

The effect of road network patterns on pedestrian safety:

A Zone-based Bayesian spatial modeling approach

Qiang GUO¹, Pengpeng XU², Xin PEI^{3*}, S.C. WONG⁴, Danya YAO⁵

¹Department of Automation, Tsinghua University, Beijing, 100084, China; email: guoqiang_620@126.com

²Department of Civil Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong; email: pengpengxu@yeah.net

^{3*}Corresponding Author, Department of Automation, Tsinghua University, Beijing, 100084, China; email: peixin@mail.tsinghua.edu.cn

⁴Department of Civil Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong; email: hhecwsc@hku.hk

⁵Department of Automation, Tsinghua University, Beijing, 100084, China; email: yaody@mail.tsinghua.edu.cn

ABSTRACT

Pedestrian safety is increasingly recognized as a major public health concern. Extensive safety studies have been conducted to examine the influence of multiple variables on the occurrence of pedestrian–vehicle crashes. However, the explicit relationship between pedestrian safety and road network characteristics remains unknown. This studies particularly focused on the role of different road network patterns on the occurrence of crashes involving pedestrians. A global integration index via space syntax was introduced to quantify the topological structures of road networks. The Bayesian Poisson-lognormal (PLN) models with conditional autoregressive (CAR) prior were then developed via three different proximity structures: contiguity, geometry-centroid distance, and road network connectivity. The models were also compared with the PLN counterpart without spatial correlation effects. The analysis was based on a comprehensive crash dataset from 131 selected traffic analysis zones in Hong Kong.

The results indicated that higher global integration was associated with more pedestrian–vehicle crashes; the irregular pattern network was proved to be safest in terms of pedestrian crashes occurrences, whereas the grid pattern was the least safe; the CAR model with a neighborhood structure based on road network connectivity was found to outperform in model goodness-of-fit, implying the importance of accurately accounting for spatial correlation when modeling spatially aggregated crash data.

Key words: pedestrian safety, road network patterns, global integration, road network connectivity, zone-based approach, Bayesian spatial model

1. Introduction

As a means of primary transportation, walking holds promise for not only supporting public health, but also reducing resource and land consumption, air and noise pollution, and greenhouse gas emissions (Gomez-Ibanez et al., 2009; Ewing et al., 2011). However, walking is also associated with negative public health outcomes related with pedestrian–vehicle crashes. Approximately 400,000 pedestrians are killed in pedestrian–vehicle crashes each year worldwide (Naci et al., 2009), whereas in Hong Kong, pedestrians accounted for more than half of the road traffic fatalities (HKTD, 2014), a proportion that was much higher than that in other high-income countries (e.g., the United States, the United Kingdom, and Japan). Hence, effective road safety strategies must be formulated and implemented to improve the safety of these vulnerable road users (Wong et al., 2002; Loo et al., 2005, 2007).

Over the past decade, numerous studies have been conducted to identify the effects of potential contributing factors, such as pedestrian and driver characteristics (e.g., age and gender), road geometric features (e.g., road width, horizontal and vertical alignment), traffic characteristics (e.g., traffic volume and travel speed), neighborhood environmental conditions (e.g., urban form, land use, and access to public transport) and other factors (e.g., weather, temporal, and lighting), on the occurrence of crashes involving pedestrians (Gårder, 2004; Lee and Abdel-Aty, 2005; Kim et al., 2008; Wier et al., 2009; Ewing and Dumbaugh, 2009; Moudon et al., 2011; Siddiqui et al., 2012, 2014; Lam et al., 2014; Yao et al., 2015; Wang et al., 2016).

In addition to the preceding contributing factors, road network structures have been identified as important in determining traffic safety, especially for pedestrians (Marshall and Garrick, 2010b; Rifaat et al., 2011, 2012; Wei and Lovegrove, 2012; Dumbaugh and Zhang, 2013; Zhang et al., 2015). Urban road networks not only influence quality of life, but also promote economic development and regional growth (Florida, 2002) and improve people’s physical fitness and health (Frank et al., 2003). According to Appleyard (1980), an ideal road network should possess several characteristics (such as located in a livable environment, located in neighborly territory, and being green and pleasant) in addition to acting as thoroughfares for carrying people and goods from one place to another. A reduction in pedestrian crashes is one of the road and traffic-related issues that has an effect on neighborhood life. The structure of road networks influences pedestrian safety through its effects on the fundamental characteristics of traffic volumes and journey patterns for both vehicles and pedestrians (Zhang et al., 2015). Road network patterns also affect travel activities, vehicle speeds, and driver behavior, thus influencing road safety (Quddus, 2008).

Previous studies have considered road network structures in different ways. One common way of classifying road network patterns is based on the concept of macro- and micro-level road networks proposed by Marshall (2004). Macro-level or citywide road networks distinguish roads that are typically continuous over a substantial part of the city and enable travel from one part of the city to another. In contrast, micro-level or neighborhood road networks mainly serve residential neighborhood travel, as these roads are on routes that are not continuous over a significant part of the city. Following Marshall and Garrick (2010a, 2010b, 2011), four types of citywide road network (linear, tributary, radial, and grid) and two types of neighborhood road network (tree and grid) are defined and combined to describe hierarchical road network patterns. An alternative method involves concentrating directly on the overall road network patterns of a community rather than the different types of road network pattern and then combining the different types to compose a new type. For

1 instance, Lovegrove and Sun (2010, 2013) classified road network patterns into five categories: fused
2 grid, commonly used grid, cul-de-sac, three-way offset, and Dutch sustainable road safety (SRS).
3 They found that fused grid and three-way offset patterns were safer than grid, cul-de-sac, and Dutch
4 SRS patterns. Southworth and Ben-Joseph (2003) classified road network patterns into five other
5 categories: gridiron, fragmented parallel, warped parallel, loops and lollipops, and lollipops on a
6 stick. Rifaat et al. (2010, 2011, 2012) further merged fragmented parallel patterns and lollipops into a
7 mixed pattern and concluded that relative to limited access patterns, the grid pattern was associated
8 with more crashes and lower pedestrian crash severity.

9 The main limitation of the preceding studies is that their road network classification methods
10 were largely based on the visual inspections and hence may be prone to subjectivity and uncertainty
11 due to the unavailability of quantitative road network indices. Moreover, these manual classification
12 approaches are time-consuming, limiting their generalization to large study regions. To overcome
13 these problems, some researchers have attempted to extract the structures of road networks
14 quantitatively based on topological measurements (Wang et al., 2012, 2013; Zhang et al., 2012,
15 2015).

16 This study introduced an alternative innovative road network modeling technique known as
17 space syntax (Hillier and Hanson, 1984; Hiller, 1996) to quantify the characteristics of different road
18 network patterns. The notion of syntax refers to relationships between different spaces, or
19 interactions between spaces and society. It recognizes that spatial patterns or structures greatly
20 influence human activities and behavior in urban environments. Typical applications of space syntax
21 include pedestrian and bicyclist modeling, criminal mapping, pollution analysis, and retail viability
22 modeling (Jiang, 1999; Ratti, 2004). Space syntax can objectively measure the effects of road
23 network patterns on accessibility (Raford et al., 2007). This study therefore used space syntax to
24 quantify the structures of road network patterns.

