose in Hong Kong (Yao and Loo, 2016). Initially, the Hong Kong Planning Department partitioned the whole territory as 289 TPUs in total in 2011. For the purpose of privacy, 80 TPUs with a sparse population were then merged with their adjacent zones by the Hong Kong Census Department when releasing the 2011 Population Census Report. The resulting 209-TPU system, with an average size of 5.31 km^2 and appromiately one-third of TPUs having an area of less than 1 km^2, was used as the spatial unit system of our analysis, because it readily matches the existing population census data and is capable of quantifying the built environmental factors at a relatively fine geographical scale.

To assign the boundary crashes, a buffer zone with a radius of 100 ft (i.e. 30.48 m) was created first around the TPU boundaries. Crashes located within the boundary buffer were then allocated equally to the adjacent TPUs. This half-to-half ratio assignment method was recommended by Washington et al. (2010) and Xu et al. (2017). Other variables were also spatially attached to their respective TPUs in an analogous manner. Fig. 1 illustrates the spatial distribution of PMV crashes in 209 TPUs on normal weekdays during 2010–2012 in Hong Kong. The number of PMV crashes across TPUs was 34 on average, with a minimum value of 0 to a maximum of 292.

The vehicle flow data were derived from the Annual Traffic Census System, which is maintained by the Hong Kong Transport Department. Each year, the Hong Kong Transport Department publishes its traffic census reports with the vehicle flow data recorded by the counting stations. In 2011, approximately 850 counting stations were surveyed, covering over 85% of trafficable roads in Hong Kong (HKTD, 2012). By multiplying the reported average annual daily traffic volume by the corresponding length of road segments, the average daily vehicle kilometers traveled was obtained.

Fig. 1. The spatial distribution of PMV crashes on normal weekdays in Hong Kong during 2010–2012.

To estimate the city-wide pedestrian exposure, the 2011 Travel Characteristics Survey (HKTD, 2014) released by the Hong Kong Transport Department was used. Between September 2011 and January 2012, a random sample of 35,401 households (approximately 1.5% of domestic households) was successfully enumerated. Respondents were asked to recall all types of activities they had engaged in on the preceding weekday (excluding Saturdays, Sundays, and public holidays). For each trip, detailed information on origin and destination, trip purpose, departure time, arrival time, and trip duration was recorded accordingly. The collected trip records were then extrapolated to the entire population and were further adjusted for underreporting by comparison with independent transportation statistics (To et al., 2005). To estimate pedestrian activities within each TPU, as illustrated in Fig. 2, all of the walking trips including the walk-only trips and walking trip legs were first plotted on the ArcGIS map, with the centroid of street blocks as a trip's origin or destination. In Hong Kong, the street blocks are the smallest enumeration units delineated by the Hong Kong Planning Department. There are 4,993 street blocks across the whole territory, with an average area of 0.22 km^2. For inter-zone trips whose origin and destination were not within the same street block, the shortest network path was calculated using the Dijkstra algorithm (Dijkstra, 1959). These shortest walking paths were then overlaid with the street block map to extract the part of the routes within the boundary of each street block. Given the reported trip duration, together with an assumption that people walked at a constant speed throughout their trips, the walking time of each inter-zone trip within corresponding street blocks could then be calculated. Finally, by spatially joining the street blocks with the TPU map, the walking time within each TPU was obtained. A similar procedure was used to estimate the walking distance at the TPU level, as detailed in Fig. 3. Here we estimated the walking distance for intra-zone trips by multiplying the self-reported walking time by average walking speed stratified by sex and age groups. Age- and sex-specific walking speeds were calculated by dividing the total distance of shortest walking paths estimated for the inter-zone trips by corresponding walking time, as presented in Table A2. Compared with Yao et al. (2015) who used solely the inter-zone trips to extract pedestrian trajectories, our proposed procedure is expected to produce more accurate estimates of pedestrian activities within each neighborhood by an integrated consideration of inter-zone and intra-zone trips. In total, the 2011 Travel Characteristics Survey estimated approximately 28.71 million walking trips (including walk-only trips and walking trip legs), corresponding to 2.40 million hours and 10.21 million kilometers walked by Hong Kong residents per weekday during 2010–2012.

In addition to the estimation of exposure measures, a range of explanatory variables related to land-use, road-network characteristics, and socio-demographic factors that potentially contribute to the frequency of PMV crashes were collected from a crowdsourced dataset in Hong Kong.

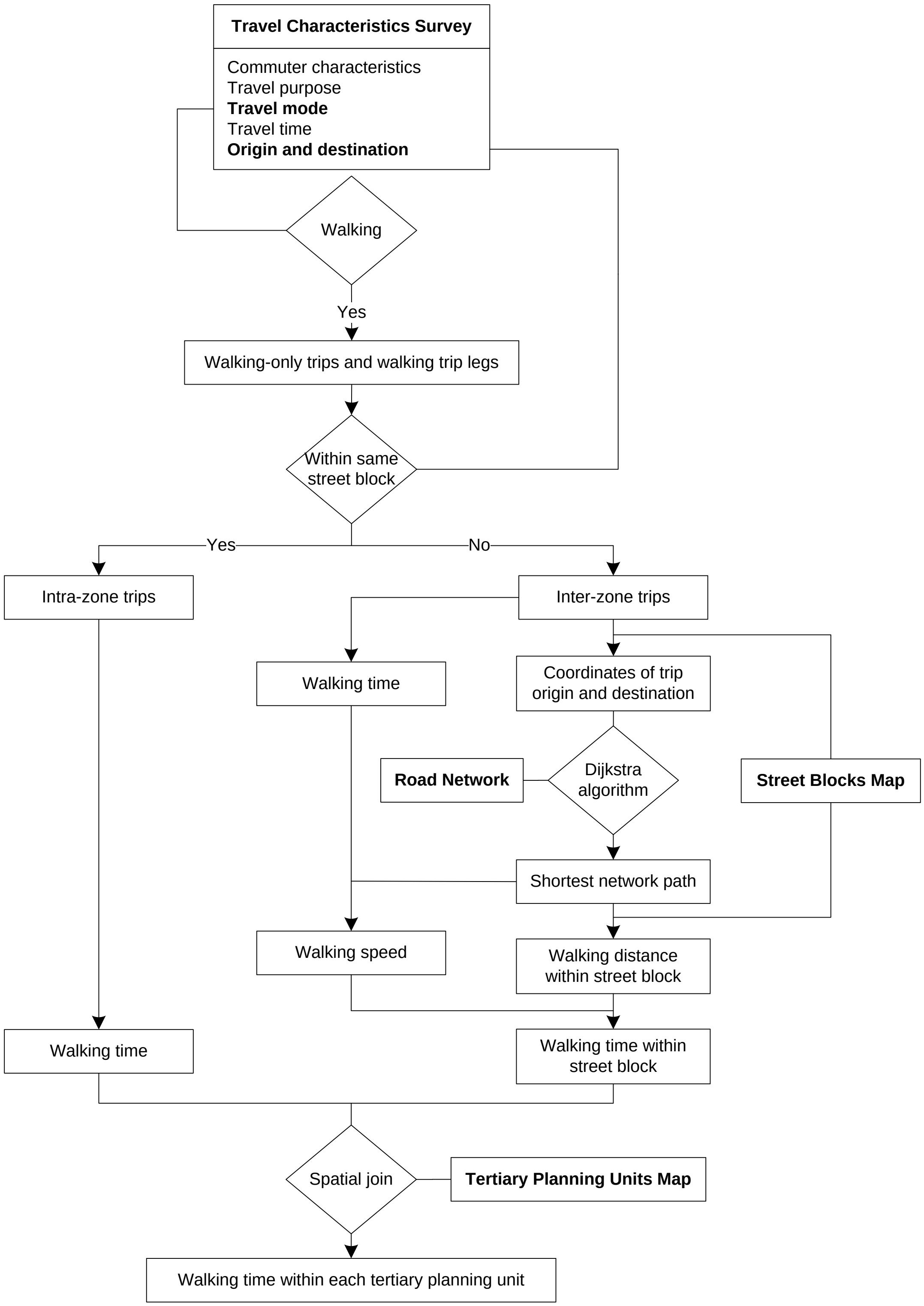

Fig. 2. Flowchart used to estimate walking time at TPU level based on the 2011 Hong Kong Travel Characteristics Survey data.

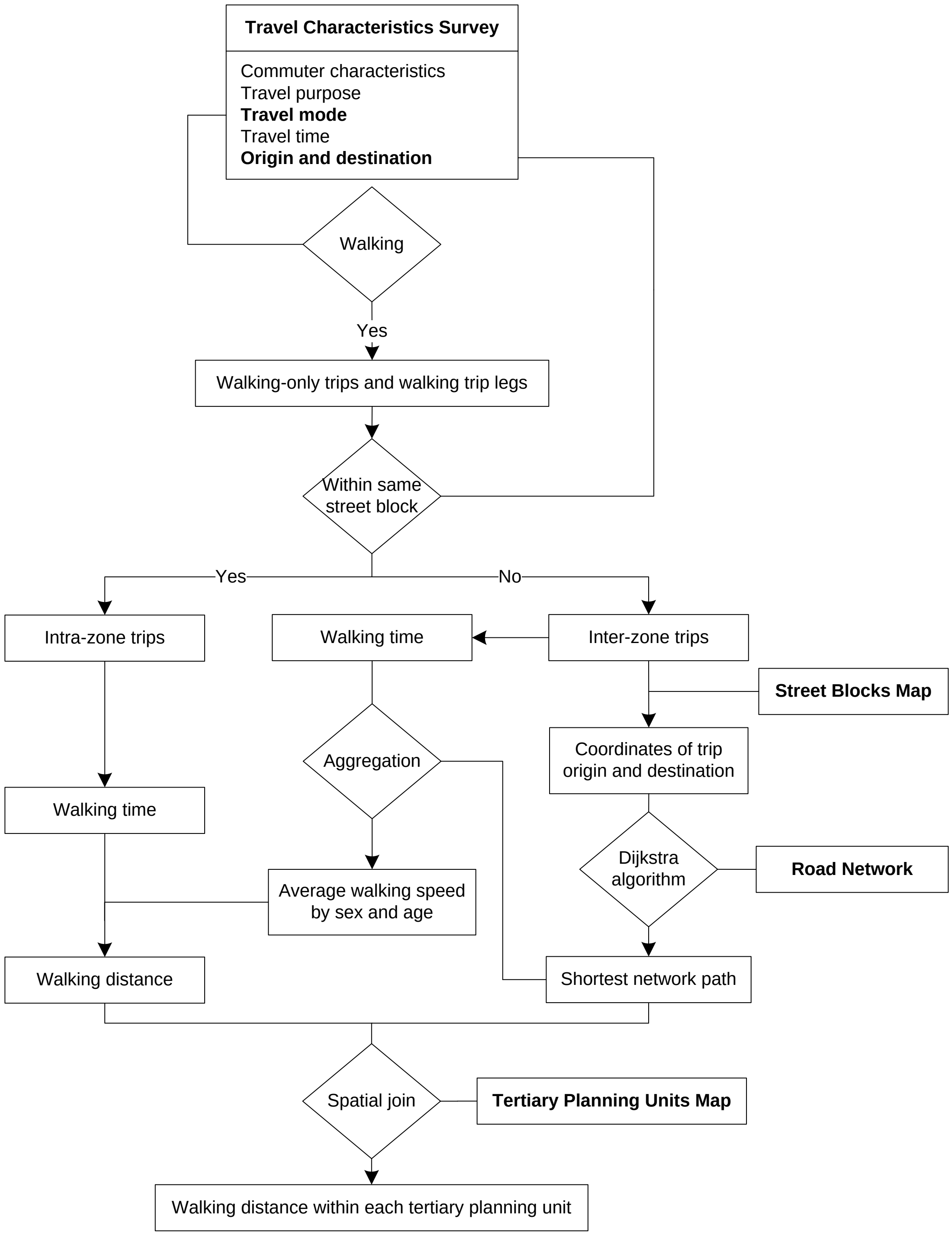

Fig. 3. Flowchart used to estimate walking distance at TPU level based on 2011 Hong Kong Travel Characteristics Survey data.

