Accident Analysis and Prevention 128 (2019) 17–24

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Exposure to pedestrian crash based on household survey data: Effect of trip purpose

N.N. Sze*, Junbiao Su, Lu Bai

Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

ARTICLE INFO

Keywords:
Pedestrian safety
Exposure
Travel purpose
Crash prediction model

ABSTRACT

Pedestrian are vulnerable to severe injury and mortality in the road crashes. Understanding the essence of the pedestrian crash is important to the development of effective safety countermeasures and improvement of social well-being. It is necessary to measure the exposure for the quantification of pedestrian crash risk. The primary goals of this study are to explore the efficient exposure measure for pedestrian crash, and identify the possible factors contributing to the incidence of pedestrian crash. In this study, amount of travel was estimated based on the Travel Characteristic Survey (TCS) data in 2011, and the crash data were obtained from the Transport Information System (TIS) of the Hong Kong Transport Department during the period from 2011 to 2015. Total population, walking frequency and walking time were adopted to represent the pedestrian exposure to road crash. The effect of trip purpose on pedestrian crash was evaluated by disaggregating the pedestrian exposure proxies by purpose. Three random-parameter negative binomial regression models were developed to compare the performances of the three pedestrian exposure proxies. It was found that the model in which walking frequency was used as the exposure proxy provided the best goodness-of-fit. Frequency of walking back home, among other trip purposes, was the most sensitive to the increase in pedestrian crash risk. Additionally, increase in the frequency of pedestrian crash was correlated to the increases in the proportions of children and elderly people. Furthermore, household size, median household income, road density, number of non-signalized intersection as well as number of zebra crossings also significantly affected the pedestrian crash frequency. Findings of this study should be indicative to the development and implementation of effective traffic control and management measures that can improve the pedestrian safety in the long run.

1. Introduction

There are more than 1.2 million road deaths every year. Pedestrians are vulnerable to death and severe injury on roads (Zegeer and Bushell, 2012; Wang and Kockelman, 2013; World Health Organization, 2015; Cai et al., 2016). 11% and 16% of road deaths were pedestrians in the United States and Canada, respectively (Miranda-Moreno et al., 2011). Particularly, in Hong Kong, 64% of road deaths were pedestrians in 2016. The likelihood of mortality of pedestrian on road was 17 times higher than that of vehicle passengers and 6 times higher than that of drivers (Lam et al., 2013; Hong Kong Transport Department, 2016). On the other hand, over 89% of daily commuter trips in Hong Kong are made by public transport where walking is the primary access to and from the public transit stations (Hong Kong Transport Department, 2012b; Li and Loo, 2016). What's more, walkability has attracted the attention from urban and transport planners where pedestrian safety is one of the key attributes. To this end, it is crucial to identify the factors contributing to the higher risk of pedestrian injury and mortality on roads (Abdel-Aty et al., 2013; Lee et al., 2015a; Amoh-Gyimah et al., 2016; Cai et al., 2016, 2017).

A number of studies have attempted to model the pedestrian crash frequency (Miranda-Moreno et al., 2011; Siddiqui et al., 2012a,b; Lam et al., 2014; Lee et al., 2015b; Amoh-Gyimah et al., 2016; Ma et al., 2017). Efficient exposure measurement is necessary for better understanding of safety assessment among different (space and time) entities. Vehicular traffic volume and its derivatives like vehicle-kilometer and vehicle-hour are commonly adopted as the exposure for vehicular crash modelling, given the advances in automatic vehicle detection techniques (Qin et al., 2006; Pei et al., 2012). However, it remains difficult to obtain accurate pedestrian counts, especially for long-term and extensive monitoring and evaluation. Despite that the collection of high-quality pedestrian count data may be possible at specified locations, it is rare that robust measure of the exposure, in terms of the walking frequency and time, for macroscopic pedestrian crash modeling has been established.

* Corresponding author.
E-mail addresses: tony.nn.sze@polyu.edu.hk (N.N. Sze), junbiao.su@connect.polyu.hk (J. Su), l.bai@polyu.edu.hk (L. Bai).

https://doi.org/10.1016/j.aap.2019.03.017
Received 10 September 2018; Received in revised form 27 March 2019
0001-4575/

Additionally, travel purpose could also affect the pedestrian behavior and therefore the pedestrian crash incidence (Bao et al., 2017). Travel is indeed a derived demand, either direct or indirect, of economic activities including work, education, trade and recreation. For example, a work-related activity involves travel between the place of residence and the workplace. There is a supply of labour in one location (residence) and a demand for labour in another (workplace). Therefore, travel (commuting) is directly derived from the relationship between the two locations. To this end, not only the time, location and mode of movements of people are of interest, but also the need of access to economic activities (e.g. work, education and recreation, etc.) should be of concern. Activity-based transport demand modeling has been receiving more and more attention in the past decades (Bowman, 1998; Bhat and Koppelman, 1999; Bhat and Singh, 2000; Bowman and Ben-Akiva, 2001; Recker, 2001; Miller and Roorda, 2003a; Ettema et al., 2004; Bhat et al., 2013). In conventional pedestrian crash prediction model, population has been widely used as the proxy to the crash exposure (Wier et al., 2009; Cottrill and Thakuriah, 2010). However, travel demand could vary among different population groups, and more specifically by time, location and activity purpose. It is therefore of essence to estimate the pedestrian crash exposure based on the derived transport demand by different travel purposes (e.g. economic activities) (Chliaoutakis et al., 1999, 2005; Abdel-Aty et al., 2013; Elias and Shiftan, 2014; Lee et al., 2014, 2015a). Furthermore, possible trip generation factors that could affect the accessibility to different economic activities, and then the pattern of travel activities by purposes and therefore the crash risk, should be examined (Naderan and Shahi, 2010; Siddiqui et al., 2012a,b; Bao et al., 2017; Zou et al., 2017).

This study attempts to explore efficient exposure measure for pedestrian crash and to identify the factors contributing to the incidence of pedestrian crash at the zonal level. Additionally, the moderating effects on the association between pedestrian crash, exposure and possible risk factors by travel purpose would be attempted.

The remainder of the paper is organized as follows. Section 2 provides a review of literature on pedestrian safety. Section 3 describes the data. Methodology and model formulation are presented in Section 4. Section 5 shows the results of the crash prediction models. After that, discussion is presented in Section 6. Lastly, Section 7 summarizes the study and provides the recommendations for future study respectively.