25 To determine the effects of road network patterns on pedestrian safety, aggregate relationships
26 between the aforementioned independent variables and pedestrian crash counts can be established
27 using macro-level or area-wide safety performance functions (SPFs). Traditional Poisson and
28 Poisson-lognormal (PLN) models ignore the issue of spatial correlation, which may lead to biased
29 parameters and incorrect inferences (Huang and Abdel-Aty, 2010). To address the spatial dependence
30 between adjacent zones, Bayesian spatial models with a conditional autoregressive (CAR) prior have
31 been widely applied in current road safety analysis (Quddus, 2008; Mitra, 2009; Aguero-Valverde
32 and Jovanis, 2008, 2010; Siddiqui et al., 2012; Wang et al., 2012, 2013; Xu et al., 2014; Xu and
33 Huang, 2015).

34 Within the framework of CAR, the spatial relationship is typically linked by virtue of a spatial
35 weight matrix. One commonly used weighting scheme includes adjacency-based first-order
36 neighbors (Miaou et al., 2003; Aguero-Valverde and Jovanis, 2008; Quddus, 2008), which can be
37 defined as all of the observation units that connect directly with the one in question. Additional
38 consideration is given to distance decay-based forms (Aguero-Valverde and Jovanis, 2010; Dong et
39 al., 2014; Wang et al., 2016). However, the pure geographical proximity may be inadequate to
40 express the spatial relationships under investigation. In reality, a more rational spatial weighting
41 configuration is required to simultaneously consider the geographical dimension and other aspects
42 (e.g., travel frequency, travel modes, and travel demand) that vary with the context of study. For this
43 purpose, this study proposed a new neighborhood structure using road network connectivity to reflect
44 the potential traffic interactions between adjacent zones.

Overall, this study described an innovative street pattern modeling technique known as space syntax to objectively quantify the structure of road networks. More specifically, the primary objective of this study was to explore how road network characteristics—in terms of global integration—are associated with zone-level pedestrian safety using both non-spatial (e.g., PLN model) and spatial (e.g., CAR model) SPFs. Another objective was to compare different spatial proximity structures at the traffic analysis zone (TAZ) level. Bayesian spatial models with a CAR prior were used with three different types of neighborhood structure: contiguity, geometry-centroid distance, and road network connectivity.

The remainder of this paper is structured as follows. Section 2 describes the data collection process. Section 3 presents the structures of non-spatial and spatial SPFs. Section 4 summarizes the results and discusses the investigation. Section 5 presents the conclusions and suggestions for the future direction of research.

2. Data

A comprehensive database relating to pedestrian–vehicle crashes and traffic-related variables for Hong Kong was established using geographical information system (GIS) techniques.

The spatial units for analysis were the TAZs from the Hong Kong Planning Vision and Strategy (PVS) zoning system, which is managed by the Hong Kong Planning Department. The PVS zoning system divides the whole of Hong Kong into different sectors and is revised every few years to better describe the zoning boundary. The PVS 338 zoning system uses a 2001-based Territorial Population and Employment Data Matrix (TPEDM). There are 338 TAZs in this zoning system. Of these, 131 TAZs (38.76%) were finally selected as having adequate traffic and geometric information. To examine the role of road network patterns and zone-related variables that contribute to pedestrian safety at a more disaggregated level, the database was organized into 4-hour periods: 07:00–11:00 (morning), 11:00–15:00 (noon), 15:00–19:00 (afternoon), 19:00–23:00 (evening), 23:00–03:00 (middle of the night), and 03:00–07:00 (dawn). Our database thus comprised 786 observations.

2.1 Measurements of road networks

Various morphological parameters of 131 TAZs in Hong Kong were derived by space syntax, including connectivity, depth, and integration. Connectivity refers to the number of lines that directly intersect a given axial line. Depth is defined as the distance in number of lines from a number of steps to the given axial line. Connectivity focuses on immediate neighborhoods, and depth focuses on k neighborhoods. Integration refers to the accessibility of a whole road network, which is derived from the mean depth.

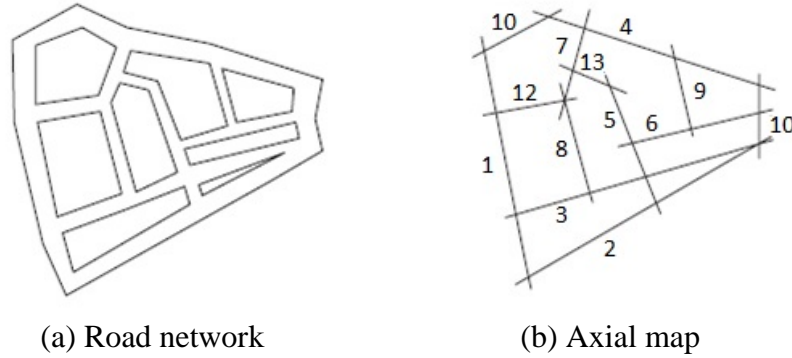
Based on the axial map, which consists of the smallest number of longest and intersected axial lines needed to cover the whole free space and to represent the road network patterns, several variables are defined and estimated accordingly.

$$D_i = \sum_{d=1}^s d \times N_d = \begin{cases} \text{connectivity} & s=1 \\ \text{local depth} & s=k, \\ \text{global depth} & s=l \end{cases} \quad (1)$$

where d , an integer, represents the shortest distance far from i th node, N_d denotes the number of

1 nodes with the shortest distance d , k is the number of neighborhoods, and l denotes the
 2 maximum shortest distance. Furthermore, it should be noted that $1 < k < l$, where k is usually equal
 3 to 3 for the calculation of local depth, and the connectivity is equivalent to local depth if $k = 1$.

4 To introduce the principles of this algorithm, Fig. 1a represents a simplified road network
 5 structure. Its axial map is derived by drawing line 1, then 2, then 3 ... and the final axial map consists
 6 of 13 axial lines as shown in Fig. 1b.



7
 8 (a) Road network (b) Axial map
 9 Fig. 1. A simplified road network and its axial map

10 Consider axial line 1 in Fig. 1b. The line directly intersects four lines, so the connectivity value
 11 of this axial line is equal to 4. The immediate neighborhoods of line 1 are lines 2, 3, 10, and 12, and
 12 their respective connectivity values are 4, 5, 2, and 3. Axial line 1 has four neighborhoods one step
 13 away, five neighborhoods two steps away, and three neighborhoods three steps away. The global
 14 depth of this axial line is thus equal to $1 \cdot 4 + 2 \cdot 5 + 3 \cdot 3$ if $s=3$, and the local depth with two steps away
 15 is equal to $1 \cdot 4 + 2 \cdot 5$ if $s=2$.