The digital land-use data were obtained from the Hong Kong Planning Department, which were categorized into seven types: commercial, residential, industrial, institutional, recreational, special utilities, and green space. In addition to using the percentage to indicate the intensity of a particular type of land-use within an area, following Wang and Kockelman (2013), Chen and Zhou (2016), and Ding et al. (2018), we calculated the entropy index to quantify the mixture of land-use as Eq. (1):

$$\text{Entropy}_i = \frac{-\sum_{j=1}^{k_i} p_j^i \ln(p_j^i)}{\ln(k_i)} \quad (1)$$

where p_j^i refers to the percentage of land-use type $j(j=1,2,...,7)$ in TPU $i(i=1,2,...,209)$. k_i denotes the number of land-use types in the ith TPU. The entropy index varies from 0 to 1, with a value towards 1 associated with a greater extent of mixed land-use.

However, the aforementioned entropy index implicitly assumes that an area is perfectly mixed if its land-use types share equal percentage, which seems theoretically inadequate. The balance index (Cervero and Duncan, 2003) was therefore introduced here to measure how different types of land-use interact in balance with each other. Let t_j the percentage of land-use type j within the whole city. Setting the entire area under investigation as a benchmark with well-balanced land-use, the balance index for the ith TPU could then be calculated as (Song et al., 2013):

$$\text{Balance}_i = 1 - \sum_{j}^{7} t_j \left| p_j^i - t_j \right| \quad (2)$$

Similar to the entropy, the balance index also ranges from 0 to 1, with higher values representing more balanced land-use.

The road-network data were extracted from the node-link road centerline system provided by the Hong Kong Lands Department and were further adjusted by the shapefile derived from OpenStreetMap. Various geometric characteristics, namely road density, intersection density, percentages of road-segment lengths with different functional classifications, and percentages of different types of intersections, were included. Road density here was defined as the length of road segments per square kilometers, while intersection density was calculated as the number of intersections divided by the length of road segments.

The points-of-interest data were grabbed from the GeoInfo Map released by the Hong Kong SAR Government. Given precise location information, the numbers of various public facilities, including the bus stops, parking lots, tram stops, metro entrances, petrol stations, supermarkets, shopping malls, convenience stores, licensed hotels, nursing homes for the elderly, child care centers, hospitals, clinics, schools, police stations, country parks, libraries, museums, playgrounds, performing venues, sports centers, and sports grounds, were thus counted within each TPU. This rich points-of-interest dataset provides us a valuable opportunity to examine the effects of some previously under-investigated factors, such as bus-stop density and the prevalence of parking lots, on the frequency of zonal PMV crashes.

Finally, the demographic, educational, economic, and household characteristics were derived from the 2011 Population Census Report. The variables available for model development, along with their descriptive statistics, are presented in Table 1.

Table 1. Characteristics of the 209 TPUs under investigation.

Variables	Mean	SD	Min	Max
Dependent variable				
Number of PMV crasheson working days during 2010–2012	33.99	43.09	0.00	292.00
Exposure variables				
Average daily vehicle kilometers traveled ($\times 10^3$)	127.70	153.79	0.36	1204.05
Resident population ($\times 10^3$)	33.83	41.05	1.02	287.90
Average daily walking trips ($\times 10^3$)	163.78	180.81	0.60	1142.53
Average daily walking time ($\times 10^3$ hours)	11.46	13.84	0.02	89.63
Average daily walking distance ($\times 10^3$ km)	48.84	60.68	0.07	416.11
Explanatory variables				
Land-use				
Percentage of commercial land-use	0.04	0.09	0.00	0.53
Percentage of residential land-use	0.22	0.16	0.00	0.81
Percentage of industrial land-use	0.02	0.07	0.00	0.67
Percentage of institutional land-use	0.10	0.10	0.00	0.67
Percentage of recreational land-use	0.08	0.10	0.00	0.47
Percentage of special utilities	0.17	0.14	0.00	0.56
Percentage of green space	0.37	0.34	0.00	0.98
Land-use mix	0.66	0.23	0.01	0.98
Land-use balance	0.67	0.20	0.41	0.99
Road-network attributes				
Road density (km/km ^{2})	10.52	6.98	0.62	37.49
Percentage of motorways	0.06	0.08	0.00	0.57
Percentage of primary roads	0.05	0.09	0.00	0.69
Percentage of secondary roads	0.11	0.09	0.00	0.36
Percentage of tertiary roads	0.10	0.11	0.00	0.78
Percentage of unclassified roads	0.68	0.17	0.20	1.00
Intersection density (/km)	4.43	1.81	1.00	14.31
Percentage of signalized intersections	0.12	0.12	0.00	0.57
Percentage of roundabouts	0.02	0.02	0.00	0.15
Percentage of three-leg intersections	0.86	0.13	0.24	1.00
Percent of four-leg intersections	0.14	0.12	0.00	0.76
Percent of intersections with five or more legs	0.05	0.01	0.00	0.08
Public facilities				
Bus-stop density (/km)	1.36	0.98	0.00	4.69
Number of parking lots	3.11	4.38	0.00	25.00

Number of tram-stops	0.61	2.21	0.00	14.00
Number of metro entrances	2.45	4.08	0.00	19.00
Number of petrol stations	0.82	1.22	0.00	5.00
Number of supermarkets	3.23	3.44	0.00	15.00
Number of shopping malls	3.27	4.48	0.00	26.00
Number of convenience stores	6.45	7.83	0.00	49.00
Number of licensed hotels	5.93	24.60	0.00	294.00
Number of nursing homes for the elderly	0.78	1.29	0.00	8.00
Number of child care centers	0.06	0.23	0.00	1.00
Number of hospitals	0.21	0.64	0.00	4.00
Number of clinics	1.18	2.04	0.00	11.00
Number of schools	13.11	17.09	0.00	122.00
Number of police stations	0.22	0.46	0.00	3.00
Number of country parks	0.11	0.40	00.0	3.00
Number of libraries	0.89	1.25	0.00	9.00
Number of museums	0.09	0.45	0.00	5.00
Number of playgrounds	0.30	0.63	0.00	3.00
Number of performing venues	0.08	0.31	0.00	2.00
Number of sports centers	0.49	0.80	0.00	4.00
Number of sports grounds	0.13	0.34	0.00	1.00
Demographic characteristics				
Proportion of male population	0.47	0.04	0.36	0.85
Proportion of population aged less than 15	0.12	0.03	0.00	0.21
Proportion of population aged between 15 and 24	0.11	0.03	0.02	0.27
Proportion of population aged between 25 and 44	0.33	0.05	0.21	0.57
Proportion of population aged between 45 and 64	0.30	0.04	0.15	0.44
Proportion of population aged 65 or above	0.13	0.06	0.00	0.44
Proportion of population of Chinese ethnicity	0.88	0.12	0.43	0.99
Educational characteristics (highest level attended)				
Proportion of population with primary education or below	0.28	0.08	0.07	0.61
Proportion of population with secondary education	0.45	0.07	0.17	0.84
Proportion of population with post -secondary education	0.27	0.13	0.00	0.75
Economic characteristics				
Labor-force participation rate	0.60	0.08	0.00	0.88
Proportion of working population	0.51	0.72	0.00	0.83
Proportion of working population with place of work at home	0.12	0.10	0.00	0.51
Median monthly income ($\times 10^3$)	13.86	5.68	0.00	40.00
Household characteristics				
Household density ($\times 10^3$/km ^{2})	10.66	12.74	0.00	58.09
Average household size	2.93	0.38	1.60	4.10
Proportion of households with three or more persons	0.56	0.12	0.00	0.82
Median monthly household income ($\times 10^3$)	33.43	29.73	0.00	170.80
Median monthly household rent ($\times 10^3$)	8.06	11.71	0.00	76.00

Median rent to income ratio	0.19	0.11	0.00	0.53
Proportion of population in public rental housing	0.16	0.24	0.00	1.00
Proportion of population in subsidized home ownership housing	0.08	0.14	0.00	0.78
Proportion of population in permanent housing	0.72	0.32	0.00	1.00
Proportion of population in non-domestic housing	0.01	0.06	0.00	0.85
Proportion of population in temporary housing	0.03	0.08	0.00	0.53

2.2 Model specification

We modeled the frequency of PMV crashes consistent with previous studies (Sebert Kuhlmann et al., 2009; Wang and Kockelman, 2013; DiMaggio, 2015; Lee et al., 2015; Guo et al., 2017; Osama and Sayed, 2017; Goel et al., 2018). Let Y_i denote the number of PMV crashes in the *i*th TPU on working days during 2010–2012. The use of aggregate crash data over a 3-year period helps to avoid confounding effects and the regression-to-the-mean phenomenon (Cheng and Washington, 2005). V_i and P_i refer to the vehicle and pedestrian volumes, respectively, and X_{ik} is the *k*th explanatory variable related to zone-specific attributes. Given the potential non-linear relationship between PMV crashes and traffic volumes (Elvik and Goel, 2019), we have:

$$Y_i \sim \text{Poisson}(\lambda_i)$$
$$\ln(\lambda_i) = \beta_1 + \beta_2 \ln(V_i) + \beta_3 \ln(P_i) + \sum_{k=4}^{p} \beta_k X_{ik} + u_i + s_i \quad (3)$$

where λ_i is the parameter of the Poisson model (i.e., the expected number of PMV crashes in the *i*th TPU; β_1 is the intercept; $\beta_k (k = 2, ..., p)$ refers to the *k*th regression coefficients to be estimated; u_i denotes the unstructured effect, which is specified as an exchangable normal prior with a mean of 0 and a variance of σ_u^2, i.e., $u_i \sim \text{Normal}(0, \sigma_u^2)$; and s_i is the spatially structured or spatially correlated effect.

One commonly used joint density for $s(s_1, s_2, ..., s_{209})$ is formulated in terms of pairwise differences in errors and a variance term of σ_s^2 (Besag et al., 1991):

$$P(s_1, s_2, ..., s_n) \propto \exp[-0.5(\sigma_s^2)^{-1} \sum_{i\sim j} c_{ij}(s_i - s_j)^2] \quad (4)$$

This results in a normal conditional prior for s_i:

$$s_i \mid s_{j\neq i} \sim \text{Normal}\left(\frac{\sum_j c_{ij} s_j}{\sum_j c_{ij}}, \frac{\sigma_s^2}{\sum_j c_{ij}}\right) \quad (5)$$

where c_{ij} represents the non-normalized weight, e.g., $c_{ij} = 1$ if TPU i is adjacent to TPU j, otherwise $c_{ij} = 0$. In our study, geographically non-contiguous zones were also considered as neighbors if they were directly connected by cross-harhor tunnels, bridges, or ferries. σ_s^2 is the variance parameter, controlling the amount of extra variations due to spatial correlation.