2. Literature review

Prevalence of pedestrian crash is one of the main streams for the pedestrian safety analysis. For the prevalence of pedestrian crash at the zonal level, increase in the population density was often found correlated to the increase in pedestrian crash frequency (Wang and Kockelman, 2013; Wang et al., 2017). Land use and transport network characteristics also affected the pedestrian safety. For instances, increase in the proportion of commercial area, road density, number of signalized intersections can increase the probability of pedestrian crash, considering the increase in the activity amount and conflicts between pedestrian and vehicular traffic (Loukaitou-Sideris et al., 2007; Wier et al., 2009). Household demographic and socio-economic characteristics including car ownership, age and household income were correlated to the prevalence of pedestrian crash (Noland and Quddus, 2004; Ponnaluri and Nagar, 2010; Lee et al., 2015b). Increase in the number of jobs was found to increase the likelihood of pedestrian crash occurrence (Kim et al., 2006; Lee et al., 2015b). The abovementioned factors can indeed affect the need of travel (derived demand of economic activity), and thus the travel behavior and crash risk. It is necessary to consider the moderating effect of the amount of travel by purpose, when measuring the association between pedestrian safety and possible factors.

Population and population density are often adopted to represent the pedestrian exposure because they can be extracted easily and directly from the census dataset (Wier et al., 2009; Cottrill and Thakuriah, 2010; Siddiqui et al., 2012a,b; Lee et al., 2015b; Wang et al., 2017). However, using population as the proxy does not consider the habit, and thus the differences in the amount of travel and behavior among different population groups. The results of association measure using population as the exposure could be biased.

Another alternative for measuring the exposure to pedestrian crash is the number of personal trips estimated using the conventional four-step model (Bao et al., 2017). However, the transport mode of every trip leg within a single trip had to be distinguished (Allsop, 2005). For instances, a multi-modal trip (home – metro – bus – office) could generate three walking trip legs, taking into account the access to public transport station and transfer between modes. Therefore, the exposure could be underestimated when only the number of Origin-Destination trips is used as a proxy (Lam et al., 2014). On the other hand, observational pedestrian counting approach may not be able to incorporate the effect of pedestrian characteristics including demographics and travel purpose when evaluating the relationship between pedestrian behavior and safety (Lam et al., 2014).

Travel characteristics survey through household interview can be exploited as a way for estimating the pedestrian exposure. It is possible to deduce the origin, destination and time of each walking trip (and trip leg) from the household travel survey data. More importantly, attributes of every trip could be linked to the personal and household characteristics of the commuter. This could help improve the understanding on the association between pedestrian crash occurrence and personal factors (Allsop, 2005; Lam et al., 2013; Papadimitriou, 2016). The effect of travel purpose on the risk of vehicular crashes had received much attentions (Elias et al., 2010; Naderan and Shahi, 2010; Lam et al., 2013). Indeed, trip purposes may also affect the pedestrian behavior, in terms of walking path choice, walking speed, crossing location and time, and propensity of convicted crossing behavior (Sze and Wong, 2007; Lavieri, 2018). For instances, distraction by mobile phone use and social interactions were correlated to the prevalence of pedestrian crash at the crosswalk (Elias and Shiftan, 2014). Therefore, the intervention effect of trip purposes for the association between pedestrian crash, exposure and possible factors should not be ignored. Temporal variations in the environment, traffic and road user behavior could also affect the association between pedestrian safety and possible risk factors (Mason, 1979). Attempts had been made to examine the temporal effect on safety, with time scales varying from annually (Castro et al., 2012), daily (Mason, 1979; Ma et al., 2017) to hourly (Lam et al., 2013). Spatial and temporal variations in the exposure are crucial to the association measure between pedestrian crash and possible factors (Pahukula et al., 2015). Some examples of pedestrian exposure measures used in previous studies are summarized in Table 1.

This study attempted to address the above issues by measuring the time-varying pedestrian exposure, in the terms of walking frequency and walking time, aggregated at zonal level for the pedestrian crash prediction model. The proposed method also considered the effect of trip purpose, personal and household characteristics on pedestrian safety, based on the data captured from the household travel survey.

3. Data

The dataset is established by integrating the information captured from different sources: (i) The police crash record from the Traffic Information System (TIS) maintained by the Hong Kong Transport Department; (ii) the household travel data from the 2011 Travel Characteristic Survey (TCS) and Annual Traffic Census (ATC) data from the Hong Kong Transport Department; (iii) the data of population census from the Hong Kong Statistics and Census Department; and (iv) the zoning scheme and land use planning data from the Hong Kong Planning Department. The above-mentioned data were matched to the specified spatial unit, Tertiary Planning Unit (TPU), using the Geographical Information System (GIS) technique (Lee et al., 2014,b; Cai et al., 2016, 2017). The entire territory of Hong Kong is divided into

Table 1
Some examples of pedestrian crash exposure measures.

Exposures	References	Advantages	Limitations
Population / population density	Wier et al., 2009 Cottrill and Thakuriah, 2010 Siddiqui et al., 2012a Lee et al., 2015b Wang et al., 2016	• Easy to extract	• Not considering the difference in the amount of travel between individual
Pedestrian volume from observational survey	Lam et al., 2014	• Considering the spatial and temporal distributions of pedestrian exposure	• Not considering the effects of demographics and other personal characteristics
Trip generation estimates from four-step model	Bao et al., 2017	• Considering the effects of land use, demographics and socio-economic characteristics on travel demand	• Not considering the multi-modal (walking and non-walking) trip legs
Walking frequency from household travel survey	Elias et al., 2010 Naderan and Shahi, 2010 Lam et al., 2013	• Considering the effects of demographics and travel purpose on safety • Considering the temporal effect on safety	• Should consider multi-modal (vehicle and pedestrian) exposures

289 TPUs. Considering the problem of prevalence of zero crash count when the data is highly disaggregated (of small zone size and/or short time period), the 289 TPUs were aggregated into 26 zones.

In this study, 12,672 pedestrian crashes occurred on the weekdays during the period from 2011 to 2015 were captured from the TIS. TIS database consists of three profiles, namely crash environment, vehicle attributes, and casualty characteristics. Since the precise information on location and time of every single crash is available, it is possible to allocate the crashes into units by space (26 zones) and time (24 h) when evaluating the spatial-temporal variations in pedestrian crash incidence. Finally, there are 624 units for the proposed pedestrian crash prediction models. The distribution of pedestrian crash (by zone) is depicted in Fig. 1.

As shown in Fig. 1, the pedestrian crashes are concentrated in Central Kowloon and North shore of Hong Kong Island (colored in red). It is not surprising since they are the Central Business Districts (CBDs), and the interactions between pedestrian and vehicular traffic are prevalent. This justifies the need to estimate the pedestrian exposure, in terms of the amount of travel by walking.