16 MD_i , the mean depth, is defined as the average distance from the i th node to other nodes. The
 17 formula for calculating mean depth is defined as

$$18 \quad MD_i = D_i / (n-1) = \sum_{d=1}^s d \times N_d / (n-1) = \begin{cases} \text{local mean depth} & s = k \\ \text{global mean depth} & s = l \end{cases} \quad (2)$$

19 Integration measures how easily accessible the i th node is from other nodes in the system. The
 20 integration value is calculated by

$$21 \quad I_i = \frac{D_n}{2(MD_i - 1) / (n-2)}, \quad (3)$$

22 where n is the number of nodes within the system and

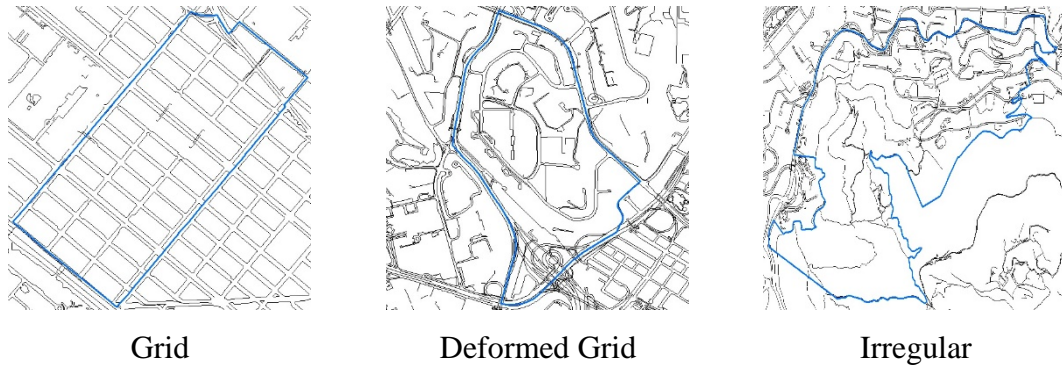
$$23 \quad D_n = 2 \left\{ n \left[\log_2 \left(\frac{n+2}{3} \right) - 1 \right] + 1 \right\} / (n^2 - 3n + 2). \quad (4)$$

24 Integration is a reciprocal function of MD_i , i.e., the integration value is high when the mean
 25 depth is low and vice versa. Depending on the mean depth used, i.e., local versus global, the
 26 corresponding integration value is called local versus global integration.

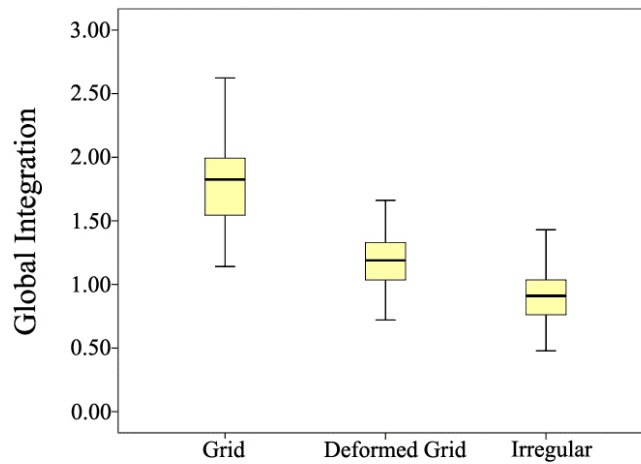
27 Continuing with this example, the global mean depth of axial line 1 is 1.9167 and the D-value is
 28 0.2755, so the global integration of this axial line is equal to 1.6529.

29 In this study, connectivity, local integration, and global integration were calculated for each of
 30 the inspected 131 TAZs. Both connectivity and local integration are local parameters, as they

1 consider immediate neighboring lines or lines with few steps. In comparison, global integration is a
 2 global parameter and considers both immediate and non-immediate neighborhoods up to a maximum
 3 number of steps away. To determine whether global integration is capable of quantitatively
 4 distinguishing different road network patterns, the network patterns of all 131 TAZs were first
 5 classified as grid, deformed grid, and irregular (Liu, 2007). These patterns are shown in Fig. 2. The
 6 values of global integration for each of these three network patterns were then calculated. The
 7 corresponding distribution is shown in Fig. 3. As illustrated, the global integration changed
 8 systematically with different road network patterns and was an appropriate measurement for
 9 describing any type of road network pattern.



10 Fig. 2. Typical network patterns of Hong Kong

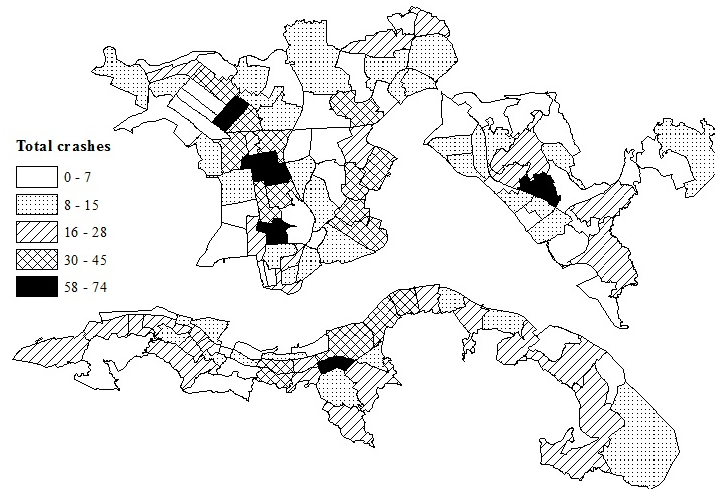


11 Fig. 3. Comparison of the global integration of different network patterns

12 *2.2 Crash data*

13 The pedestrian–vehicle crash data used in this study were obtained from the traffic information
 14 system (TIS) maintained by the Hong Kong Police Force and Transport Department. The TIS
 15 consists of three types of statistical information: crash environment profile, vehicle involvement
 16 profile, and casualty injury profile (Wong et al., 2007; Huang et al., 2016). In the TIS, all crashes are
 17 categorized according to crash severity level: fatal, serious, and slight. A fatal crash refers to a crash
 18 in which at least one person is killed or is injured and subsequently dies within 30 days of the crash.
 19 A serious crash refers to a crash in which one or more persons are injured and detained in hospital for
 20 more than 12 hours, and a slight crash is one in which one or more persons are injured, but not to the
 21 extent that a hospital stay of more than 12 hours is required. In this study, fatal and serious crashes

1 were classified together as killed and seriously injured (KSI) crashes (Pei et al., 2011). Fig. 4 shows
2 the spatial distribution of aggregated crashes that occurred in the 131 TAZs of Hong Kong during
3 2011. Of the 2,168 pedestrian–vehicle crashes, 462 (21.31%) were KSI crashes.