Although the unvirate conditional prior distribution in Eq. (5) is well defined, the corresponding joint prior distribution for $\mathbf{S}$ is improper with undefined mean and infinite variance (Sun et al., 1999). This fact probably leads to problems in convergence and identifiability (Eberly and Carlin, 2000).

An alternative strategy to gain properity is based on the strength of a single set of random effects $\mathbf{v}(v_1, v_2, \ldots, v_{209})$:

$$\ln(l_i) = b_1 + b_2 \ln(V_i) + b_3 \ln(P_i) + \sum_{k=4}^{p} b_k X_{ik} + v_i \tag{6}$$

Following Lee (2011), v_i here is specified as the CAR prior proposed by Leroux et al. (1999):

$$v_i \left| v_{j \neq i} \right. \sim \text{Normal}\left(\frac{r_v \sum_j c_{ij} v_j}{1 - r_v + r_v \sum_j c_{ij}}, \frac{s_v^2}{1 - r_v + r_v \sum_j c_{ij}}\right) \tag{7}$$

where $r_v (0 \le r_v \le 1)$ is the spatial correlation parameter, with $r_v = 0$ simplifying to an independently and identically distributed normal prior, and a value closer to 1 indicating a stronger spatial correlation. Accordingly, setting $r_v = 1$ corresponds to the intrinsic CAR, as in Eq. (5).

Based on the factorization theorem, $\mathbf{V}$ results in a joint multivariate Gaussian distribution:

$$\mathbf{v} \sim \text{MVN}(\mathbf{0}, \sigma_v^2 [\rho_v \mathbf{K} + (1 - \rho_v)\mathbf{I}]^{-1}) \tag{8}$$

where $\mathbf{I}$ is an $n \times n$ matrix, and the elements of $\mathbf{K}$ are calculated as:

$$K_{ij} = \begin{cases} \sum_j c_{ij} & \text{if } i = j \\ -c_{ij} & \text{if } i \neq j \end{cases} \tag{9}$$

Although the covariance structure in Eq. (6) incorporates the local relationships, the outputs from the preceding models still consist of a set of global parameter estimates. Intuitively, the local variations can be addressed by setting the regression coefficients as random effects, allowing the effects of covariates to vary spatially:

$$\ln(\lambda_i) = \beta_1 + \beta_{i2} \ln(V_i) + \beta_{i3} \ln(P_i) + \sum_{k=4}^{p} \beta_{ik} X_{ik} + v_i \tag{10}$$

where β_{ik} is the coefficient of the kth explanatory variable for TPU i.

To account for both the unstructured and spatially structured variations in model regression coefficients, following Xu et al. (2017), we have:

$$\boldsymbol{\beta}_\mathbf{k} \sim \text{MVN}(\boldsymbol{\mu}_\mathbf{k}, \sigma_k^2 [\rho_k \mathbf{K} + (1 - \rho_k)\mathbf{I}]^{-1}) \tag{11}$$

Unlike Eq. (8), Eq. (11) has a constant non-zero mean $\boldsymbol{\mu}_\mathbf{k}(\mu_k, \ldots, \mu_k)$, in which μ_k is the overall estimate of the regression slope, representing the average of the posterior estimates of $\boldsymbol{\beta}_\mathbf{k}(\beta_{1k}, \beta_{2k}, \ldots, \beta_{209k})$. The precision matrix is now given by $\rho_k \mathbf{K} + (1 - \rho_k)\mathbf{I}$, which is a weighted average of spatially correlated and independent structures denoted as $\mathbf{K}$ and $\mathbf{I}$, respectively. This specification is capable of accounting for a range of weak and strong spatial correlation in regression coefficients, with $\rho_k = 0$ reducing to spatially independent random effects only, while an increase in ρ_k toward 1 represnts more spatial smoothing.

Accordingly, the univariate full conditional distribution for Eq. (11) is:

$$\beta_{ik} \left| \beta_{jk} \sim \text{Normal}\left(\frac{\rho_k \sum_j c_{ij}\beta_{jk} + (1-\rho_k)\mu_k}{1-\rho_k+\rho_k\sum_j c_{ij}}, \frac{\sigma_k^2}{1-\rho_k+\rho_k\sum_j c_{ij}}\right)\right. \quad (12)$$

Specifically, the conditonal expectation of β_{ik} is a weighted average of the random effects at neighboring zones and the overall mean μ_k. When β_{ik} exhibits a strong spatial correlation, ρ_k is close to 1 and the conditional variance approaches $\sigma_k^2 / \sum_j c_{ij}$. This variance configuration recognizes that in the presence of strong spatial correlation, the more neighbors a neighborhood has, the more information the data contain on the value of its random effects. In comparison, if the random effect is spatially independent, the conditional variance becomes σ_k^2. Evidently, the parameter $\rho_k (0 \le \rho_k \le 1)$ serves as an indicator to assess the relative strength of spatial and unstructured variations in the estimated coefficients. In addition, if there is no significant heterogeneity in $\boldsymbol{\beta_k}$, σ_k^2 becomes dispersive with the mean of its posterior distribution lower than the standard deviation (Barua et al., 2015; Xu et al., 2017). In this case, the regression slopes are better modeled as fixed effects.

Obtaining the full Bayesian posterior estimates requires the specification of prior distributions. Prior distributions are typically used to reflect prior knowledge about the parameters of interest. If such information is available, it is encouraged to formulate the so-called informative priors (Yu and Abdel-Aty, 2013; Heydari et al., 2014). In the absence of sufficient prior knowledge, non-informative priors, were applied to model the parameters here (Dong et al., 2016):

$$\begin{aligned} \beta_k &\sim \text{Normal}(0, 1000) \\ \mu_k &\sim \text{Normal}(0, 1000) \end{aligned} \quad (13)$$

Consistent with Congdon (2008), the spatial correlation parameters ρ_v and ρ_k were assigned as $\text{uniform}(0,1)$. A $\text{uniform}(0,10)$ was also specified for σ_v and σ_k, respectively, following Gelman (2006), Lee (2011), and Xu et al. (2017).

2.3 Model-performance comparison measures

For model comparison, three commonly used measures were adopted here, i.e., the mean absolute deviance (MAD), mean squared prediction error (MSPE), and deviance information criterion (DIC).

The MAD was calculated as follows (Xu et al., 2015; Yao et al., 2015):

$$\text{MAD} = \frac{1}{209}\sum_{i=1}^{209}\left|\hat{Y}_i - Y_i\right| \quad (14)$$

where $\hat{Y}_i$ denotes the predicted number of PMV crashes estimated by the fitted models. A smaller value of MAD suggests that on average the model predicts the observed data better.

The MSPE was also used to provide a measure of model predictive performance (Yao et al., 2015), which was formulated as:

$$\mathrm{MSPE} = \frac{1}{209}\sum_{i=1}^{209}(\hat{Y}_i - Y_i)^2 \qquad (15)$$

Similar to the MAD, models with a lower value of MSPE indicate a better predictive performance.

Meanwhile, as a penalized goodness-of-fit measure, the DIC was used here to take model complexity into account:

$$\mathrm{DIC} = D(\bar{\theta}) + 2p_D = \bar{D} + p_D \qquad (16)$$

where $D(\bar{\theta})$ is the deviance evaluated at $\bar{\theta}$, the posterior means of the parameters; p_D is the effective number of parameters in the model; and $\bar{D}$ is the posterior mean of the deviance statistic $D(\theta)$. The lower the DIC value, the better the model fit. In general, for a pair of models with a difference in DIC value of more than 10, the model with a higher DIC value is definitely ruled out; model pairs with DIC differences between 5 and 10 are considered substantially different; and a difference of less than 5 indicates that the two models are not statistically different (Spiegelhalter et al., 2002).

3. Results and discussion

The freeware WinBUGS (Spiegelhalter et al., 2005) was used to calibrate the models. Three parallel chains with diverse starting points were tracked. The first 50,000 iterations were discarded as burn-ins, and then 5,000 iterations were performed for each chain, resulting in a sample distribution of 15,000 for each parameter. The model's convergence was monitored by the Brooks-Gelman-Rubin statistic (Brooks and Gelman, 1998), visual examination of the Markov chain Monte Carlo chains, and the ratios of Monte Carlo errors relative to the respective standard deviations of the estimates. As a rule of thumb, these ratios should be less than 0.05.

For model specification, a correlation test was conducted first to ensure the non-inclusion of strongly correlated variables. Our correlation analysis indicated a strong correlation between percentage of land-use categorized as special utilities and road density, between labor force participation rate and proportion of working population, and between average household size and proportion of population in permanent housing, with the estimated Spearman's correlation parameters (Washington et al., 2011) greater than 0.70. Likewise, the proportion of working population with place of work at home, median monthly income, median monthly household income, and median monthly household rent were also highly correlated, suggesting that these four variables should not be simultaneously added to the models. Similar conclusions hold true for the variables of the number of supermarkets, the number of shopping malls, and the number of convenience stores, given their Spearman's correlation parameters all greater than 0.80. Other variables showed weak collinearity, as their Spearman's correlation parameters were less than 0.50. In the initial model, we included all of the uncorrelated variables (Xie et al., 2018; Zhou et al., 2020). The DIC was then used to compare alternative models with different covariate subsets. The model producing a lower DIC value was considered statistically superior.

For comparison purposes, in addition to the BSVC model, we developed the Bayesian spatially fixed coefficients (BSFC) model with a spatially correlated error term. To highlight the role of pedestrian exposure, models with population, walking trips, walking time, and walking distance as the exposure measure, respectively, were estimated and compared. As such, eight models were calibrated. The performance of these models is presented below, followed by the presentation and interpretation of the parameter estimates.

3.1 Model-performance comparison

Table 2 shows the results of goodness-of-fit measures for the calibrated models. The values of MAD, MSPE, and DIC in the BSVC models were not substantially different from those derived from the BSFC models, indicating that our data were fairly robust to model configuration. More specifically, the BSVC model with walking trips as the measure of pedestrian exposure performed best, given the lowest value of DIC. Based on a similar dataset but aggregating pedestrian crashes to 26 districts in Hong Kong, Sze et al. (2019) also reported that the model using walking frequency as the proxy for pedestrian exposure was superior to the other two counterparts using zonal population and walking time. Although population information is readily available as it is routinely reported by local authorities, the use of such an aggregated data as a surrogate for pedestrian exposure completely neglects the variations in pedestrian activities within an area of interest. This negligence may produce biased results for central business districts with a sparse resident population but a prevalance of pedestrian activities during workday rush-hours.

Table 2. Goodness-of-fit measures for BSFC and BSVC models for the frequency of PMV crashes in 209 TPUs in Hong Kong during 2010-2012.

Model type	Description	Pedestrian exposure	MAD	MSPE	DIC
BSFC-1	Bayesian spatially fixed coefficients model	Population	5.53	68.05	1318.36
BSFC-2		Walking trips	5.52	67.91	1314.36
BSFC-3		Walking time	5.54	68.01	1323.28
BSFC-4		Walking distance	5.54	68.10	1322.31
BSVC-1	Bayesian spatially varying coefficients model	Population	5.54	68.10	1320.27
BSVC-2		Walking trips	5.52	67.86	1313.66
BSVC-3		Walking time	5.54	68.23	1324.51
BSVC-4		Walking distance	5.54	68.20	1322.81

MAD, MSPE, and DIC: mean absolute deviance, mean squared prediction error, and deviance information criterion, respectively.