For the pedestrian exposure, the walking frequency and walking time were estimated based on the Travel Characteristics Survey (TCS) which was conducted by the Hong Kong Transport Department. The survey was conducted during the period from September 2011 to March 2012. Of the 2.36 million households in Hong Kong, 35,401 households (1.5%) were surveyed in TCS. The survey captured the details of 136,122 trips (including 122,237 motorized trips and 13,885 walk-only trips) made by 101,385 people during the weekdays. Such 136,122 reported trips of selected households could be projected to that of all households in Hong Kong using a set of expansion factors, in

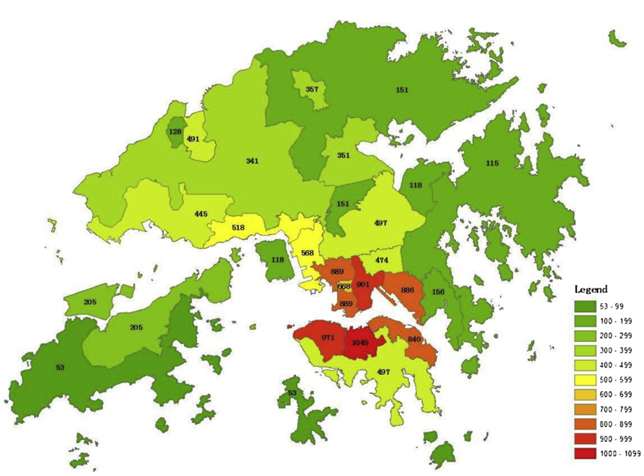

Fig. 1. Spatial distribution of pedestrian crash.

accordance to the demographics and socio-economic characteristics of the households. For instances, the 122,237 motorized trips reported could be transformed into 12.6 million person-trips per day for the whole Hong Kong. To eliminate the biases of trip estimates attributed to the under-reporting of surveyed households, the trip estimates were validated using the independent datasets including the screenline traffic counts and public transport ridership statistics (Hong Kong Transport Department, 2012a).

A motorized trip was defined as the trip that involved at least one motorized mode (i.e. bus, metro, taxi and car, etc.). One motorized trip can have more than one walking trip leg (e.g. variable of interest in current study). For each trip, the information on origin location, destination location, departure time, arrival time and transport mode of each trip leg was captured. In this study, walking frequency refers to the total number of walking trip legs, and walking time refers to the total time spent on walking respectively. The total walking frequency and walking time for the whole Hong Kong were estimated using the expansion factors concerned. For instances, the total walking frequency derived was 28.7 million per day. The distribution of walking frequency by the 624 analysis units (i.e. 26 zones x 24 h) is depicted in Fig. 2. As shown in Fig. 2, four analysis units had zero walking trip (leg). They were all in the rural area and before dawn (e.g. from 2.00 a.m. to 4.59 a.m.).

Variations in the pedestrian crash frequency and walking frequency, over time of the day, are depicted in Fig. 3. The peak walking frequencies occurred during the periods from 8:00 am to 8:59 am and from 6:00 pm to 6:59 pm, respectively. Except during the period from midnight to dawn (e.g. from 11.00 pm to 5:59 am), walking frequency was the lowest in the middle of the day while the pedestrian crash frequency was consistently high throughout the period from 8:00 am to 7:00 pm. Because of the discrepancy between crash and walking frequency, bias may exist in the association measure when only the aggregated walking frequency was used as the exposure for the pedestrian crash prediction model. In this study, the walking frequency and walking time of different trip purposes in different time periods were considered. Specifically, trip purpose could be categorized into six: home, work, school, shopping, dining and others (Miller and Roorda, 2003a,2003b; Bhat et al., 2013; Allahviranloo and Recker, 2015).

Taking into account the effect of multi-modal (vehicle and pedestrian) exposure, the vehicular traffic was also considered in the proposed crash prediction model. For instances, the information on annual average daily traffic (AADT) (and hourly count) was obtained from the Annual Traffic Census (ATC) dataset in 2011 maintained by the Hong Kong Transport Department. ATC covers 87% of all public roads in Hong Kong. Additionally, information on age, employment status, household size and household income were obtained from the population census. Furthermore, the transport network and traffic control attributes including number of signalized and non-signalized intersections, and number of zebra crossings, were also incorporated into the

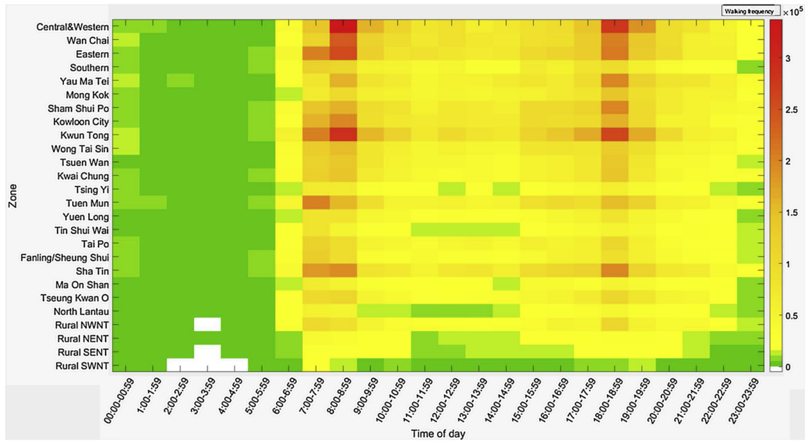

Fig. 2. Space-time distribution of walking frequency.

proposed model. The above data was matched with the crash, traffic count and pedestrian exposure data using the GIS technique. Table 2 summarizes the characteristics of the candidate variables.

As shown in Table 2, walking frequency and time of back home trip were predominant among the six travel purposes. Back home trip constituted to 45% of walking frequency and 46% of walking time respectively. For the mandatory travel (e.g. commuting), work trip constituted to 25% of walking frequency and 24% of walking time respectively. For the demographic characteristics, vulnerable group (i.e. elderly and children) constituted to one-fourth of the overall population. In Hong Kong, the average household size was 2.8 person, and the proportion of household with more than three members was 34%.

Prior to the association measure between pedestrian crash, exposure and possible factors, the variance inflation factor (VIF) was estimated to identify any possible collinearity among the variables. The candidate variables with VIF greater than five (e.g. proportion of unemployment, number of signalized intersections) were not considered in the consolidated model. As the main purpose of current study is to identify the significant contributory factors (and their effects) to pedestrian crash, it is necessary to exclude the (theoretically less important) correlated variables, even which the estimates of remaining correlated variables might be biased (Tay, 2017).

4. Method

Poisson and negative binomial regression models are the most commonly used approaches for crash occurrence (Wong et al., 2007). Results of over-dispersion test indicated that the data was subject to significant over-dispersion at the 1% level. Therefore, negative binomial regression model was preferred to the Poisson model.