4 Fig. 4. Spatial distribution of pedestrian crashes in the 131 TAZs of Hong Kong

5 2.3 Speed data

6 Travel speed is believed to be a determinant factor of pedestrian crash occurrence. In this study,
7 speed data from global positioning system (GPS) probe vehicles were used. GPS probes are present
8 in 480 taxis traveling on the roadway network in Hong Kong. Instantaneous information on the
9 location, speed, and travel direction of each GPS-equipped taxi is transmitted to and stored in a
10 control center at 30-second intervals. Pei et al. (2009) adopted the journey time survey data
11 maintained by the Department of Transportation to explore the validity of the GPS probe data and
12 found that taxi speed was a reasonable proxy for actual travel speed. The average travel speed for
13 each TAZ of interest was calculated for every 4-hour period using these raw data.

14 2.4 Exposure

15 Two exposure indicators—vehicle hours (VH) and pedestrian activities (PA)—were considered
16 in the proposed models. As suggested by Pei et al. (2012), the time exposure measured by VH is
17 more reasonable as an exposure index. To estimate the variable, the average annual daily traffic
18 (AADT) data and its temporal (in relation to month, day of the week, and hour of the day) and
19 directional multiplicative factors were obtained from the Hong Kong annual traffic census (ATC)
20 system, which is maintained by the Hong Kong Transport Department (Lam et al., 2003; Tong et al.,
21 2003). The ATC system provides the hourly traffic volume for approximately 100 core stations in
22 Hong Kong. According to Wong and Wong’s (2015, 2016) studies, a linear data projection method
23 was adopted to combine the AADT and taxi GPS dataset. First, the scaling factors for each ATC core
24 station were calculated. Second, the scaling factors for each TAZ were predicted according to the
25 spatial correlation of the related core stations. Consequently, the hourly traffic spots within each TAZ
26 could thus be estimated, and time exposure could further be estimated in relation to the last time for
27 each spot, which was supposed to be equal to the update time of the taxi GPS data.

28 To calculate another exposure indicator, pedestrian trip data were first extracted from the
29 trip-based data file (named C501A) of the Travel Characteristics Survey 2002 (TCS 2002). This

1 survey is conducted by the Hong Kong Transport Department about once a decade to understand the
2 normal weekday (excluding public holidays) trip information and other information, such as the
3 household and personal details of Hong Kong residents. The C501A file contains all trip records
4 reported by the respondents. Two sub-files were extracted from C501A for separate analysis: one
5 containing the data for walk-only trips, and the other containing the data for trips involving any of
6 the 10 major mechanized transport modes. In this study, analysis of pedestrian trips data was based
7 on walk-only trips. The PA of each inspected TAZ consisted of three parts: trip generations, trip
8 attractions, and internal pedestrian activities.

9 To explore the effect of exposure on safety performance, logarithmically transformed VH and PA
10 were incorporated into the models as explanatory factors.

11 *2.5 Other explanatory factors*

12 For a better estimation of the effects on pedestrian safety, the confounding factors of roadway
13 geometric characteristics, demographics, land use patterns, and temporal distribution were controlled
14 in the developed models.

15 Roadway geometric characteristics are commonly considered as contributing factors to the
16 occurrence of crashes involving pedestrians at the TAZ level. Various macro-level geometric
17 attributes (Noland and Quddus, 2004; Wier et al., 2009; Siddiqui et al., 2012), including the number
18 of intersections, total road length, road density, and intersection density, are considered as possible
19 descriptive factors. Road density is equal to road length divided by zone area per TAZ. Intersection
20 density is defined as the number of intersections divided by road length per TAZ. In this study, road
21 density and intersection density were included in the models.

22 Demographic factors were extracted from TPEDM and included different kinds of estimated
23 values (such as the resident population, the age cohort of residents with regular residence, and the
24 number of domestic households) in the base year listed for each TAZ. In this study, the percentage of
25 the population aged 65 years and older was considered in the models.

26 Land use patterns were extracted from the outline zoning plan (OZP) maintained by the Hong
27 Kong Town Planning Board. The OZP uses colored hatching on the base map to represent and
28 identify different permitted land uses, such as government/institution/community, commercial, and
29 residential. In this study, the land use pattern of each selected TAZ was estimated by the percentage
30 of trips in each area according to the trip pattern and purpose. The land use patterns were divided into
31 four different categories—residential, commercial, mixed (combined commercial and residential use),
32 and others—by comparing the dominant trip purpose (more than 50%) and incorporated into the
33 models as categorical variables, with residential as the control.

34 The effect of temporal distribution on pedestrian–vehicle crash occurrences by time of day was
35 also controlled to neutralize the confounding effect of unidentified factors. For example, the six
36 periods were included in the regression models as categorical variables with the dawn period (03:00–
37 07:00) as the control.

38 *2.6 Summary statistics of variables*

39 Table 1 shows the descriptive statistics for the variables included in the final models. To ensure
40 that all of the contributory factors included were statistically independent, the multicollinearity
41 between them was calculated. One of the quantified factors, the variance inflation factor (VIF), was
42 used to identify the severity of the multicollinearity in the regression model. The VIF values of all of

1 the independent variables were less than 5, showing no evidence of a potentially serious collinearity
 2 problem or a correlation problem affecting the regression.

3

4 Table 1 Summary Statistics of Crash Occurrences and Contributory Variables

	Min	Max	Mean	S.D.	Count	Proportion
Number of observations = 786						
Total crashes	0	30	2.758	3.704		
KSI crashes	0	6	0.588	0.995		
Slight crashes	0	26	2.171	3.078		
Average speed (km/h)	7.758	77.972	26.820	12.622		
Global integration	0.479	3.624	1.146	0.428		
Road density (km/km ²)	21.198	109.392	50.074	14.962		
Intersection density (Number of intersections per km)	1.210	11.240	3.559	1.474		
Senior percentage (older than 65)	0.000	0.232	0.131	0.040		
Ln(PA)	0.000	6.238	3.153	1.819		
Ln(VH)	5.371	13.516	11.044	1.250		
07:00–11:00					131	16.7%
11:00–15:00					131	16.7%
15:00–19:00					131	16.7%
19:00–23:00					131	16.7%
23:00–03:00					131	16.7%
03:00–07:00					131	16.7%
Residential					462	58.7%
Commercial					150	19.1%
Mixed					102	13.0%
Others					72	9.2%

5 3. Methodology

6 For the purpose of comparison, two types of SPFs—non-spatial and spatial models—were
 7 established and estimated using the Markov chain Monte Carlo (MCMC) approach and full Bayesian
 8 method (Gelman et al., 2014), respectively. Both of these modeling techniques are briefly presented
 9 as follows.