Among the three activity-based exposure measures, the model of walking trips seems to better account for the cross-sectional variability in zonal counts of PMV crashes. One plausible explanation is that due to the self-reported nature of travel-diary data, information on waking trips is more reliable than that on walking time, given the potential recall bias and inconsistency of time perception among individual respondents. Likewise, the estimated walking distance in our study depends on a strong assumption that pedestrians chose the shortest path. In reality, the route choice of pedestrians, however, is considerably complex and dynamic, because people do not always choose the shortest path when walking from one place to another (Guo and Loo, 2013). As a consequence, the measurement-errors introduced in the process of estimating walking time and walking distance probably lead to the reduced model performance.

3.2 Parameter estimates

Tables 3 and 4 summarize the parameter estimates in the BSFC and BSVC models applied to the frequency of PMV crashes in 209 TPUs in Hong Kong, respectively. A 5% level of significance was used as the threshold to determine whether the parameters differed significantly from 0. Variables insignificant in all eight models were then excluded.

Several general observations are worthy of mention. First, the significant variables were not entirely identical between the models of population and activity-based exposure measures. For example, the percentage of residential land-use and the percentage of roundabouts were statistically significant in the models of population, but became totally insignificant in the models of activity-based exposure measures. The same holds true for the variable of median monthly income, as this variable was only significant in the models of activity-based exposure measures. Second, relative to the models of resident population, the effects of several risk factors, i.e., road density, the percentage of motorways, and the number of parking lots, changed substantially in the models of activity-based exposure measures. Specifically, the coefficient of road density in the BSVC model decreased sharply from 0.52 to 0.37 once the number of walking trips was used as the exposure measure. Similar results were also observed for the effects of the percentage of motorways and the number of parking lots. These findings raise an alarm that the extensive use of population or population density to represent pedestrian exposure in previous studies (LaScala et al., 2000; Graham and Glaister, 2003; Noland and Quddus, 2004; Priyantha Wedagama et al., 2006; Loukaitou-Sideris et al., 2007; Wier et al., 2009; Chakravarthy et al., 2010; Cottrill and Thakuriah, 2010; Ha and Thill, 2011; Ukkusuri et al., 2011, 2012; Siddiqui et al., 2012; Dumbaugh and Zhang, 2013; Graham et al., 2013; Noland et al., 2013; DiMaggio, 2015; Lee et al., 2015; Wang et al., 2016; Gai et al., 2017a; Goel et al., 2018; Rothman et al., 2020) has very likely resulted in biased estimates and incorrect inferences.

Table 3. Results of the BSFC model with a spatially correlated error term for the frequency of PMV crashes in 209 TPUs in Hong Kong during 2010–2012.

	BSFC-1			BSFC-2			BSFC-3			BSFC-4		
Variables	Mean	SD	95% BCI	Mean	SD	95% BCI	Mean	SD	95% BCI	Mean	SD	95% BCI
Intercept	–3.62*	0.49	(–4.61, –2.65)	–4.41*	0.61	(–5.82, –3.32)	–2.45*	0.39	(–3.23, –1.73)	–3.06*	0.48	(–4.07, –2.13)
Ln (vehicle km traveled)	0.39*	0.05	(0.29, 0.50)	0.34*	0.06	(0.21, 0.45)	0.36*	0.06	(0.25, 0.47)	0.37*	0.06	(0.26, 0.48)
Ln (population)	0.48*	0.05	(0.38, 0.59)									
Ln (walking trips)				0.51*	0.06	(0.40, 0.64)						
Ln (walking time)							0.43*	0.05	(0.34, 0.52)			
Ln (walking distance)										0.43*	0.05	(0.33, 0.53)
Residential land-use (%)	–0.25*	0.05	(–0.36, –0.15)	–0.06	0.05	(–0.15, 0.04)	–0.06	0.05	(–0.16, 0.03)	–0.06	0.05	(–0.16, 0.03)
Road density	0.51*	0.05	(0.41, 0.61)	0.31*	0.05	(0.21, 0.42)	0.34*	0.05	(0.23, 0.44)	0.33*	0.06	(0.23, 0.44)
Motorways (%)	–0.18*	0.05	(–0.27, –0.08)	–0.14*	0.05	(–0.24, –0.04)	–0.16*	0.05	(–0.26, –0.06)	–0.18*	0.05	(–0.28, –0.08)
Intersection density	0.23*	0.06	(0.12, 0.34)	0.15*	0.06	(0.04, 0.26)	0.15*	0.06	(0.03, 0.27)	0.14*	0.06	(0.03, 0.26)
Roundabouts (%)	–0.12*	0.05	(–0.22, –0.03)	–0.08	0.05	(–0.17, 0.01)	–0.09	0.05	(–0.18, 0.01)	–0.08	0.05	(–0.18, 0.01)
Bus-stop density	0.21*	0.05	(0.11, 0.31)	0.16*	0.05	(0.05, 0.27)	0.18*	0.05	(0.08, 0.29)	0.18*	0.06	(0.07, 0.29)
Number of parking lots	0.19*	0.05	(0.10, 0.28)	0.12*	0.05	(0.02, 0.22)	0.13*	0.05	(0.02, 0.22)	0.12*	0.05	(0.01, 0.22)
Median monthly income	0.01	0.05	(–0.11, 0.09)	–0.13*	0.05	(–0.23, –0.04)	–0.11*	0.05	(–0.21, –0.02)	–0.11*	0.05	(–0.21, –0.01)
σ_v^2	0.31*	0.09	(0.20, 0.55)	0.30*	0.09	(0.19, 0.53)	0.33*	0.10	(0.20, 0.58)	0.35*	0.11	(0.21, 0.62)
ρ_v	0.08	0.07	(0.00, 0.26)	0.07	0.06	(0.00, 0.25)	0.08	0.07	(0.00, 0.27)	0.09	0.08	(0.00, 0.30)

BSFC: Bayesian spatially fixed coefficients model.

SD: standard deviation.

BCI: Bayesian credible interval.

* denotes significance at 95% CI.

Table 4. Results of the BSVC model for the frequency of PMV crashes in 209 TPUs in Hong Kong during 2010–2012.

	BSFC-1			BSFC-2†			BSFC-3			BSFC-4		
Variables	Mean	SD	95% BCI	Mean	SD	95% BCI	Mean	SD	95% BCI	Mean	SD	95% BCI
Intercept	–3.89*	0.51	(–4.84, –2.92)	–4.37*	0.55	(–5.48, –3.34)	–2.46*	0.41	(–3.28, –1.66)	–2.97*	0.45	(–3.86, –2.09)
Ln (vehicle km traveled)	0.42*	0.06	(0.31, 0.54)	0.33*	0.05	(0.24, 0.44)	0.36*	0.05	(0.26, 0.47)	0.37*	0.05	(0.27, 0.49)
Ln (population)	0.50*	0.05	(0.40, 0.60)									
Ln (walking trips)				0.50*	0.05	(0.40, 0.61)						
Ln (walking time)							0.43*	0.05	(0.34, 0.53)			
Ln (walking distance)										0.42*	0.05	(0.32, 0.51)
Residential land-use (%)	–0.24*	0.05	(–0.35, –0.13)	–0.08	0.05	(–0.18, 0.02)	–0.09	0.05	(–0.19, 0.01)	–0.09	0.05	(–0.19, 0.01)
Road density	0.52*	0.05	(0.42, 0.63)	0.37*	0.08	(0.22, 0.54)	0.39*	0.08	(0.24, 0.56)	0.40*	0.08	(0.23, 0.57)
$\sigma^2_{\text{Road density}}$	——	——	——————	*0.15**	*0.12*	*(0.01, 0.44)*	*0.16**	*0.12*	*(0.01, 0.48)*	*0.17**	*0.13*	*(0.01, 0.49)*
$\rho_{\text{Road density}}$	——	——	——————	*0.47**	*0.29*	*(0.02, 0.98)*	*0.45**	*0.29*	*(0.01, 0.98)*	*0.46**	*0.29*	*(0.02, 0.97)*
Motorways (%)	–0.26*	0.08	(–0.42, –0.10)	–0.16*	0.05	(–0.26, –0.06)	–0.18*	0.05	(–0.28, –0.08)	–0.19*	0.05	(–0.29, –0.10)
$\sigma^2_{\text{motorways}}$	*0.21**	*0.16*	*(0.01, 0.63)*	——	——	——————	——	——	——————	——	——	——————
$\rho_{\text{motorways}}$	*0.37**	*0.28*	*(0.01, 0.95)*	——	——	——————	——	——	——————	——	——	——————
Intersection density	0.21*	0.06	(0.10, 0.32)	0.14*	0.06	(0.04, 0.26)	0.15*	0.06	(0.03, 0.27)	0.14*	0.06	(0.03, 0.26)
Roundabouts (%)	–0.16*	0.05	(–0.27, –0.06)	–0.08	0.05	(–0.17, 0.01)	–0.09	0.05	(–0.18, 0.01)	–0.09	0.05	(–0.18, 0.01)
Bus-stop density	0.17*	0.05	(0.07, 0.28)	0.17*	0.05	(0.06, 0.27)	0.19*	0.05	(0.08, 0.29)	0.19*	0.05	(0.08, 0.29)
Number of parking lots	0.20*	0.05	(0.11, 0.29)	0.13*	0.05	(0.04, 0.22)	0.13*	0.05	(0.03, 0.22)	0.13*	0.05	(0.03, 0.22)
Median monthly income	0.01	0.05	(–0.09, 0.11)	–0.12*	0.05	(–0.21, –0.03)	–0.10*	0.05	(–0.19, –0.01)	–0.10*	0.05	(–0.19, –0.01)
σ^2_v	0.27*	0.11	(0.12, 0.54)	0.25*	0.09	(0.13, 0.48)	0.29*	0.11	(0.15, 0.59)	0.30*	0.12	(0.15, 0.60)
ρ_v	0.15	0.16	(0.00, 0.62)	0.10	0.10	(0.00, 0.38)	0.13	0.13	(0.00, 0.49)	0.14	0.14	(0.00, 0.52)

BSVC: Bayesian spatially fixed coefficients model.

SD: standard deviation.

BCI: Bayesian credible interval.

* denotes significance at 95% CI.

† Detailed WinBUGS code for the BSVC-2 model was presented in the Appendix.

More importantly, unlike the BSFC models whose coefficients were restricted to be constant, the BSVC models allowed the regression coefficients to vary spatially. Hence, using BSFC approach, a single model was applied for the entire area, whereas different coefficients could be estimated for each neighborhood by virtue of the BSVC model. It is also interesting to observe that the error variability (σ_v^2) decreased slightly when variations were introduced to the regression coefficients. This result is intuitively reasonable, because the heterogeneity in the regression slopes can capture some of the extra variations previously explained by the random effects in the error term (Xu et al., 2017).

Given that the BSVC model with walking trips as the measure of pedestrian exposure performed best, we chose it to interpret our results in the subsequent section. As Table 4 shows, eight variables had a significant association with the frequency of PMV crashes: vehicle kilometers traveled, walking trips, road density, percentage of motorways, junction density, bus-stop density, number of parking lots, and median monthly income. The signs of these parameters were generally consistent with empirical judgments and the results of previous studies (Chen and Zhou, 2016; Guo et al., 2017; Tasic et al., 2017).