Negative binomial regression can be derived from the Poisson model by introducing an unobserved heterogeneity error term τ, following the gamma distribution with mean of 1.

$$y_{it} \sim f(y_{it}|\mu_{it}, \tau_{it}) = \frac{\exp(-\mu_{it}\cdot\tau_{it})(\mu_{it}\cdot\tau_{it})^{y_{it}}}{y_{it}!} \tag{1}$$

Let $E(y_{it}) = \mu_{it}$ be the expected number of pedestrian crashes in zone i (i =1,2,3…,26) for time period t (t =1,2,3…,24). The density function of negative binomial distribution $p(\cdot)$ can be further derived as:

$$p(y_{it}) = \frac{\Gamma(y_{it} + \alpha^{-1})}{y_{it}\Gamma(\alpha^{-1})}\left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu_{it}}\right)^{\alpha^{-1}}\left(\frac{\mu_{it}}{\alpha^{-1} + \mu_{it}}\right)^{y_{it}} \tag{2}$$

where the mean of y_{it} was given by $E(y_{it}) = \mu_{it}$ and $\text{var}(y_{it}) = \mu_{it}(1 + \mu_{it})$. $\Gamma(\cdot)$ stands for Gamma function and α presents the over-dispersion parameter. Furthermore, μ_{it} was given as follow:

$$\ln(\mu_{it}) = \beta_0 + \boldsymbol{x}_{it}^{T}\cdot\boldsymbol{\beta} \tag{3}$$

where $\boldsymbol{x}_{it}$ was the column vector of exogenous variables corresponding to entity i and time period t. $\boldsymbol{\beta}$ was a vector of parameters of $\boldsymbol{x}_{it}$ while β_0 is the constant term.

For the conventional count data (Poisson and negative binomial) models, the parameters were assumed to be fixed. However, as the crash data were aggregated over specified zone and time period, the effect of a possible factor on crash incidence may vary. For example, risk perception could affect the pedestrian behavior, and therefore contribute to crash incidence. The association between risk perception and crash frequency was not revealed in the aggregated data. In this case, risk perception varied across different individuals, but was not explicitly observed. It was known as unobserved heterogeneity. Therefore, assuming the parameters to be constant across observations

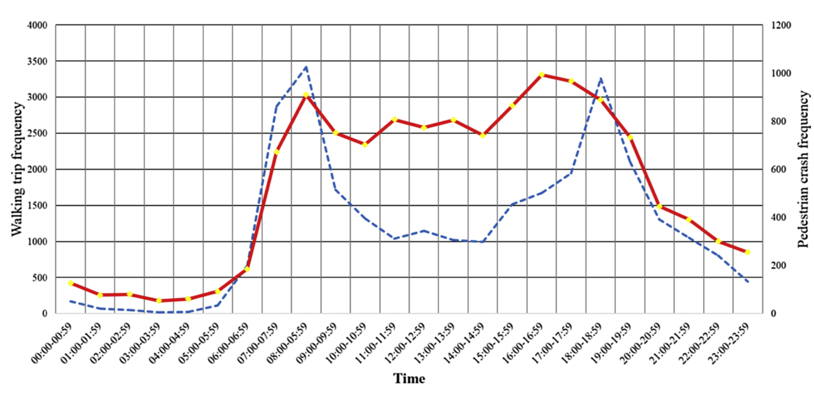

Fig. 3. Temporal distributions of pedestrian crash frequency and walking frequency.

Table 2
Summary statistics of the variables used in current study.

Variable	Mean	Std. Dev.	Min	Max
Pedestrian crash count	20.31	20.69	0	89
Walking frequency				
Home	20,711.00	29,755.74	0	219,524
Work	11,282.38	26,590.24	0	288,345
School	3,725.92	10,508.50	0	92,389
Shopping	3,792.33	5,023.76	0	32,209
Dining	2,547.89	4,251.15	0	35,439
Others	3,953.78	4,699.43	0	25,726
Walking time (minute)				
Home	106,281.40	156,019.10	0	1,360,410
Work	54,653.49	135,696.30	0	1,410,589
School	18,066.02	53,136.65	0	544,495
Shopping	21,320.60	28,749.95	0	168,726
Dining	11,684.00	20,491.71	0	205,374
Others	19,035.26	22,239.84	0	124,815
Annual average hourly traffic	1,514	1,098	10	4,646
Zonal population	271,938	146,182.50	57,338	622,152
Proportion of population				
Age below 15 years	0.12	0.01	0.10	0.15
Age 15-64 years	0.75	0.26	0.72	0.80
Age above 64 years	0.13	0.03	0.07	0.18
Unemployed	0.33	0.03	0.27	0.37
Non-Chinese	0.08	0.06	0.02	0.21
Proportion of household with more than 3 members	0.34	0.04	0.25	0.40
Median household income	23,365	5,461.49	7,500	12,500
Road density (km/km^2)	9.18	7.58	0.33	31.28
Number of signalized intersections	56.81	39.60	0	141
Number of non-signalized intersections	798.19	342.02	156	1,544
Number of zebra crossings	10.23	8.89	0	31

could result in biased parameter estimations when they were indeed varying across observations (Savolainen et al., 2011; Xiong and Mannering, 2013; Mannering et al., 2016).

To capture the variation in the effect of possible factor on the crash incidence across individuals, the random parameter approach has been applied (Anastasopoulos and Mannering, 2009). Specifically, parameter corresponding to the specific factor is no longer treated as constants but vary cross individuals (Train, 2009). Assumed that $\boldsymbol{\beta}$ was associated with the density function of $f(\boldsymbol{\beta})$, Eq. (3) becomes:

$$P(y_{it}) = \int p(y_{it}|\boldsymbol{\beta})f(\boldsymbol{\beta})d\boldsymbol{\beta} \tag{4}$$

If random parameter vector $\boldsymbol{\beta}$ collapsed to the case that only constant was found randomly distributed, random parameter model was consistent to random-effect model. Halton draws have been used to solve the random-parameter models and provided stable results (Train, 2001). Three indicators are used for evaluating the goodness-of-fit of the random-parameter models, namely Pseudo ρ^2, Akaike information criterion (AIC) and Bayesian information criterion (BIC).

5. Estimation results

Three random-parameter negative binomial regression models were established to measure the association between pedestrian crash frequency and possible risk factors. The exposure to pedestrian crash was proxied by the zonal population in Model 0, walking frequency in Model 1 and walking time in Model 2 respectively. Results of the three estimated models were given in Table 3.