10 3.1 Non-spatial models

11 Crash data are nonlinear, discrete, and non-negative data that approximately follow the Poisson
 12 distribution (Lord and Mannering, 2010). One underlying assumption of the Poisson model is that
 13 the mean of the data is restricted to be equal to the variance, i.e., $E(y_i) = Var(y_i)$. If this equality
 14 does not hold, the data are said to be under-dispersed ($E(y_i) < Var(y_i)$) or over-dispersed
 15 ($E(y_i) > Var(y_i)$), and the coefficient's vector is biased if that is not taken into account. In most
 16 practical cases, the variance of pedestrian crash data is greater than the mean, so the data are
 17 considered to be over-dispersed (Noland and Quddus, 2004). Over-dispersion can be resolved by

1 assuming that the crash data are distributed in a negative binomial (NB) or PLN distribution. To test
 2 the appropriateness of using the traditional count models (e.g., NB) rather than zero-inflated models
 3 (e.g., zero-inflated negative binomial (ZINB)), the Vuong test was conducted to compare the
 4 predictive probability of two candidate models (Vuong, 1989). Finally, the Vuong statistic for testing
 5 the ZINB versus the normal-count NB was computed as -3.421, which was less than -1.96 and
 6 favored the traditional NB model. Although the NB model has been widely applied in road safety
 7 analysis (Washington et al., 2011), Lord and Miranda-Moreno (2008) recently found that the PLN
 8 model could be a better alternative in cases of small sample size and low sample mean values.

9 Consequently, the PLN distribution is preferred over other mixed-Poisson distributions by
 10 including the lognormal error distribution in the Poisson distribution. The PLN model (Model 1) is
 11 specified as follows:

$$12 \quad y_i | \lambda_i \sim \text{Poisson}(\lambda_i), \quad (5)$$

$$13 \quad \lambda_i = \exp\left(\beta_0 + \beta_1 \ln(\text{VH}_i) + \beta_2 \ln(\text{PA}_i) + \sum_{k=3}^p \beta_k X_{ik} + \theta_i\right), \quad (6)$$

$$14 \quad \theta_i \sim \text{Normal}(0, \tau_\theta^2), \quad (7)$$

15 where y_i is the number of crashes for TAZ_i per period, λ_i is the expected number of crashes
 16 per period, β_0 is the intercept term, $\beta = (\beta_1, \beta_2, \dots, \beta_p)'$ is a regression coefficient vector, VH_i
 17 and PA_i are the exposure variables of TAZ_i , X_i is a vector of explanatory variables, θ_i is the
 18 unobserved (i.e., uncorrelated) heterogeneity error component of the model, and τ_θ^2 is the precision
 19 (inverse of the variance) parameter with a prior gamma distribution $\text{Ga}(0.5, 0.0005)$, as suggested
 20 by Wakefield et al. (2000) and used by Aguero-Valverde and Jovanis (2008, 2010).

21 3.2 Spatial models

22 The PLN model is able to capture unstructured over-dispersion or unobserved heterogeneity
 23 caused by omitted variables. However, this model ignores the issue of spatial dependence among
 24 spatial TAZs, which violates the traditional Gauss-Markov assumption involved in regression
 25 modeling (Huang and Abdel-Aty, 2010). To account for the heterogeneity that occurs due to spatial
 26 correlation, a second random effect component ϕ_i is added to the model, and Equation 6 can be
 27 rewritten as

$$28 \quad \lambda_i = \exp\left(\beta_0 + \beta_1 \ln(\text{VH}_i) + \beta_2 \ln(\text{PA}_i) + \sum_{k=3}^p \beta_k X_{ik} + \theta_i + \phi_i\right), \quad (8)$$

29 where ϕ_i is the spatial correlation error component.

1 The effect of spatial correlation, ϕ_i , is assigned as intrinsic CAR with a distribution $N(\bar{\phi}_i, \tau_i^2)$
 2 as recommended by Besag et al. (1991), where

$$3 \quad \bar{\phi}_i = \frac{\sum_{i \neq j} \omega_{ij} \phi_j}{\sum_{i \neq j} \omega_{ij}} \quad \text{and} \quad \tau_i^2 = \frac{\tau_\phi^2}{\sum_{i \neq j} \omega_{ij}}, \quad (9)$$

4 in which ω_{ij} is the adjacency index for TAZ_i and TAZ_j in the spatial weight matrix W and
 5 τ_ϕ^2 is also assumed a gamma prior with $\text{Ga}(0.5, 0.0005)$ as suggested by Wakefield et al. (2000).

6 To determine the contribution of spatial correlation, the proportion of variability in the random
 7 effect components is calculated as

$$8 \quad \alpha = \frac{\text{SD}(\phi)}{\text{SD}(\phi) + \text{SD}(\theta)}. \quad (10)$$

9 where SD is the empirical marginal standard deviation function.

10 As there is no prior information available, a non-information prior distribution for β is
 11 specified as Normal(0, 100,000). A uniform prior distribution is assumed for β_0 .

12 3.3 Proximity matrix

13 In the CAR models, the specification of spatial weight matrix W is a major concern because
 14 the choice of proximity matrix may have a substantive effect on the results. In the present study,
 15 three different forms of neighborhood structure constructed by contiguity (Model 2a),
 16 geometry-centroid distance (Model 2b), and road network connectivity (Model 2c) were used.

17 *Model 2a proximity structure based on contiguity*

18 The simplest form for W is a “0/1” contiguity matrix, wherein $\omega_{ij} = 1$ if TAZ_i and TAZ_j
 19 are spatially adjacent and 0 otherwise.

20 *Model 2b proximity structure based on geometry-centroid distance*

21 A more general setup uses distance to reflect the correlation of TAZs. An alternative form of W ,
 22 defining ω_{ij} as a function of the distance between TAZs, is

$$23 \quad \omega_{ij} = C(d_{ij}), \quad (11)$$

24 where d_{ij} is the geometry-centroid distance between TAZ_i and TAZ_j, and $C(d_{ij})$ is usually a
 25 decreasing function of d_{ij} so that the TAZs close to each other have a large covariance. In this study,
 26 the exponential function $C(d_{ij}) = \exp(-d_{ij} / \omega)$, where ω is a spatial decay parameter, was used to

1 calculate the spatial weight matrix. According to a previous study by Carlin and Ma (2007), ω was
2 set to the average distance among TAZs in this study.

3 *Model 2c proximity structure based on road network connectivity*

4 In this study, another alternative approach to defining the spatial weights matrix W was based
5 on road network connectivity. The general expression for W is provided by

$$6 \quad \omega_{ij} = Q(e_{ij}) * C(d_{ij}), \quad (12)$$

7 where e_{ij} represents the links connecting TAZ_{*i*} and TAZ_{*j*}. The links are defined as the road

8 segments between two TAZs. $Q(e_{ij})$ is the traffic volume of the links. The total number of road

9 segments between two TAZs is calculated as a quantitative index to represent the traffic volume of

10 links. $C(d_{ij})$ is a negative exponential function of the geometry-centroid distance d_{ij} , i.e.,

11 $C(d_{ij}) = \exp(-d_{ij} / \omega)$. ω was similarly set to the average distance among TAZs.