Both vehicle kilometers traveled and walking trips were significant and positive, with coefficients estimated at 0.33 and 0.50, respectively. This nonlinear relationship between pedestrian volume and the number of PMV crashes has been widely reported (Geyer et al., 2006; Schneider et al., 2010; Miranda-Moreno et al., 2011; Elvik et al., 2013; Elvik, 2016; Mooney et al., 2016; Guo et al., 2017; Osama and Sayed, 2017; Tasic et al., 2017; Xie et al., 2018; Sze et al., 2019; Stipancic et al., 2020), suggesting that as the number of pedestrians increases, the absolute number of PMV crashes also increases, whereas the risk of collisions by motor vehicles for each individual pedestrian decreases. This is known as the "safety-in-numbers" effect (Jacobsen, 2003; Elvik and Bjørnskau, 2017; Elvik and Goel, 2019; Xu et al., 2019b; Cai et al., 2020). One plausible explanation is that motorists adjust their behavior when they encounter a group of people walking. This hypothesis is evidenced by the greater visibility of pedestrians in greater numbers (Jacobsen et al., 2015). As motorists is less likely to collide with pedestrians when more people are walking, policies that encourage walking are claimed as effective measures to improve the safety of pedestrians (Jacobsen, 2003; Osama and Sayed, 2017). However, this conclusion based on a cross-sectional research design should be interpreted with great caution, because it is impossible to determine whether this safety-in-numbers effect is a causal relationship or merely a statistical association (Bhatia and Wier, 2011; Xu et al., 2019b). Bhatia and Wier (2011) also cautioned that treating the promotion of walking as a safety intervention would mask efforts to create an inherently safe environment for all pedestrians. Further efforts are therefore warranted to investigate the underlying mechanisms.

Instead of simply encouraging people to walk in groups to draw motorists' notice, an alternatively sound measure to improve pedestrian safety would be to restrict the usage of motor vehicles. In addition to the benefits of less congestion, fewer emissions of pollutants, and less traffic noise, the strategies to reduce vehicle volume would lower both the number of PMV crashes and the crash risk for pedestrians. According to our results, all things being equal, a neighrborhood halving its vehicle kilometers traveled would expect an approximately 20% decrease in PMV crashes ($1-0.50^{0.33}= 0.20$). Therefore, a modal shift from motor vehicles to other travel modes, such as public transit and walking, should be vigorously advocated, especially in dense urban settings.

Interestingly, road density had a spatially varying coefficient with a posterior mean of 0.37 and a variance of 0.15. The magnitude of this coefficient ranged from –0.01 to 0.60, as shown in Fig. 4. Given these distributional parameters, almost all of the TPUs (i.e., 208 out of 209) exhibited a positive association between road density and the frequency of PMV crashes. This heterogeneous effect is likely to reflect the variations in road conditions across neighborhoods and may result partially from some unobserved factors, such as topology of road networks, speed limits, and the presence of pedestrian facilities. In addition, the spatial correlation parameter $\rho_{\text{Road density}}$ produced a posterior estimate with a mean of 0.47 and a standard deviation of 0.29, implying that a moderate proportion of heterogeneity was explained by the spatially correlated effects. The corresponding 95% CI was (0.02, 0.98), which significantly differed from both 0 and 1. This finding demonstrated the presence of both unstructured and spatially structured variations in the effects of related risk factors on the frequency of zonal PMV crashes.

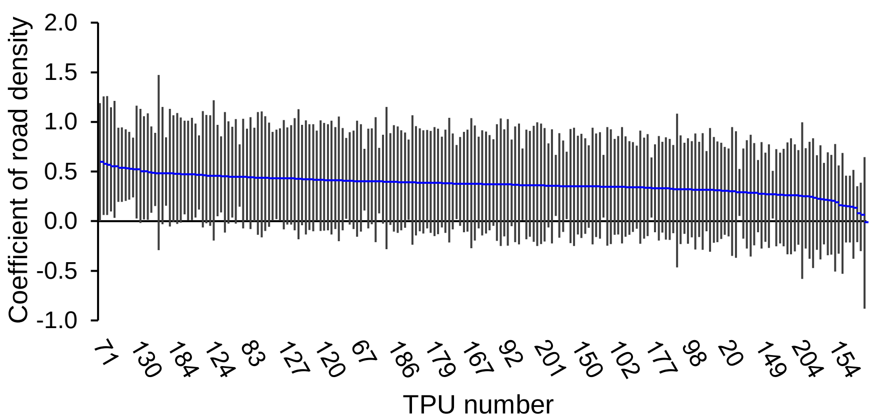

Fig. 4. Mean and 95% Bayesian credible interval for the variable of road density, estimated by the BSVC model with walking trips as the measure of pedestrian exposure (ranked in descending order by mean values; dots: mean; solid lines: 95% Bayesian credible interval).

One major advantage of the BSVC model is its capability to explain spatially non-stationary relationships. A mapping of local parameters would therefore help local authorities to explicitly identify neighborhoods where a particular risk factor has a greater influence on the frequency of PMV crashes. To illustrate, Fig. 5 identifies the overall pattern of the regression coefficients of road density as spatial clustering. Special attention should be paid to rural areas located in the New Territories, as road density in these neighborhoods was found to be highly significant, resulting in a more pronounced effect.

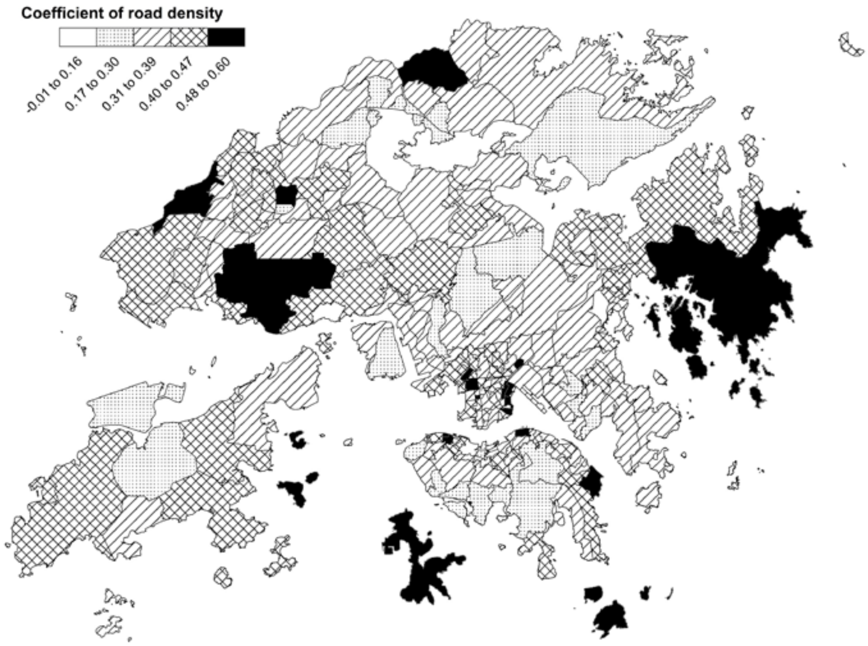

Fig. 5. Spatial distribution of the coefficients for the variable of road density, estimated by the BSVC model with walking trips as the measure of pedestrian exposure.

Road functional classification also had a significant influence on the frequency of PMV crashes. According to our results, neighborhoods with a higher percentage of motorways were associated with a lower risk of PMV crashes. This result is highly expected, because as limited-access roads, motorways mainly serve fast-moving vehicles while pedestrians are prohibited (Tasic et al., 2017).

Intersections are well-known as hazardous locations in road networks. Given equal road length, more intersections imply less continuous road networks and more conflict points between pedestrians and motor vehicles. It is therefore not surprising that intersection density had a significantly positive relationship with the frequency of PMV crashes. Similar findings were also reported by Priyantha Wedagama et al. (2006), Guo et al. (2017), Osama and Sayed (2017), and Tasic et al. (2017).

With respect to the variables related to public facilities, bus-stop density was associated with an increased risk of PMV crashes. A visual examination via Google Street View suggests that traffic conditions near bus stops in Hong Kong are fairly complicated and mixed, as various road users share the same activity spaces. A previous study by Lam et al. (2014) reported that pedestrians were more likely to jaywalk to board buses without noticing the approaching vehicles. The stationary buses at bus stops may also obscure the visibility of pedestrians when crossing the road (Chen and Zhou, 2016). As a consequence, the frequent interactions between pedestrians and motor vehicles near bus stops inevitably increase the risk of collisions (Retting et al., 2003; Zegeer and Bushell, 2012; Stoker et al., 2015; Goel et al., 2018). Similar explanations hold true for the effect of the number of parking lots.

Finally, area deprivation level is closely associated with safety awareness, driving behavior, and road facilities, and thus has an indirect influence on the frequency of PMV crashes. Consistent with Siddiqui et al. (2012), Noland et al. (2013), Jermprapai and Srinivasan (2014), Cai et al. (2017a), and Rothman et al. (2020), the negative relationship between median monthly income and the frequency of PMV crashes found in our study is very likely attributable to two causes. First, people with higher incomes may have better awareness of safe walking. Second, deprived neighborhoods probably have fewer favorable pedestrian facilities such as marked crosswalks, overpasses or underpasses, and refuge islands to completely separate pedestrians from motor vehicles on roads (Rothman et al., 2020).

4. Conclusions

This study sought to investigate the factors that contribute to the frequency of PMV crashes at zonal levels, using a rich dataset collected in 209 TPUs in Hong Kong over a 3-year period. Detailed activity-based exposure data were integrated with land-use, road-network features, accessibility of public facilities, and socio-demographic characteristics to construct our dataset. A procedure was proposed to extract pedestrian trajectories from publicly available travel-diary survey data. A BSVC model was then developed to account for the spatially heterogeneous effects of risk factors. To highlight the role of exposure, models with population, walking trips, walking time, and walking distance as the measure of pedestrian exposure, respectively, were calibrated and compared.

Several key findings are worthy of note. First, although pedestrian volume is indispensable in determining the incidence of pedestrian crashes, the major challenge lies in the unavailability of reliable pedestrian activity data for the whole area under investigation. Fortunately, the household travel-diary survey provides a straightforward and invaluable means to estimate territory-wide pedestrian activities. Given the rich trip information recorded in the survey, by virtue of an integrated use of crowdsourced datasets, pedestrian trajectories can be easily estimated based on our proposed procedure. By incorporating these activity-based exposure measures into PMV crash-frequency models, we explicitly demonstrate that the use of population or population density as a surrogate for pedestrian exposure when modeling the frequency of zonal PMV crashes will lead to biased estimations and incorrect inferences, as our empirical results indeed indicated substantial inconsistencies in the the effects of several risk factors between the models of population and activity-based exposure measures

Second, among the three activity-based exposure measures, walking trips were empirically proved adequate in accounting for the spatial variations in zonal counts of PMV crashes. Although the model using walking trips as the measure of pedestrain exposure resulted in a slightly better goodness-of-fit, the models of walking distance and walking time could also help in the evaluation of safety effects of specific transport policies, such as quantifying the safety benefits associated with a modal shift from motor vehicles to walking for short-distance trips in metropolitan areas. Indeed, only when we understand the differences in how much people walk, will we be able to measure how safe walking is, and further evaluate the effectiveness of specific countermeasures in improving pedestrian safety.