As shown in Table 3, Pseudo ρ^2 of Model 0, Model 1 and Model 2 were 0.04, 0.27 and 0.26, AICs of Model 0, Model 1 and Model 2 were 4,837.80, 3,714.35 and 3,751.12, and BICs of Model 0, Model 1 and Model 2 were 4,875.73, 3,798.64 and 3,824.54, respectively. Model 1 was superior to the two other counterparts, given the highest Pseudo ρ^2 while the lowest AIC and BIC. This indicated that model using the walking frequency as the proxy for pedestrian exposure had the best prediction performance.

For the model using walking frequency as proxy for pedestrian exposure (Model 1), as shown in Table 3, walking frequency (for back home, work, study and shopping), traffic flow, proportion of population age below 15, age above 64, proportion of household with more than 3 members, median household income, road density, number of non-signalized intersection and zebra crossing all significantly affected the crash frequency, at the 1% level. In particular, increase in walking frequency (coefficient ranged from 0.03 to 0.25), proportion of vulnerable population (age below 15, coefficient = 8.02; and above 64, coefficient = 4.15), median household income (0.002), road density (0.02), and number of non-signalized intersection (0.04) and zebra crossing (0.01), were correlated to the increase in pedestrian crash frequency. In contrast, increases in traffic flow (-0.03) and household with more than 3 members (-2.44) were correlated to the reduction in pedestrian crash frequency. Additionally, effects of proportion of young population, and household with more than 3 members and median household income on crash frequency varied across the observations. Similar results could be revealed for the model using walking time as the proxy for pedestrian exposure (Model 2).

Table 4 illustrates the elasticity estimates of walking frequency by travel purpose. Increase in pedestrian crash frequency is less proportionate to that of walking frequency, as the elasticities are all less than one (0.25, 0.11, 0.08, 0.07 and 0.03 for back home, work, school, shopping and other trips respectively). Nevertheless, the effect of back home travel is dominant among other travel purposes.

6. Discussion

In this study, pedestrian crash prediction model using walking frequency as the proxy for pedestrian exposure was found superior to that using zonal population and walking time. Though population is often used as the proxy for exposure as it is readily available in the census dataset, it is not necessary that every individual would make the same number of (walking) trips in a specified period of time. On the other hand, population could only reflect the number of residents in the area concerned, it should not necessarily reflect the exposure of commuters

Table 3
Estimation results of random parameter negative binomial models.

Variables	Model 0		Model 1		Model 2	
	Coef.	p-value	Coef.	p-value	Coef.	p-value
ln (Walking frequency)						
Home			0.25	< 0.001		
Work			0.11	< 0.001		
School			0.08	< 0.001		
Shopping			0.07	< 0.001		
Dining			0.02	0.07		
Others			0.03	0.02		
ln (Walking time)						
Home					0.22	< 0.001
Work					0.09	< 0.001
School					0.06	< 0.001
Shopping					0.05	< 0.001
Dining					0.03	< 0.001
Others					0.02	< 0.001
Annual average hourly traffic	0.09	0.06	−0.03	0.01	−0.02	0.03
Zonal population	1.59	< 0.001				
Proportion of population						
Age below 15 years			8.02	< 0.001	6.78	< 0.001
std. of age below 15 years			*1.47*	*< 0.001*	*2.09*	*< 0.001*
Age above 64 years	6.53	0.01	4.15	< 0.001	3.92	< 0.001
Proportion of household with more than 3 members	−4.77	0.003	−2.44	< 0.001	−2.33	< 0.001
std. of household with more than 3 members			*0.45*	*< 0.001*		
Median household income	0.003	0.001	0.002	< 0.001	0.003	< 0.001
std. median household income			*0.45*	*< 0.001*	*0.000*	*< 0.001*
Road density (km/km^2)	0.03	< 0.001	0.02	< 0.001	0.03	< 0.001
Number of non-signalized intersection	0.02	< 0.001	0.04	< 0.001	0.04	< 0.001
Number of zebra crossing	−0.007	< 0.001	0.01	< 0.001	0.01	< 0.001
Constant	−0.214	0.320	−3.22	< 0.001	−3.39	< 0.001
Log likelihood with constant only	−2514.17		−2514.17		−2513.17	
Log likelihood at convergence	−2408.90		−1838.17		−1857.56	
Number of observations	624		624		624	
Pseudo ρ^2	0.04		0.27		0.26	
AIC	4837.80		3714.35		3751.12	
BIC	4875.73		3798.64		3824.54	

Table 4
Elasticity of pedestrian crash frequency.

Variable	Elasticity		
Walking frequency	Mean	t-statistics	(*p*-value)
Home	0.25	26.52	(< 0.001)
Work	0.11	11.47	(< 0.001)
School	0.08	10.53	(< 0.001)
Shopping	0.07	6.89	(< 0.001)
Dining	0.02	1.81	(0.069)
Other	0.03	2.28	(0.023)

travelling in the area.

Besides, model using walking frequency as the proxy better reflected the prevalence of vehicular-pedestrian interaction, as compared to that using walking time. For instances, increase in the number of walking trip by ten times was more likely correlated to the increase in the frequency of vehicular-pedestrian interaction by the same extent, however, it might not be the case for increasing the walking time by ten times (it could be attributed to the increase in the time spent on the footpath or waiting at the crosswalk) (Siddiqui et al., 2012a,b; Cai et al., 2017; Wang et al., 2017). On the other hand, in the household travel survey, the reliability of (self-reported) information on walking frequency could be higher than that of walking time, because of the variation of time perception among individuals and thus the unreliability of self-reported walking time (Chu, 2003; Greene-Roesel et al., 2007). On the other hand, increase in traffic flow was correlated to the reduction in pedestrian crash. It could be attributed to the increase in safety awareness of the pedestrian due to the increase in traffic volume. Then, the propensity of reckless crossing behavior could be reduced, and therefore, the risk of pedestrian crash reduced.

6.1. Travel purpose

Increase in the pedestrian exposure was found correlated to the increase in the frequency of pedestrian crash. However, as shown in Fig. 2, the pedestrian crash rates (per unit walking frequency) were remarkably high during the noon (12.00 pm – 1.59 pm) and late afternoon (3:00 pm- 4:59 pm), even though the overall walking frequency was very low during the periods. It could be attributed to the reckless behavior of the pedestrian, since it is the prime time for back home travel during the periods (Clarke et al., 2006; Anderson, 2008; Devlin et al., 2010; Liu and Yue, 2011; Palamara and Broughton, 2013). Considering the distribution of walking frequency by travel purpose, back home travel was dominating during the noon (about 35%) and late afternoon (about 75%) periods. Indeed, pedestrian inattentiveness was one of the top contributory factors to road crashes in recent years (Hong Kong Transport Department, 2012-2016Hong Kong Transport Department, 2012-2016). These are consistent to the findings of earlier studies that the travel purpose and time of the day can affect the travelers' behavior and thus the safety level on roads (Elias et al., 2010; Naderan and Shahi, 2010). This could shed light on the setting of road safety target, and then better planning of road safety strategies that can enhance the safety of vulnerable pedestrian group at specified time periods and locations (Wong et al., 2006; Sze et al., 2014; Sze and Christensen, 2017). In particular, it is necessary to improve the road user education and enforcement measures that could enhance the safety awareness of pedestrian on their way back home (Wong et al., 2008). Yet, the association between safety perception and safety risk of individual pedestrian was blurred when the crash data was aggregated by time and zone. It would be worth exploring the factors affecting the safety perception of pedestrian and thus crash involvement using observational and/or perceptional survey.