12 *3.4 Model comparison*

13 To compare the different models, the deviance information criterion (DIC), an extension of the
14 Akaike information criterion (AIC) proposed by Spiegelhalter et al. (2002), was used. The DIC

15 provides a Bayesian measure of model complexity and fitting. Based on the deviance statistic $D(\theta)$,

16 the criterion DIC is defined as

$$17 \quad \text{DIC} = \overline{D(\theta)} + p_D, \quad (13)$$

18 where $\overline{D(\theta)}$ is the posterior expectation deviance that can be a Bayesian measure of fitting, i.e.,

19 $\overline{D(\theta)} = E(D(\theta) | y)$. p_D is the effective number of parameters in relation to a model and is defined

20 as

$$21 \quad p_D = E(D(\theta) | y) - D(E(D(\theta) | y)). \quad (14)$$

22 Similar to the AIC, models with lower DIC values are preferred.

23 **4. Results and discussion**

24 Depending on the complexity of the models, three MCMC chains from different starting points
25 were estimated to ensure convergence. The first 5,000 iterations in each chain were discarded as

26 burn-ins. A further 5,000 iterations were then performed for each chain, resulting in a sample

27 distribution of 15,000 for each parameter. Convergence of the models was monitored by visual

28 examination of the MCMC chains, autocorrelation plots, and Gelman-Rubin statistic plots.

1 4.1 Model comparison

2 The model performance results are shown in Table 2. The model performance results are shown
 3 in Table 2. The posterior means for the standard deviations of spatial correlation ϕ for all CAR
 4 models are statistically significant, suggesting that the observed pedestrian-vehicle crashes are
 5 spatially correlated among the TAZs. Similar to ϕ , the posterior means for the standard deviations
 6 of uncorrelated heterogeneity θ are also statistically significant.

7 As shown by the DIC values, the CAR models substantially outperform their PLN counterpart.
 8 This result is not surprising to some extent, as the CAR models not only account for the unstructured
 9 over-dispersion, but also take the spatial correlation between adjacent TAZs into account.

10 By comparing the DIC values of CAR models (DIC = 2,576.49 for Model 2a, 2,564.74 for
 11 Model 2b, and 2,559.17 for Model 2c), it can be seen that the consideration of proximity structure
 12 based on road network connectivity among adjacent TAZs is more reasonable. This result is expected,
 13 as both Model 2a and Model 2b only consider contiguity or distance relationships. In contrast, Model
 14 2c is a quantitative measurement of transport infrastructures, which actually concentrates on
 15 transport activities (e.g., vehicular trips and pedestrian trips). For example, the relationships between
 16 entities that are geographically connected to neighbors are likely to be loose, as neighboring entities
 17 have low road network connectivity, resulting in a total number of vehicular and pedestrian trip
 18 productions and attractions that is small or very close to zero.

19 In summary, it is essential to identify and explore the spatial correlation of pedestrian-vehicle
 20 crashes. Moreover, the CAR model with a neighborhood structure based on road network
 21 connectivity can properly account for the spatial heterogeneity between the observations.

22

23 Table 2 Model Comparison

Model	PLN model		CAR models	
	Model 1	Model 2a	Model 2b	Model 2c
S.D. of θ	0.364 (0.297, 0.434)*	0.132 (0.027, 0.260)	0.127 (0.028, 0.253)	0.114 (0.025, 0.228)
S.D. of ϕ	-	0.415 (0.344, 0.495)	0.389 (0.325, 0.456)	0.394 (0.328, 0.463)
α	-	0.769 (0.602, 0.938)	0.764 (0.598, 0.935)	0.783 (0.625, 0.941)
DIC	2,609.870	2,576.490	2,564.740	2,559.170

24 * Estimated mean (95% Bayesian credible interval) for the parameter

25 4.2 Parameter estimation

26 The estimation results of Model 2c are presented in Table 3. The estimation results of Model 2c
 27 are presented in Table 3. The main objective of this study was to explore the effect of road network
 28 structures on the occurrence of pedestrian-vehicle crashes. For this purpose, a global integration
 29 index with space syntax was derived and used. As shown in Table 3, an increase in global integration
 30 increases the likelihood of a pedestrian-vehicle crash occurring at the 5% significance level (mean =
 31 0.255; 95% CI = [0.090, 0.424]). A road network with higher global integration undoubtedly
 32 indicates roads that provide greater accessibility to a large number of other roads acting as conduits
 33 for more origin/destination pairs. These networks consequently experience more use than do less
 34 connected and less accessible networks. A higher global integration also indicates that a road within a
 35 network is likely to include a greater number of connections (intersections) to reach other road. As

1 the number of intersections increases, so does the probability of complex vehicular maneuvers. In
 2 addition, a network with higher global integration tends to have more complicated routes, which
 3 increases the potential for pedestrian–vehicle conflicts. To further illustrate the relationship between
 4 road networks and global integration, the values of global integration for each TAZ were calculated.
 5 According to the three studied network patterns, nine samples from one hundred and thirty-one TAZs
 6 were chosen. Each category contained three samples. The global integration of the road network for
 7 each sample zone is shown in Table 4.










8
 9 Table 3 Modeling Results of Model 2c

Variable	Posterior Mean	Posterior S.D.	(95% CI)
Constant	-3.186*	0.364	(-3.920,-2.475)
Global integration	0.255*	0.087	(0.090, 0.424)
Road density	0.002	0.002	(-0.002, 0.007)
Intersection density	0.088*	0.028	(0.034, 0.142)
Average speed	-0.021*	0.003	(-0.028, -0.015)
Senior percentage	0.568	1.045	(-1.474, 2.550)
Commercial	-0.237*	0.119	(-0.470, -0.003)
Mixed	0.574*	0.108	(0.365, 0.789)
Others	-0.660*	0.207	(-1.076, -0.260)
Residential	(control)		
07:00–11:00	0.959*	0.155	(0.667, 1.282)
11:00–15:00	0.947*	0.162	(0.643, 1.284)
15:00–19:00	1.144*	0.153	(0.849, 1.460)
19:00–23:00	0.998*	0.142	(0.731, 1.289)
23:00–03:00	0.400*	0.163	(0.083, 0.726)
03:00–07:00	(control)		
Ln(PA)	0.235*	0.033	(0.168, 0.298)
Ln(VH)	0.168*	0.035	(0.108, 0.242)

10 * Statistically significant at the 5% level.

11

12 Table 4 Global Integration of Nine Samples

	G1	G2	G3	D1	D2	D3	I1	I2	I3
Network									
Integration	3.624	1.997	1.773	0.915	1.046	0.909	0.768	0.955	0.891

13 Note: G = grid; D = deformed grid; I = irregular.

14 The results presented in Table 4 show that a greater global integration value means the greater
 15 likelihood of a grid pattern. In addition, the grid pattern network is more compact because road
 16 density tends to be higher with higher global integration. The safety of the studied network patterns

1 generally improved as global integration decreased, with a grid being the least safe pattern and an
2 irregular pattern the safest. This finding is similar to that of Wang et al. (2013). Deformed grid and
3 irregular design patterns generally have less connectivity, frequent curves, and loops. The grid
4 pattern generally has straight roadways, which makes the process of driving comfortable and
5 predictable. However, the favorable driving environment on the grid pattern often increases fatigue
6 in drivers and reduces their attention, which may enhance the possibility of pedestrian–vehicle
7 crashes occurring in critical situations (Rifaat et al., 2011, 2012). In contrast, the presence of frequent
8 curves and loops in the deformed grid and irregular design patterns may provoke drivers to be more
9 cautious. Further, given that the irregular pattern may restrict drivers' sight distance and make
10 vehicle control more difficult, drivers are likely to compensate with increased attention (Hedlund,
11 2000; Levym and Miller, 2000; McCarthy and Talley, 1999). Moreover, the grid network pattern has
12 higher potential for pedestrian–vehicle conflict points and less traffic calming than the deformed grid
13 and irregular design patterns.