Third, by virtue of the BSVC model, we add new insights to existing studies that in addition to the unstructured variability, the heterogeneity in the effects of explanatory variables on the frequency of PMV crashes can also arise from spatially correlated effects. The developed BSVC model provides a sound methodological alternative to investigate the spatially non-stationary relationships, as the varying regression coefficients are modeled via a single set of random effects and a spatial correlation parameter, with extreme values corresponding to pure unstructured or pure spatially correlated random effects. Given that pedestrian crash data are typically collected in spatial proximity, we expect our study to promote the awareness among traffic professionals that spatial heterogeneity should not be neglected when modeling pedestrian crashes involving contiguous spatial units.

Eight variables were ultimately found to have a significant association with the frequency of PMV crashes in neighborhoods. The nonlinear relationship between pedestrian volume and the frequency of PMV crashes was confirmed, with an estimated coefficient of 0.50, 0.43, and 0.42 for walking trips, walking time, and walking distance, respectively. Vehicle kilometers traveled, road density, intersection density, bus-stop density, and the number of parking lots were found to be positively associated with PMV crash frequency, whereas the percentage of motorways and median monthly income had negative effects on the risk of PMV crashes. These findings are expected to assist local authorities in the formulation of effective strategies to improve walkability and pedestrian safety in neighborhoods. Such countermeasures may include restricting the use of motor vehicles, promotion of a shift from motor vehicles to walking for walkable-distance trips, traffic calming in residential neighborhoods, updating of pedestrian facilities to completely separate pedestrians from motor vehicles on very busy roads, reducing conflicts between pedestrians and motor vehicles particularly near bus stops and at entrances of parking lots, and publicity on safe road-crossing behaviors.

Our study is not without limitations. The TPUs used in our analysis are delineated mainly for planning purposes and may not be the optimal spatial units for zonal pedestrian crash analysis. Given the potential of the modifiable areal unit problem (Xu et al., 2014; Lee et al., 2014; Cai et al., 2017b; Xu et al., 2018; Obelheiro et al., 2020), more efforts are warranted to validate the robustness of our findings via aggregation of data at various spatial resolutions. In addition, although the neighborhood-level analysis is useful to investigate the effects of area-wide variables on the frequency of PMV crashes, pedestrian safety is actually a microscopic concern, because PMV crashes are usually caused by interactions between pedestrian and motor vehicles (Yue et al., 2020). Future studies towards an integration of cross-sectional research designs with in-depth crash-causation investigations and accident-reconstruction simulations are highly recommended to achieve deeper insights into the causes of pedestrian crashes.

Acknowledgments: We would like to thank the Hong Kong Police Force and the Hong Kong Transport Department for providing access to the database used in this study. The views expressed are the authors' own and do not necessarily represent the views of the Hong Kong Police Force and the Hong Kong Transport Department.

Funding: This work was supported by the grants from the Natural Science Foundation of China (Project No. 71601163), National Key R&D Program of China (2016YFC0802208), and Sichuan Science and Technology Program (Project No. 2020YFH0035; 2020YJ0268; 2020YJ025; 2020JDRC0032). Pengpeng Xu was supported by the Y S and Christabel Lung Postgraduate Scholarship and a Research Postgraduate Studentship from the University of Hong Kong. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: We declare that no competing interests exist.

References

Ariannezhad, A., Karimpour, A., Wu, Y., 2020. Incorporating mode choice into safety analysis at the macroscopic level. *Journal of Transportation Engineering, Part A: Systems* 146(4), 04020022.

Bao, J., Liu, P., Yu, H., Xu, C., 2017. Incorporating twitter-based human activity information in spatial analysis of crashes in urban areas. *Accident Analysis and Prevention* 106, 358–369.

Barua, S., El-Basyouny, K., Islam, M.T., 2015. Effects of spatial correlation in random parameter collision count-data models. *Analytic Methods in Accident Research* 5–6, 28–42.

Besag, J., York, J., Molli, E.A., 1991. Bayesian image restoration with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* 43(1), 1–59.

Bhatia, R., Wier, M., 2011. "Safety in Numbers" re-examined: Can we make valid or practical inferences from available inference? *Accident Analysis and Prevention* 43(1), 235–240.

Brooks, S.P., Gelman, A., 1998. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 7(4), 434–455.

Cai, Q., Abdel-Aty, M., Lee, J., 2017a. Macro-level vulnerable road users crash analysis: A Bayesian joint modeling approach of frequency and proportion. *Accident Analysis and Prevention* 107, 11–19.

Cai, Q., Abdel-Aty, M., Lee, J., Eluru, N., 2017b. Comparative analysis of zonal systems for macro-level crash modeling. *Journal of Safety Research* 61, 157–166.

Cai, Q., Abdel-Aty, M., Castro, S., 2020. Explore effects of bicycle facilities and exposure on bicycle safety at intersections. International Journal of Sustainable Transportation DOI: 10.1080/15568318.2020.1772415

Cai, Q., Lee, J., Eluru, N., Abdel-Aty, M., 2016. Macro-level pedestrian and bicycle crash analysis: Incorporating spatial spillover effects in dual state count models. *Accident Analysis and Prevention* 93, 14–22.
Cervero, R., Duncan, M., 2003. Walking, bicycling, and urban landscapes: Evidence from the San Francisco Bay Area. *American Journal of Public Health* 93(9), 1478–1483.
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.
Chen, P., Zhou, J., 2016. Effects of the built environment on automobile-involved pedestrian crash frequency and risk. *Journal of Transport and Health* 3(4), 448–456.
Cheng, W., Washington, S.P., 2005. Experimental evaluation of hotspot identification methods. *Accident Analysis and Prevention* 37(5), 870–881.
Congdon, P., 1997. Bayesian models for spatial incidence: A case study of suicide using the BUGS program. *Health and Place* 3(4), 229–247.
Congdon, P., 2008. A spatially adaptive conditional autoregressive prior for area health data. *Statistical Methodology* 5(6), 552–563.
Congdon, P., 2014. *Applied Bayesian Modelling (2nd edition)*. John Wiley & Sons, Chichester, UK.
Cottrill, C.D., Thakuriah, P., 2010. Evaluating pedestrian crashes in areas with high low-income or minority populations. *Accident Analysis and Prevention* 42(6), 1718–1728.
Cressie, N., 1993. *Statistics for Spatial Data*. John Wiley & Sons, New York, US.
Delmelle, E.C., Thill, J.C., Ha, H.H., 2012. Spatial epidemiology analysis of relative collision risk factors among urban bicyclists and pedestrians. *Transportation* 39(2), 433–448.
Dijkstra, E., 1959. A note on two problems in connection with graphs. *Numerische Mathematik* 1(1), 269–271.
DiMaggio, C., 2015. Small-area spatiotemporal analysis of pedestrian and bicyclist injuries in New York City. *Epidemiology* 26(2), 247–254.
Ding, C., Chen, P., Jiao, J., 2018. Non-linear effects of the built environment on automobile-involved pedestrian crash frequency: A machine learning approach. *Accident Analysis and Prevention* 111, 116–126.
Dong, N., Huang, H., Lee, J., Gao, M., Abdel-Aty, M., 2016. Macroscopic hotspots identification: A Bayesian spatio-temporal interaction approach. *Accident Analysis and Prevention* 92, 256–264.
Dumbaugh, E., Zhang, Y., 2013. The relationship between community design and crashes involving older drivers and pedestrians. *Journal of Planning Education Research* 33(1), 83–95.
Eberly, L., Carlin, B. , 2000. Identifiability and convergence issues for Markov chain Monto Carlo fitting of spatial models. *Statistics in Medicine* 19(17–18), 2279–2294.

Elvik, R., Sørensen, M.W.J., Nævestad, T.O., 2013. Factors influencing safety in a sample of marked pedestrian crossing selected for safety inspections in the city of Oslo. *Accident Analysis and Prevention* 59, 64–70.

Elivk, R., 2016. Safety-in-numbers: Estimates based on a sample of pedestrian crossings in Norway. *Accident Analysis and Prevention* 91, 175–182.

Elvik, R., Bjørnskau, T., 2017. Safety-in-numbers: A systematic review and meta-analysis of evidence. *Safety Science* 92, 274–282.

Elvik, R., Goel, R., 2019. Safety-in-numbers: An updated meta-analysis of estimates. *Accident Analysis and Prevention* 129, 136–147.

Fotheringham, A.S., Brunsdon, C., Charlton, M.E., 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationship*. Wiley, Chichester, UK.

Gelman, A., 2006. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis* 1, 515–533.

Geyer, J., Raford, N., Pham, T., Ragland, D.R., 2006. Safety in numbers: Data from Oakland, California. *Transportation Research Record: Journal of the Transportation Research Board* 1982, 150–154.

Goel, R., Jain, P., Tiwari, G., 2018. Correlates of fatality risk of vulnerable road users in Delhi. *Accident Analysis and Prevention* 111, 86–93.

Gomes, M.M., Pirdavani, A., Brijs, T., Pitombo, C.S., 2019. Assessing the impacts of enriched information on crashes prediction performance. *Accident Analysis and Prevention* 122, 162–171.

Graham, D.J., Glaister, S., 2003. Spatial variation in road pedestrian casualties: The role of urban scale, density and land-use mix. *Urban Studies* 40(8), 1591–1607.

Graham, D.J., McCoy, E.J., Stephen, 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.

Guo, Q., Xu, P., Pei, X., Wong, S.C., Yao, D., 2017. The effect of road network patterns on pedestrian safety: A zone-based Bayesian spatial modeling approach. *Accident Analysis and Prevention* 99, 114–124.

Guo, Z., Loo, B.P.Y., 2013. Pedestrian environment and route choice: Evidence from New York City and Hong Kong. *Journal of Transport Geography* 28, 124–136.

Ha, H.H., Thill, J.C., 2011. Analysis of traffic hazard intensity: A spatial epidemiology case study of urban pedestrians. *Computers, Environment, and Urban System* 35(3), 230–240.

Hadayeghi, A., Shalaby, A.S., Persaud, B.N., 2010. Development of planning level transportation safety tools using geographically weighted Poisson regression. *Accident Analysis and Prevention* 42(2), 676–688.

Heydari, S., Miranda-Moreno, L.F., Lord, M., Fu, L., 2014. Bayesian methodology to estimate and update safety performance functions under limited data conditions: A sensitivity analysis. *Accident Analysis and Prevention* 64, 41–51.
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 and Prevention* 131, 15–24.
Hong Kong Transport Department, 2012. *The Annual Traffic Census 2011*. https://www.td.gov.hk/en/publications_and_press_releases/publications/free_publications/the_annual_traffic_census_2011/index.html.
Hong Kong Transport Department, 2014. *Travel Characteristics Survey 2011 Final Report*. https://www.td.gov.hk/filemanager/en/content_4652/tcs2011_eng.pdf.
Huang, Y., Wang, X., Patton, D., 2018. Examining spatial relationships between crashes and the built environment: A geographically weighted regression approach. *Journal of Transport Geography* 69, 221–233.
Jacobsen, P.L., 2003. Safety in numbers: More walkers and bicyclists, safer walking and bicycling. *Injury Prevention* 9(3), 205–209.
Jacobsen, P.L., Ragland, D.R., Komanoff, C., 2015. Safety in numbers for walkers and bicyclists: Exploring the mechanisms. *Injury Prevention* 21(4), 217–220.
Jermprapai, K., Srinivasan, S., 2014. Planning-level model for assessing pedestrian safety. *Transportation Research Record: Journal of the Transportation Research Board* 2464, 109–117.
Lam, W.W.Y., Yao, S., Loo, B.P.Y., 2014. Pedestrian exposure measures: A timespace framework. *Travel Behaviour and Society* 1(1), 22–30.
LaScala, E.A., Gerber, D., Gruenewald, P.J., 2000. Demographic and environmental correlated of pedestrian injury collisions: A spatial analysis. *Accident Analysis and Prevention* 32(5), 651–658.
Lee, D., 2011. A comparison of conditional autoregressive models used in Bayesian disease mapping. *Spatial and Spatio-Temporal Epidemiology* 2(2), 79–89.
Lee, J., Abdel-Aty, M., Jiang, X., 2014. Development of zone system for macro-level traffic safety analysis. *Journal of Transport Geography* 38, 13–21.
Lee, J., Abdel-Aty, M., Choi, K., Huang, H., 2015. Multi-level hot zone identification for pedestrian safety. *Accident Analysis and Prevention* 76, 64–73.
Lee, J., Abdel-Aty, M., Huang, H., Cai, Q., 2019. Transportation safety planning approach for pedestrians: An integrated framework of modeling walking duration and pedestrian fatalities. *Transportation Research Record: Journal of the Transportation Research Board* 2673(4), 898-906.
Leroux, B., Lei, X., Breslow, N., 1999. Estimation of disease rates in small areas: A new mixed model for spatial dependence. In Halloran, M., Berry, D., (eds), *Statistical Models in Epidemiology, the Environment and Clinical Trials*. New York: Springer-Verlag, 135–178.