6.2. Socio-demographics

Age could also affect the frequency of pedestrian crash. In particular, increases in the proportions of children and elderly people were correlated to the increase in the frequency of pedestrian crash (Fontaine and Gourlet, 1997; Noland and Quddus, 2004; Palamara and Broughton, 2013; Jiménez-Mejías et al., 2016). It could be attributed to the variations in the obedience of traffic rule, risk perception and safety awareness across different age groups (Ponnaluri and Nagar, 2010; Siddiqui et al., 2012a,b). In particular, the children and younger teens had a higher propensity to involve in reckless behavior on roads (Tay, 2003). For the elderly pedestrian, the high crash risk could be attributed to the degradation of cognitive performance, gap acceptance and impaired mobility (Fontaine and Gourlet, 1997; Noland and Quddus, 2004; Palamara and Broughton, 2013; Jiménez-Mejías et al., 2016). Therefore, targeted road user education programs should be conducted in the schools and community centers.

Increase in household median income was found correlated to the increase in pedestrian crash incidence. It could be attributed to the difference in safety perception and choice decision across population of different income levels. In particular, the trade-off between travel cost and crash risk could vary with the value of time of commuter (Tay, 2003). Such finding could be indicative to the planning and design of built environment, especially for the enhancement of accessibility, and therefore the safety performance, in accordance to the need of different population demographics and socio-economics (AECOM Asia, 2010; Mannings (Asia) Consultants, 2010; MVA Hong Kong, 2013). Additionally, increase in the proportion of household with more than 3 members was found correlated to the reduction in the frequency of pedestrian crash. It could be because the vulnerable road users (i.e. children and elderly) are usually accompanied by other household members, in particular the domestic helper or health care taker, for the household with more members, which was commonly seen along the streets in Hong Kong. It could then enhance the safety awareness and reduce the crash risk on roads (Huang et al., 2010; Cai et al., 2017). This is indicative to the road safety education and promotional strategies targeted to the caretakers of vulnerable pedestrian group (*Sze and Christensen, 2018*).

6.3. Transport system characteristics

Last but not the least, transport system characteristics including road density (road length per unit area), number of non-signalized intersection and number of zebra crossing could affect the pedestrian safety. Specifically, increase in road density was found correlated to the increase in pedestrian crash. It could be attributed to the increase in the potential vehicular-pedestrian interactions, and thus the incidence of pedestrian crash. This should also be applicable to the association between pedestrian crash, number of non-signalized intersection, and number of zebra crossing (Varhelyi, 1997). This implied that installation of signal control at the intersection and crosswalk (segregating the conflicting vehicular and pedestrian streams by time and space) might effectively reduce the incidence of pedestrian crash. (Castro et al., 2012; Siddiqui et al., 2012a,b). This is indicative to the planning of pedestrian infrastructure and traffic management & control strategies that could mitigate the vehicular-pedestrian interaction, and therefore the pedestrian crash potential (Sze and Wong, 2007). It was noticed that the signs of annual average hourly traffic and number of zebra crossing were difference across the three models. When the amount of walking is not considered (using zonal population as the proxy), it is expected that the crash incidence would increase with traffic volume and decrease with the presence of zebra crossing. The pedestrian perception and behavior are indeed correlated to the traffic volume and traffic control. Therefore, when the amount of walking is used as the proxy (effect of pedestrian behavior is incorporated implicitly), the effects of traffic volume and presence of zebra walking could be reversed. Therefore, increase in pedestrian crash is correlated to the reduction in annual average hourly traffic and the presence of zebra crossing.

7. Conclusion and future research

This paper contributes to the literature by developing a pedestrian crash prediction model, at zonal level, that takes into account the effects of pedestrian exposure and travel purpose. The contribution of this paper is twofold. Firstly, we sought an alternative approach for measuring the pedestrian exposure based on the comprehensive household travel characteristics survey data. Therefore, the effect of the difference in the amount of walking across population groups on crash incidence can be considered. In particular, the model that applied the walking frequency as the proxy to pedestrian exposure was superior to that using zonal population and walking time. Secondly, the effect of travel purpose on pedestrian crash risk was examined. The exposure was discretized by hours of the day and travel purposes. Results indicated that the crash risk of back home walking trip was more profounding than that of other travel purposes. Additionally, effects of confounding factors including age, household size, household income, road density and intersection control on pedestrian safety were revealed. Results should be indicative to the targeted road user education, enforcement and traffic control measures that could enhance the safety awareness of vulnerable pedestrian group, and therefore improve the pedestrian safety in the long run. This is an essence to the promotion of walkability in an aging society like Hong Kong.

Yet, in the current study, the association between pedestrian crash incidence, exposure and possible risk factors was measured using the aggregated data by time and zone. Besides, we only considered the pedestrian crash exposure on weekdays, given the availability of travel data. It would be worth exploring the effects of demographics, socio-economics, built environment, and multi-modal (vehicular and pedestrian) traffic flow on the safety perception, and in turn the propensity of reckless behavior of pedestrian, when the trajectory data of individual vehicle and pedestrian are available through extensive observational survey in the extended study.

Acknowledgements

The work that is described in this paper was supported by the grants from the Research Grants Council of Hong Kong (15209818) and Research Committee of the Hong Kong Polytechnic University (1-ZE5V).