14 Consistent with previous studies by Ewing and Dumbaugh (2009) and Siddiqui et al. (2012),
15 our results show that increased intersection density is associated with a higher chance of pedestrian–
16 vehicle crashes at the 5% significance level (mean = 0.088; 95% CI = [0.034, 0.142]). Given that the
17 intersection density is defined by the number of intersections divided by road length per TAZ, more
18 intersections and shorter road lengths in particular TAZs result in higher intersection density. The
19 higher intersection density of a given TAZ increases pedestrian exposure with a higher probability of
20 pedestrian–vehicle interactions, as intersections are often locations with a relatively high number of
21 pedestrians crossing streets and with unexpected driver behavior.

22 The effect of average speed on the occurrence of pedestrian–vehicle crashes was also explored.
23 The results imply that average speed reduces the likelihood of a pedestrian–vehicle crash occurring
24 at the 5% significance level (mean = -0.021; 95% CI = [-0.028, -0.015]). Contrary to the findings of
25 Gårder (2004) and Cuadrado et al. (2008), the results show that higher average speed is associated
26 with a lower pedestrian crash occurrence. The reduction in the number of pedestrian–vehicle crashes
27 in zones with higher average speeds may be due to favorable built environments in which road
28 network design features (such as separation of pedestrians from vehicles by time or space, and
29 measures that increase the visibility and conspicuity of pedestrians) are more reasonable and road
30 infrastructures (such as pedestrian refuge islands, roadway lighting, overpasses, and underpasses) are
31 relatively good at reducing the exposure of pedestrians to potential harm, both on the roadside and
32 when crossing (Garber and Gadirau, 1998).

33 Four land use patterns were examined to identify their influence on the occurrence of crashes
34 involving pedestrians: residential, commercial, mixed, and others. Compared with residential areas,
35 the likelihood of a pedestrian–vehicle crash occurring in the commercial areas (mean = -0.237; 95%
36 CI = [-0.470, -0.003]) and other areas (mean = -0.660; 95% CI = [-1.076, -0.260]) of Hong Kong is
37 significantly reduced at the 5% level, and the likelihood of a pedestrian–vehicle crash occurring in
38 mixed areas (mean = 0.574; 95% CI = [0.365, 0.789]) is significantly increased at the 5% level.
39 Contrary to the findings of previous studies by Kim et al. (2008) and Wier et al. (2009), the reduction
40 of pedestrian–vehicle crash occurrences in commercial and other areas of Hong Kong may be due to
41 the presence of additional overpasses. Fig. 5 illustrates two figures comparing the commercial and
42 residential areas during the weekday afternoon off-peak period on the same day. It can be seen that
43 vehicle flow and travel speed in commercial areas may be higher than in the residential areas
44 throughout the day. However, overpasses in commercial areas usually provide separate pedestrian

1 linkages to reduce the chances of conflicts compared with the level crossings in residential areas.
2 Moreover, in most of the residential areas of Hong Kong, ground-floor shops are popular and more
3 shops are generally associated with more attractions, restaurants, and dwelling units. This could
4 eventually increase the probability of pedestrian–vehicle crashes. The relationship between
5 pedestrian crashes and mixed areas has been supported by the finding of Ewing and Dumbaugh
6 (2009) that pedestrians in neighborhoods with mixed land use patterns have more difficulty
7 perceiving the environment as safe than do their counterparts in neighborhoods dominated by
8 monotonous land use patterns.



Commercial area

Residential area

Fig. 5. Comparison between commercial and residential areas

10
11 The effect of time of day on the occurrence of pedestrian-related crashes was also investigated.
12 A day was divided into six categories: 07:00–11:00 (morning), 11:00–15:00 (noon), 15:00–19:00
13 (afternoon), 19:00–23:00 (evening), 23:00–03:00 (middle of the night), and 03:00–07:00 (dawn).
14 Compared with the dawn period, the temporal distribution in the models represents the crash
15 heterogeneity at different times of day. The afternoon off-peak period from 15:00 to 19:00 has the
16 highest probability of pedestrian–vehicle crashes (mean = 1.144; 95% CI = [0.849, 1.460]). Although
17 the traffic flow in this period is lower than that in the peak period, the number of daily activities
18 (leisure, shopping, meetings, etc.) may be higher. The second and third highest probabilities of
19 pedestrian–vehicle crashes occur in the evening peak period (19:00–23:00) (mean = 0.998; 95% CI =
20 [0.731, 1.289]) and morning peak period (07:00–11:00) (mean = 0.959; 95% CI = [0.667, 1.282]).
21 During the peak periods, each TAZ is expected to have higher internal and external pedestrian
22 activities, for example, morning and evening walks, which increase the probability of
23 pedestrian-related crashes. The period with the fourth highest probability of accidents was the noon
24 off-peak period from 11:00–15:00 (mean = 0.947; 95% CI = [0.643, 1.284]). As more workers go to
25 restaurants for meals during this period, there may be an increase in pedestrian and vehicle trips. In
26 turn, this may increase pedestrian vulnerability to crashes. The period with the lowest probability of
27 accidents happening is the midnight period (23:00–03:00) (mean = 0.400; 95% CI = [0.083, 0.726]).
28 Although there are extremely low traffic flows during the midnight period, both drivers and
29 pedestrians may engage in higher risk behavior in poor lighting conditions, such as driving while
30 fatigued, aggressive driving, and driving under the influence of alcohol/drugs, which leads to more
31 pedestrian crashes. This finding is consistent with the results of previous studies by Lee and

1 Abdel-Aty (2005) and Moudon et al. (2011).

2 An increase in log-transformed VH significantly increases the likelihood of a pedestrian–vehicle
3 crash occurring at the 5% level (mean = 0.168; 95% CI = [0.108, 0.242]). This finding shows that
4 crash occurrences increase with increases of VH. The higher probability of pedestrian–vehicle
5 interactions due to increased vehicular traffic volumes explains why log-transformed VH is
6 positively correlated with the occurrence of crashes involving pedestrians. An increase in
7 logarithmically transformed PA (mean = 0.235; 95% CI = [0.168, 0.298]) also significantly increases
8 the likelihood of a pedestrian–vehicle crash occurring at the 5% level. This finding is consistent with
9 previous studies by Lam et al. (2014) and Yao et al. (2015). The higher log-transformed PA of
10 particular TAZs is highly correlated with more consequent pedestrian activities (compulsory trips,
11 shopping trips, escort trips, etc.). A higher volume of pedestrian activities is expected to increase the
12 possibility of pedestrian–vehicle crashes.