Li, Z., Wang, W., Liu, P., Bigham, J.M., Ragland, D.R., 2013. Using geographically weighted Poisson regression for county-level crash modeling in California. *Safety Science* 58, 89–97.

Loo, B.P.Y., 2006. Validating crash locations for quantitative spatial analysis: A GIS-based approach. *Accident Analysis and Prevention* 38(5), 879–886.

Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice* 44(5), 291–305.

Loukaitou-Sideris, A., Liggett, R., Sung, H., 2007. Death on the crosswalk: A study of pedestrian-automobile collisions in Log Angeles. *Journal of Planning Education and Research* 26(3), 338–351.

Maibach, E., Steg, L., Anable, J., 2009. Promoting physical activity and reducing climate change: Opportunities to replace short car trips with active transportation. *Preventive Medicine* 49(4), 325–327.

Mannering, F.L., Bhat, C.R., 2014. Analytic methods in accident research: Methodological frontier and future directions. *Analytic Methods in Accident Research* 1, 1–22.

Meng, F., Xu, P., Wong, S.C., Huang, H., Li, Y.C., 2017. Occupant-level injury severity analyses for taxis in Hong Kong: A Bayesian space-time logistic model. *Accident Analysis and Prevention* 108, 297–307.

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 and Prevention* 43(5), 1624–1634.

Mooney, S.J., DiMaggio, C.J., Lovasi, G.S., Neckerman, K.M., Bader, M.D., Teitler, J.O., Sheehan, D.M., Jack, D.W., Rundle, A.G., 2016. Use of Google Street View to assess environmental contributions to pedestrian injury. *American Journal of Public Health* 106(3), 462–469.

Noland, R.B., Quddus, M.A., 2004. Analysis of pedestrian and bicycle casualties with regional panel data. *Transportation Research Record: Journal of the Transportation Research Board* 1897, 28–33.

Noland, R.B., Klein, N.J., Tulach, N.K., 2013. Do lower income areas have more pedestrian casualties? *Accident Analysis and Prevention* 59, 337–345.

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.

Osama, A., Sayed, T., 2017. Evaluating the impact of connectivity, continuity, and topology of sidewalk network on pedestrian safety. *Accident Analysis and Prevention* 107, 117–125.

Pirdavani, A., Bellemans, T., Brijs, T., Kochan, B., Wets, G., 2014. Assessing the road safety impacts of a teleworking policy by means of geographically weighted regression method. *Journal of Transport Geography* 39, 96–110.

Priyantha Wedagama, D.M., Bird, R.N., Metcalfe, A.V., 2006. The influence of urban land-use on non-motorised transport casualties. *Accident Analysis and Prevention* 38(6), 1049–1057.
Retting, R.A., Ferguson, S.A., McCartt, A.T., 2003. A review of evidence-based traffic engineering measures designed to reduce pedestrian–motor vehicle crashes. *American Journal of Public Health* 93(9), 1456–1463.
Richardson, S., Guihenneuc, C., Lasserre, V., 1992. Spatial linear models with autocorrelated error structure. *The Statistician* 41(5), 539–557.
Rothman, L., Cloutier, M., 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, 229–233.
Schneider, R.J., Diogenes, M.C., Arnold, L.S., Attaset, V., Griswold, J., Ragland, D.R., 2010. Association between roadway intersection characteristics and pedestrian crash risk in Alameda County, California. *Transportation Research Record: Journal of the Transportation Research Board* 2198, 41–51.
Sebert Kuhlmann, A.K., Brett, J., Thomas, D., Sain, S.R., 2009. Environmental characteristics associated with pedestrian–motor vehicle collisions in Denver, Colorado. *American Journal of Public Health* 99(9), 1632–1637.
Shariat-Mohaymany, A., Shahri, M., Mirbagheri, B., Matkan, A.A., 2015. Exploring spatial non-stationarity and varying relationships between crash data and related factors using geographically weighted Poisson regression. *Transactions in GIS* 19(2), 321–337.
Siddiqui, C., Abdel-Aty, M., Choi, K., 2012. Macroscopic spatial analysis of pedestrian and bicycle crashes. *Accident Analysis and Prevention* 45, 382–391.
Song, Y., Merlin, L., Rodriguez, D., 2013. Comparing measures of urban land use mix. *Computers, Environment, and Urban System* 42, 1–13.
Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van Der Linde, A., 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 64(4), 583–639.
Spiegelhalter, D.J., Thomas, A., Best, N., Lunn, D., 2005. *WinBUGS User Manual*. MRC Biostatistics Unit, Cambridge, UK.
Steinbach, R., Edwards, P., Green, J., 2014. Controlling for exposure changes the relationship between ethnicity, deprivation and injury: An observational study of child pedestrian injury rates in London. *Injury Prevention* 20(3), 159–166.
Stipancic, J., Miranda-Moreno, L., Strauss, J., Labbe, A., 2020. Pedestrian safety at signalized intersections: Modelling spatial effects of exposure, geometry and signalization in a large urban network. *Accident Analysis and Prevention* 134, 105265.

Stoker, P., Garfinkel-Castro, A., Khayesi, M., Odero, W., Mwangi, M.N., Peden, M., Ewing, R., 2015. Pedestrian safety and the built environment: A review of the risk factors. *Journal of Planning Literature* 30(4), 377–392.

Sun, D., Tsutakawa, R., Speckman, P.L., 1999. Bayesian inference for CAR(1) models with noninformative priors. *Biometrika* 86(2), 341–350.

Sze, N.N., Su, J., Bai, L., 2019. Exposure to pedestrian crash based on household survey data: Effect of trip purpose. *Accident Analysis and Prevention* 128, 17–24.

Tasic, I., Elvik, R., Brewer, S., 2017. Exploring the safety in numbers effect for vulnerable road users on a macroscopic scale. *Accident Analysis and Prevention* 109, 36–46.

To, D.K.B., Yau, K.T., Lam, A., 2005. Travel characteristics survey – method of expanding household interview survey data. *Transportmetrica* 1(3), 247–260.

Ukkusuri, S., Hasan, S., Aziz, H.M.A., 2011. Random parameter model used to explain effects of built-environment characteristics on pedestrian crash frequency. *Transportation Research Record: Journal of the Transportation Research Board* 2237, 98–106.

Ukkusuri, S., Miranda-Moreno, L.F., Ramadurai, G., Isa-Tavarez, J., 2012. The role of built environment on pedestrian crash frequency. *Safety Science* 50(4), 1141–1151.

Wang, X., Yang, J., Lee, C., Ji, Z., You, S., 2016. Macro-level safety analysis of pedestrian crashes in Shanghai, China. *Accident Analysis and Prevention* 96, 12–21.

Wang, Y., Kockelman, K.M., 2013. A Poisson-lognormal conditional-autoregressive model for multivariate spatial analysis of pedestrian crash counts across neighborhoods. *Accident Analysis and Prevention* 60, 71–84.

Washington, S.P., Karlaftis, M.G., Mannering, F.L., 2011. *Statistical and Economic Methods for Transportation Data Analysis (2nd Edition)*. CRC Press, New York, US.

Washington, S.P., Van Schalkwyk, I., You, D., Shin, K., Samuelson, J.P., 2010. *Forecasting the Safety Impacts of Socio-demographic Changes and Safety Countermeasures*. Transportation Research Board, Washington, DC., US.

Wen, H., Zhang, X., Zeng, Q., Sze, N., 2019. Bayesian spatial-temporal model for the main and interaction effects of roadway and weather characteristics on freeway crash incidence. *Accident Analysis and Prevention* 132, 105249.

Wheeler, D.C., Calder, C.A., 2007. An assessment of coefficient accuracy in linear regression models with spatially varying coefficients. *Journal of Geographical Systems* 9(2), 145–166.

Wier, M., Weintraub, J., Humphreys, E.H., Seto, E, Bhatia, R., 2009. An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. *Accident Analysis and Prevention* 41(1), 137–145.

Xie, K., Ozbay, K., Kurkcu, A., Yang, H., 2017. Analysis of traffic crashes involving pedestrians using big data: Investigation of contributing factors and identification of hotspots. *Risk Analysis* 37(8), 1459–1467.

Xie, S.Q., Dong, N., Wong, S.C, Xu, P., 2018. Bayesian approach to model pedestrian crashes at signalized intersections with meas urement errors in exposure. *Accident Analysis and Prevention* 121, 295–294.

Xu, P., Dong, N., Wong, S.C., Huang, H., 2019a. Cyclists injured in traffic crashes in Hong Kong: A call for action. *PLOS ONE* 14(8), e0220785.

Xu, P., Huang, H., 2015. Modeling crash spatial heterogeneity: Random parameter versus geographically weighting. *Accident Analysis and Prevention* 75, 16–25.

Xu, P., Huang, H., Dong, N., 2018. The modifiable areal unit problem in traffic safety: Basic issue, potential solutions and future research. *Journal of Traffic and Transportation Engineering (English Edition)* 5(1), 73–82.

Xu, P., Huang, H., Dong, N., Abdel-Aty, M., 2014. Sensitivity analysis in the context of regional safety modeling: Identifying and assessing the modifiable areal unit problem. *Accident Analysis and Prevention* 70, 110–120.

Xu, P., Huang, H., Dong, N., Wong, S.C., 2017. Revisiting crash spatial heterogeneity: A Bayesian spatially varying coefficients approach. *Accident Analysis and Prevention* 98, 330–337.

Xu, P., Xie, S., Dong, N., Wong, S.C., Huang, H. 2019b. Rethinking safety in numbers: Are intersections with more crossing pedestrians really safer? *Injury Prevention* 25, 20–25.

Yang, L., Chau, K.W., Szeto, W.Y., Cui, X., Wang, X., 2020. Accessibility to transit, by transit, and property prices: Spatially varying relationships. *Transportation Research Part D: Transport and Environment* 85, 102387.

Yao, S., Loo, B.P.Y., 2016. Safety in numbers for cyclists beyond national-level and city-level data: A study on the non-linearity of risk within the city of Hong Kong. *Injury Prevention* 22(6), 379–395.

Yao, S., Loo, B.P.Y., Lam, W.W.Y., 2015. Measures of activity-based pedestrian exposure to the risk of vehicle-pedestrian collisions: Spacetime path vs. potential path tree methods. *Accident Analysis and Prevention* 75, 320–332.