References

Abdel-Aty, M., Lee, J., Siddiqui, C., Choi, K., 2013. Geographical unit based analysis in the context of transportation safety planning. Transp. Res. Part A Policy Pract. 49, 62–75.
AECOM Asia, 2010. Study on Pedestrian Subways and Related Traffic Improvement Measures in Causeway Bay - Executive Summary. Hong Kong Transport Department.
Allahviranloo, M., Recker, W., 2015. Mining activity pattern trajectories and allocating activities in the network. Transportation 42, 561–579.
Allsop, R.E., 2005. Travel survey data required to inform transport safety policy and practice. Transportmetrica 1, 241–245.
Amoh-Gyimah, R., Saberi, M., Sarvi, M., 2016. Macroscopic modeling of pedestrian and bicycle crashes: a cross-comparison of estimation methods. Accid. Anal. Prev. 93, 147–159.
Anastasopoulos, P.C., Mannering, F.L., 2009. A note on modeling vehicle accident frequencies with random-parameters count models. Accid. Anal. Prev. 41, 153–159.
Anderson, R., 2008. Pedestrian Collisions in South Australia. Centre for Automotive

Safety Research.
Bao, J., Liu, P., Yu, H., Xu, C., 2017. Incorporating twitter-based human activity information in spatial analysis of crashes in urban areas. Accid. Anal. Prev. 106, 358–369.
Bhat, C.R., Koppelman, F.S., 1999.). Activity-based modeling of travel demand. Handbook of Transportation Science. Springer, pp. 35–61.
Bhat, C.R., Singh, S.K., 2000. A comprehensive daily activity-travel generation model system for workers. Transp. Res. Part A-Policy Pract. 34, 1–22.
Bhat, C.R., Goulias, K.G., Pendyala, R.M., Paleti, R., Sidharthan, R., Schmitt, L., Hu, H.-H., 2013. A household-level activity pattern generation model with an application for Southern California. Transportation 40, 1063–1086.
Bowman, J.L., 1998. The Day Activity Schedule Approach to Travel Demand Analysis. Massachusetts Institute of Technology.
Bowman, J.L., Ben-Akiva, M.E., 2001. Activity-based disaggregate travel demand model system with activity schedules. Transp. Res. Part A-Policy Pract. 35, 1–28.
Cai, Q., Lee, J., Eluru, N., Abdel-Aty, M., 2016. Macro-level pedestrian and bicycle crash analysis: incorporating spatial spillover effects in dual state count models. Accid. Anal. Prev. 93, 14–22.
Cai, Q., Abdel-Aty, M., Lee, J., Eluru, N., 2017. Comparative analysis of zonal systems for macro-level crash modeling. J. Safety Res. 61, 157–166.
Castro, M., Paleti, R., Bhat, C.R., 2012. A latent variable representation of count data models to accommodate spatial and temporal dependence: application to predicting crash frequency at intersections. Transp. Res. Part B Methodol. 46, 253–272.
Chliaoutakis, J.E., Darviri, C., Demakakos, P.T., 1999. The impact of young drivers' lifestyle on their road traffic accident risk in greater Athens area. Accid. Anal. Prev. 31, 771–780.
Chliaoutakis, J.E., Koukouli, S., Lajunen, T., Tzamalouka, G., 2005. Lifestyle traits as predictors of driving behaviour in urban areas of Greece. Transp. Res. Part F Traffic Psychol. Behav. 8, 413–428.
Chu, X., 2003. The fatality risk of walking in America: a time-based comparative approach. Walk 21 (IV) Conference Proceedings.
Clarke, D.D., Ward, P., Bartle, C., Truman, W., 2006. Young driver accidents in the UK: the influence of age, experience, and time of day. Accid. Anal. Prev. 38, 871–878.
Cottrill, C.D., Thakuriah, P.V., 2010. Evaluating pedestrian crashes in areas with high low-income or minority populations. Accid. Anal. Prev. 42, 1718–1728.
Devlin, A., Hoareau, E., Logan, D., Corben, B.F., Oxley, J., 2010. Towards zero pedestrian trauma: literature review and serious casualty analysis. Proc. Australas. Road Safety Res. Policing Educ. Conf.
Elias, W., Shiftan, Y., 2014. Analyzing and modeling risk exposure of pedestrian children to involvement in car crashes. Accid. Anal. Prev. 62, 397–405.
Elias, W., Toledo, T., Shiftan, Y., 2010. The effect of daily-activity patterns on crash involvement. Accid. Anal. Prev. 42, 1682–1688.
Ettema, D., Ashiru, O., Polak, J.W., 2004. Modeling timing and duration of activities and trips in response to road-pricing policies. Travel Behav. Values 2004, 1–10.
Fontaine, H., Gourlet, Y., 1997. Fatal pedestrian accidents in France: a typological analysis. Accid. Anal. Prev. 29, 303–312.
Greene-Roesel, R., Diogenes, M.C., Ragland, D.R., 2007. Estimating Pedestrian Accident Exposure: Protocol Report.
Hong Kong Transport Department, 2012a. Hong Kong Annual Tranport Digest: Major Contributory Factors of Road Traffic Accidents 2011-2015. 2016. The Hong Kong SAR Government, Hong Kong.
Hong Kong Transport Department, 2012b. Travel Characteristics Survey 2011. The Hong Kong SAR Government, Hong Kong.
Hong Kong Transport Department, 2016. Road Traffic Casualties by Class of Road User and Degree of Injury 2016. The Hong Kong SAR Government, Hong Kong.
Huang, H., Abdel-Aty, M., Darwiche, A., 2010. County-level crash risk analysis in Florida: bayesian spatial modeling. Transp. Res. Rec.: J. Transp. Res. Board 27–37.
Jiménez-Mejías, E., Martínez-Ruiz, V., Amezcua-Prieto, C., Olmedo-Requena, R., de Dios Luna-del-Castillo, J., Lardelli-Claret, P., 2016. Pedestrian-and driver-related factors associated with the risk of causing collisions involving pedestrians in Spain. Accid. Anal. Prev. 92, 211–218.
Lam, W.W.Y., Loo, B.P., Yao, S., 2013. Towards exposure-based time-space pedestrian crash analysis in facing the challenges of ageing societies in Asia. Asian Geogr. 30, 105–125.
Lam, W.W.Y., Yao, S., Loo, B.P., 2014. Pedestrian exposure measures: a time-space framework. Travel Behav. Soc. 1, 22–30.
Lavieri C.R.B, Patricia S., 2018. Modeling Individuals' Willingness to Share Trips with Strangers in an Autonomous Vehicle Future. Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin.
Lee, J., Abdel-Aty, M., Jiang, X., 2014. Development of zone system for macro-level traffic safety analysis. J. Transp. Geogr. 38, 13–21.
Lee, J., Abdel-Aty, M., Choi, K., Huang, H., 2015a. Multi-level hot zone identification for pedestrian safety. Accid. Anal. Prev. 76, 64–73.
Lee, J., Abdel-Aty, M., Jiang, X., 2015b. Multivariate crash modeling for motor vehicle and non-motorized modes at the macroscopic level. Accid. Anal. Prev. 78, 146–154.