Yu, R., Abdel-Aty, M., 2013. Investigating different approaches to develop informative priors in hierarchical Bayesian safety performance functions. *Accident Analysis and Prevention* 75, 16–25.

Yue, L., Abdel-Aty, M., Wu, Y., Zheng, O., Yuan, J., 2020. In-depth approach for identifying crash causation patterns and its implications for pedestrian crash prevention. *Journal of Safety Research* 73, 119–132.

Zegeer, C.V., Bushell, M., 2012. Pedestrian crash trends and potential countermeasures from around the world. *Accident Analysis and Prevention* 44(1), 3–11.

Zeng, Q., Wen, H., Huang, H., Abdel-Aty, M., 2017. A Bayesian spatial random parameters Tobit model for analyzing crash rates on roadway segments. *Accident Analysis and Prevention* 100, 37–43.

Zeng, Q., Guo, Q., Wong, S.C., Wen, H., Huang, H., Pei, X., 2019. Jointly modeling area-level crash rates by severity: a Bayesian multivariate random-parameters spatio-temporal Tobit regression. *Transportmetrica A: Transport Science* 15(2), 1867–1884.

Zeng, Q., Wen, H., Wong, S.C., Huang, H., Guo, Q., Pei, X., 2020. Spatial joint analysis for zonal daytime and nighttime crash frequencies using a Bayesian bivariate conditional autoregressive model. *Journal of Transportation Safety & Security* 12(4), 566–585.

Zhao, R., Zhan, L., Yao, M., Yang, L., 2020. A geographically weighted regression model augmented by Geodetector analysis and principal component analysis for the spatial distribution of PM2.5. *Sustainable Cities and Society* 56, 102106.

Zhou, H., Yuan, C., Dong, N., Wong, S.C., Xu, P., 2020. Severity of passenger injuries on public buses: A comparative analysis of collision injuries and non-collision injuries. *Journal of Safety Research* DOI: 10.1016/j.jsr.2020.04.003

Ziakopoulos, A., Yannis, G., 2020. A review of spatial approaches in road safety. *Accident Analysis and Prevention* 135, 105323.

Appendix

Table A1. Studies of factors that influence the frequency of pedestrian crashes at zonal levels over the past two decades.

Authors	Study region	Study period	Observations	Research method	Exposure measures		Risk factors included				
					Motor vehicles	Pedestrians	Land-use	Road network	POI	Population census	Climate
LaScala et al. (2000)	San Francisco, US	1990	149 census tracts	Spatial lag model and spatial error model	Average daily traffic	Population	×	✓	✓	✓	×
Graham and Glaister (2003)	UK	1999–2000	8,413 wards	Negative-binomial model	×	Population density	×	✓	×	✓	✓
Noland and Quddus (2004)	UK	1979–1998	11 standard statistical regions	Negative-binomial model	Total number of vehicles registered	Population	×	✓	×	✓	×
Priyantha Wedagama et al. (2006)	Newcastle upon Tyne, UK	1998–2001	185 enumeration districts	Negative-binomial model	Road length	Population density	✓	✓	×	✓	×
Loukaitou-Sideris et al. (2007)	Los Angeles, US	1994–2001	860 census tracts	Multiple linear model	Average annual daily traffic	Population density and employment density	✓	×	×	✓	×
Sebert Kuhlmann et al. (2009)	Denver, US	2000–2003	134 census tracts	Bayesian spatial model with CAR prior	Population density	Commuting by walking	✓	×	×	✓	×
Wier et al. (2009)	San Francisco, US	2001–2005	176 census tracts	Multiple linear model	Average daily traffic	Population	✓	✓	×	✓	×
Chakravarthy et al. (2010)	Orange County, California, US	2000–2004	577 census tracts	Negative-binomial model	×	Population	×	×	×	✓	×
Cottrill and Thakuriah (2010)	Chicago, US	2005	1,832 census tracts	Poisson-lognormal model with exogenous underreporting	Annual average daily traffic	Population density		✓		✓	×

Ha and Thill (2011)	Buffalo, US	2003–2004	90 census tracts	Spatial lag model and spatial error model	×	Population	×	✓	✓	✓	×
Ukkusuri et al. (2011)	New York City, US	2002–2006	2216 census tracts	Random-parameters negative-binomial model	×	Population	✓	✓	✓	✓	×
Delmelle et al. (2012)	Buffalo, US	2003–2004	90 census tracts	Spatial error model	×	Commuting by walking	×	✓	✓	✓	×
Siddiqui et al. (2012)	Florida (District Seven), US	2005–2006	1,479 traffic analysis zones	Bayesian spatial model with CAR prior	×	Population	×	✓	×	✓	×
Ukkusuri et al. (2012)	New York City, US	2002–2006	180 ZIP codes 2216 census tracts	Generalized negative-binomial model	×	Population	✓	✓	✓	✓	×
Dumbaugh and Zhang (2013)	San Antonio, US	2003–2007	938 census block groups	Negative-binomial model	Vehicle miles traveled	Population	✓	✓	×	✓	×
Graham et al. (2013)	UK	2001–2007	1,820 wards	Bayesian spatial longitudinal generalized linear mixed model	×	Population	×	✓	×	✓	×
Noland et al. (2013)	New Jersey, US	2003–2007	6,460 census block groups	Bayesian spatial model with CAR prior	×	Population	×	✓	×	✓	×
Wang and Kockelman (2013)	Austin, US	2007–2009	218 census tracts	Bayesian multivariate CAR model	Vehicle miles traveled	Walking miles	✓	✓	×	✓	×
Jermprapai and Srinivasan (2014)	Florida, US	2005–2009	11,390 census block groups	Negative-binomial model	Work trips per week	×	✓	✓	✓	✓	×

DiMaggio (2015)	New York City, US	2001–2010	1,908 census tracts	Bayesian hierarchical model with space-time interaction effects	Vehicle kilometers traveled per day per square kilometer	Population	×	×	×	✓	×
Lee et al. (2015)	Florida, US	2009–2011	983 ZIP codes	Bayesian simultaneous equations model with CAR prior	Vehicle miles traveled	Population	×	✓	✓	✓	×
Cai et al. (2016)	Florida, US	2010–2012	8,518 traffic analysis zones	Bayesian dual-state model with spatial spillover effects	Vehicle miles traveled	Commuting by walking	×	✓	×	✓	×
Chen and Zhou (2016)	Seattle, US	2008–2012	863 traffic analysis zones	Bayesian spatial model with CAR prior	Total number of trips	Walking trips	✓	✓	✓	✓	×
Wang et al. (2016)	Shanghai, China	2009	263 traffic analysis zones	Bayesian spatial model with CAR prior	×	Population	✓	✓	×	✓	×
Cai et al. (2017a)	Florida, US	2010–2012	594 traffic analysis districts	Bayesian joint model of frequency and proportion	Vehicle miles traveled	Population density	×	✓	×	✓	×
Guo et al. (2017)	Hong Kong, China	2011	131 traffic analysis zones	Bayesian spatial model with CAR prior	Vehicle hours traveled	Walk-only trips	✓	✓	×	✓	×
Osama and Sayed (2017)	Vancouver, Canada	2009–2013	134 traffic analysis zones	Bayesian spatial model with CAR prior	Vehicle kilometers traveled	Walking trips	×	✓	×	×	×
Tasic et al. (2017)	Chicago, US	2005–2012	801 census tracts	Generalized additive model	Vehicle miles traveled	Walking trips	✓	✓	✓	✓	×
Xie et al. (2017)	Manhattan, New York City, US	2008–2012	6,204 grid cells	Tobit model	Vehicle miles traveled	×	✓	✓	×	✓	×

Ding et al. (2018)	Seattle, US	2008–2012	863 traffic analysis zones	Multiple additive Poisson regression tree model	Total number of trips	Walking mode share	✓	✓	✓	×	×
Goel et al. (2018)	Delhi, India	2011–2012	282 wards	Bayesian spatial model with CAR prior	Vehicle kilometers traveled	Population	×	✓	×	✓	×
Lee et al. (2019)	US	2014–2016	47 metropolitan statistical areas	Bayesian integrated model	×	Walking hours	×	×	×	✓	✓
Sze et al. (2019)	Hong Kong, China	2011–2015	26 broad districts	Random-parameters negative-binomials model	Annual average hourly traffic	Population Walking frequency Walking time	×	✓	×	✓	×
Rothman et al. (2020)	Toronto, Canada	2001–2010	102 census tracts	Multivariate logistic regression model	Roadway length	Population	×	✓	×	✓	×

POI: points of interest.

CAR: conditional autoregressive prior.

Table A2. Walking speeds estimated for Hong Kong residents based on the 2011 Hong Kong Travel Characteristics Survey data, stratified by sex and age groups (unit: m/s).

Age groups	≤ 14	15–24	25–34	35–44	45–54	55–64	≥ 65
Male	1.00	1.01	0.92	0.91	1.05	1.00	0.91
Female	1.06	1.14	1.05	0.99	0.96	1.00	0.82

The WinBUGS code for the BSVC-2 model:

```
Model
{
for (i in 1:209) {
    y[i] ~ dpois(lambda[i])
    log(lambda[i]) <- beta[1]+beta[2]*LnVKT[i]+beta[3]*LnW_trp[i]+beta[4]*
P_Res_LU[i]+beta5[i]*D_Rd[i]+beta[6]*P_Mt_Rd[i]+beta[7]*D_Jun[i]+beta[8]*
P_Rad_Jun[i]+ beta[9]*D_Bus[i]+beta[10]*Parking[i]+beta[11]*MINC[i]+s[i]
    beta5[i] <- beta[5]+phi5[i]
    ypred[i] ~ dpois(lambda[i]) # Predictive value based on the complete data
    PPL[i] <- abs(ypred[i]-y[i])
    PPL2[i] <- pow(ypred[i]-y[i],2)
}
MAD <- mean(PPL[])
MSPE <- mean(PPL2[])
# Proper CAR prior of Leroux et al. (1999). Detailed explanations for the specification
of this CAR prior should be referred to Congdon (2014; page 337–343).
for (i in 1:209) {
     s[i] ~ dnorm(s.bar[i],tau[i])
     s.bar[i] <- rho.s*sum(Wsp.s[cumsum[i]+1:cumsum[i+1]])/(1-rho.s+rho.s*num[i])
     tau[i] <- tau.s*(1-rho.s+rho.s*num[i])
     phi5[i] ~ dnorm(phi5.bar[i],tau.phi5[i])
     phi5.bar[i] <- rho.5*sum(Wsp.5[cumsum[i]+1:cumsum[i+1]])/(1-rho.5+rho.5*num[i])
     tau.phi5[i] <- tau.5*(1-rho.5+rho.5*num[i])
}
for (i in 1:sumNumNeigh) { # sumNumNeigh refers to the length of adj[]
    Wsp.s[i] <- s[adj[i]]
    Wsp.5[i] <- phi5[adj[i]]
}
rho.s ~ dunif(0,1) # Correlation parameter
tau.s <- pow(sig.s,-2)
var.s <- pow(sig.s,2)
sig.s ~ dunif(0,10)
rho.5 ~ dunif(0,1)
tau.5 <- pow(sig.5,-2)
var.5 <- pow(sig.5,2)
sig.5 ~ dunif(0,10)
for (k in 1:11) {beta[k] ~ dnorm(0.0,1.0E-5)}
}
```