Li, L., Loo, B.P., 2016. Towards people-centered integrated transport: a case study of Shanghai Hongqiao Comprehensive Transport Hub. Cities 58, 50–58.
Liu, L.-L., Yue, W.L., 2011. Pedestrian Accidents in CBD of Adelaide. Eastern Asia Society for Transportation Studies.
Ma, X., Chen, S., Chen, F., 2017. Multivariate space-time modeling of crash frequencies by injury severity levels. Anal. Methods Accid. Res. 15, 29–40.
Mannering, F.L., Shankar, V., Bhat, C.R., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Anal. Methods Accid. Res. 11, 1–16.
Mannings (Asia) Consultants, 2010. Study on Improvements to Pedestrian Links in Mong Kok, Exexutive Summary. In: T. Department (Ed.), .
Mason, K., 1979. Accident patterns by time-of-day and day-of-week of injury occurrence. J. Occup. Accid. 2, 159–176.
Miller, E., Roorda, M., 2003a. Prototype model of household activity-travel scheduling. Transp. Res. Rec.: J. Transp. Res. Board 114–121.
Miller, E.J., Roorda, M.J., 2003b. Prototype model of household activity-travel scheduling. Travel Demand Land Use 2003, 114–121.
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. Accid. Anal. Prev. 43, 1624–1634.
MVA, Hong, Kong, 2013. Traffic Study on Pedestrian Links for the West Kowloon Development Area and Its Connections with Surrounding Districts - Executive Summary.
Naderan, A., Shahi, J., 2010. Aggregate crash prediction models: introducing crash generation concept. Accid. Anal. Prev. 42, 339–346.
Noland, R., Quddus, M., 2004. Analysis of pedestrian and bicycle casualties with regional panel data. Transp. Res. Rec.: J. Transp. Res. Board 28–33.
Pahukula, J., Hernandez, S., Unnikrishnan, A., 2015. A time of day analysis of crashes involving large trucks in urban areas. Accid. Anal. Prev. 75, 155–163.
Palamara, P., Broughton, M., 2013. An investigation of pedestrian crashes at traffic intersections in the Perth Central Business. J. Public Health 93, 1456–1463.
Papadimitriou, E., 2016. Towards an integrated approach of pedestrian behaviour and exposure. Accid. Anal. Prev. 92, 139–152.
Pei, X., Wong, S., Sze, N.-N., 2012. The roles of exposure and speed in road safety analysis. Accid. Anal. Prev. 48, 464–471.
Ponnaluri, R.V., Nagar, F., 2010. Road Crash Risk among Vulnerable Population Groups in Andhra Pradesh, India.
Qin, X., Ivan, J.N., Ravishanker, N., Liu, J., Tepas, D., 2006. Bayesian estimation of hourly exposure functions by crash type and time of day. Accid. Anal. Prev. 38, 1071–1080.
Recker, W.W., 2001. A bridge between travel demand modeling and activity-based travel analysis. Transp. Res. Part B-Methodol. 35, 481–506.
Savolainen, P.T., Mannering, F.L., Lord, D., Quddus, M.A., 2011. The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. Accid. Anal. Prev. 43, 1666–1676.
Siddiqui, C., Abdel-Aty, M., Choi, K., 2012a. Macroscopic spatial analysis of pedestrian and bicycle crashes. Accid. Anal. Prev. 45, 382–391.
Siddiqui, C., Abdel-Aty, M., Huang, H., 2012b. Aggregate nonparametric safety analysis of traffic zones. Accid. Anal. Prev. 45, 317–325 2012/01/25 ed.
Sze, N.N., Christensen, K.M., 2017. Access to urban transportation system for individuals with disabilities. Int. Assoc. Traffic Saf. Sci. (IATSS) Res. 41, 66–73.
Sze, N.N., Wong, S.C., 2007. Diagnostic analysis of the logistic model for pedestrian injury severity in traffic crashes. Accid. Anal. Prev. 39, 1267–1278.
Sze, N.N., Wong, S.C., Lee, C.Y., 2014. The likelihood of achieving quantified road safety targets: a binary logistic regression model for possible factors. Accid. Anal. Prev. 73, 242–251.
Tay, R., 2003. The efficacy of unemployment rate and leading index as predictors of speed and alcohol related crashes in Australia. Int. J. Transp. Econ./Rivista internazionale di economia dei trasporti 363–375.
Tay, R., 2017. Correlation, variance inflation and multicollinearity in regression model. J. East. Asia Soc. Transp. Stud. 12, 2006–2015.
Train, K., 2001. Halton sequences for mixed logit. Econometrics.
Train, K., 2009. Discrete Choice Methods with Simulation. Cambridge University Press.
Varhelyi, A., 1997. Dynamic Speed Adaptation Based on Information Technology-a Theoretical Background, vol. 142 Bulletin/University of Lund, Lund Institute of Technology, Department of Traffic Planning and Engineering.
Wang, Y., Kockelman, K.M., 2013. A Poisson-lognormal conditional-autoregressive model for multivariate spatial analysis of pedestrian crash counts across neighborhoods. Accid. Anal. Prev. 60, 71–84.
Wang, X., Yang, J., Lee, C., Ji, Z., You, S., 2016. Macro-level safety analysis of pedestrian crashes in Shanghai, China. Accid. Anal. Prev. 96, 12–21 2016/08/01 ed.
Wang, J., Huang, H., Zeng, Q., 2017. The effect of zonal factors in estimating crash risks by transportation modes: motor vehicle, bicycle and pedestrian. Accid. Anal. Prev. 98, 223–231.
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. Accid. Anal. Prev. 41, 137–145.
Wong, S.C., Sze, N.N., Yip, H.F., Loo, B.P.Y., Hung, W.T., Lo, H.K., 2006. Association between setting quantified road safety targets and road fatality reduction. Accid. Anal. Prev. 38, 997–1005.
Wong, S.C., Sze, N.N., Li, Y.C., 2007. Contributory factors to traffic crashes at signalized intersections in Hong Kong. Accid. Anal. Prev. 39, 1107–1113.
Wong, S.C., Wong, C.W., Sze, N.N., 2008. Attitudes of public light bus drivers to penalties to combat red light violations in Hong Kong. Transp. Policy 15, 43–54.
World Health Organization, W., 2015. Global Status Report on Road Safety 2015. World Health Organization.
Xiong, Y.G., Mannering, F.L., 2013. The heterogeneous effects of guardian supervision on adolescent driver-injury seventies: a finite-mixture random-parameters approach. Transp. Res. Part B-Methodol. 49, 39–54.
Zegeer, C.V., Bushell, M., 2012. Pedestrian crash trends and potential countermeasures from around the world. Accid. Anal. Prev. 44, 3–11.
Zou, W., Wang, X., Zhang, D., 2017. Truck crash severity in New York city: an investigation of the spatial and the time of day effects. Accid. Anal. Prev. 99, 249–261.