13 5. Conclusions

14 The present study attempted to investigate the role of road network patterns and zone-related
15 variables in zone-level pedestrian safety analysis. To determine the explicit relationship between road
16 network characteristics and pedestrian safety, three topological indices (connectivity, local
17 integration, and global integration) of road network patterns via space syntax were derived and used
18 to quantitatively represent the structures of road networks for each TAZ. Global integration
19 outperformed the other two local measurements due to its ability to effectively quantify the different
20 road network patterns.

21 The relationships between pedestrian–vehicle crashes and various contributory factors were
22 developed, compared, and analyzed using both non-spatial PLN (Model 1) and spatial CAR models
23 with different neighborhood structures (Models 2a–2c). The results showed that the CAR models
24 outperformed the PLN model in terms of goodness-of-fit, indicating that spatial correlation should be
25 addressed when modeling pedestrian–vehicle crashes for spatially aggregated data. Moreover, the
26 model with the proximity structure based on road network connectivity (Model 2c) was found to
27 have significantly better model performance than those based solely on contiguity (Model 2a) and
28 geometry-centroid distance (Model 2b). This neighborhood structure used the total number of road
29 segments as a quantitative index of traffic volumes to reflect the connections between TAZs. This
30 configuration was appealing in the case of neighboring zones with different vehicular traffic and
31 pedestrian volumes. The goal was to determine the geographical adjacency and traffic association
32 simultaneously.

33 A significant positive relationship between global integration and crashes involving pedestrians
34 could be observed, meaning higher global integration was related to more pedestrian–vehicle crashes.
35 A greater global integration value clearly revealed closer associations between the connections
36 within road networks and was correlated with more potential for pedestrian–vehicle conflicts. Among
37 the three mentioned road network patterns, the grid pattern network had the highest global
38 integration value and the irregular pattern network had the smallest value. Thus, a grid pattern
39 network was the least safe, a deformed grid pattern network was the second safest, and an irregular
40 pattern network was the safest.

41 For future study, one direction is to conduct spatial analysis for both larger-scale (e.g., broad
42 districts of Hong Kong) and smaller (e.g., census tracts) area aggregations. In some applications
43 there is a need to consider different scales of aggregation within analysis. Scale change issues are

1 often considered modifiable areal unit problems that may lead to biased inferences.

2 **Acknowledgments**

3 The research described here is jointly supported by grants from the National Natural Science
4 Foundation of China (71301083), the Research Foundation of Tsinghua University (No.
5 20151080412), the Research Grants Council of the Hong Kong Special Administrative Region,
6 China (Project No. 717512), and the Joint Research Scheme of the National Natural Science
7 Foundation of China/Research Grants Council of Hong Kong (Project No. 71561167001 &
8 N_HKU707/15). We would like to thank Concord Pacific Satellite Technologies Limited and Motion
9 Power Media Limited for proving the GPS taxi data for the research work that is reported in this
10 paper. We also appreciate the constructive comments from anonymous reviewers.

11 **References**

- 12 Aguero-Valverde, J., Jovanis, P.P., 2008. Analysis of road crash frequency with spatial models. *Transportation*
13 *Research Record*, 2061, 55–63.
- 14 Aguero-Valverde, J., Jovanis, P.P., 2010. Spatial correlation in multilevel crash frequency models: Effects of
15 different neighboring structures. *Transportation Research Record*, 2165, 21–32.
- 16 Appleyard, D., 1980. Livable streets: Protected neighborhoods? *Annals of the American Academy of Political and*
17 *Social Science*, 451, 106–117.
- 18 Besag, J., York, J., Molli, E.A., 1991. Bayesian image restoration with two applications in spatial statistics. *The*
19 *Annals of the Institute of Statistics and Mathematics* 43(1), 1-59.
- 20 Carlin, B.P., Ma, H., 2007. Bayesian multivariate areal wombling for multiple disease boundary analysis. *Bayesian*
21 *Analysis*, 2(2), 281–302.
- 22 Cuadrado, J., Dahdah, S., Motevalli, V., 2008. Analysis of the 1994-1998 pedestrian crash data study (PCDS)
23 conducted by the national automotive sampling system (NASS). *Transportation Research Board 87th Annual*
24 *Meetings*, 13-17, Washington, DC.
- 25 Dong, N., Huang, H., Xu, P., Ding, Z., Wang, D., 2014. Evaluating spatial-proximity structures in crash prediction
26 models at the level of traffic analysis zones. *Transportation Research Record*, 2432, 46–52.
- 27 Dumbaugh, E., Zhang, Y., 2013. The relationship between community design and crashes involving older drivers
28 and pedestrians. *Journal of Planning Education and Research*, 33, 83–95.
- 29 Ewing, R., Dumbaugh, E., 2009. The built environment and traffic safety: A review of empirical evidence. *Journal*
30 *of Planning Literature*, 23(4), 347–367.
- 31 Ewing, R., Meakins, G., Bjarnson, G., Hilton, H., 2011. Transportation and land use. In: A.L. Dannenberg, H.
32 Frumkin, & R.J. Jackson (Eds.), *Making healthy places: Designing and building for health, well-being, and*
33 *sustainability* (pp. 149–169). Washington, DC: Island Press.
- 34 Florida, R., 2002. *The rise of the creative class*. New York: Basic Books.
- 35 Frank, L.D., Engelke, P., Schmid, T., 2003. *Health and community design: The impact of the built environment on*
36 *physical activity*. Washington, DC: Island Press.
- 37 Garber, N.J., Gadirau, R., 1998. *Speed variance and its influence on accidents*. AAA Foundation for Traffic Safety,
38 Washington, DC.
- 39 Gårder, P.E., 2004. The impact of speed and other variables on pedestrian safety in Maine. *Accident Analysis and*
40 *Prevention*, 36(4), 533–542.
- 41 Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2014. *Bayesian data analysis*. Boca Raton, FL, USA: Chapman

1 and Hall/CRC.

2 Gomez-Ibanez, D.J., Boarnet, M.G., Brake, D.R., Cervero, R.B., Cotugno, A., Downs, A., Hanson, S., Kockelman,
3 K.M., Mokhtarian, P.L., Pendall, R.J., Santini, D.J., Southworth, F., 2009. *Driving and the built environment:
4 The effects of compact development on motorized travel, energy use, and CO₂ emissions*. USA: Oak Ridge
5 National Laboratory.

6 Hedlund, J., 2000. Risky business: safety regulations, risk compensation, and individual behavior. *Injury Prevention*,
7 6, 82–90.

8 Hillier, B., 1996. *Space is the machine: A configurational theory of architecture*. UK: Cambridge University Press.

9 Hillier, B., Hanson, J., 1984. *The social logic of space*. UK: Cambridge University Press.

10 Hong Kong Transport Department (HKTD), 2014. *Road traffic accident statistics*. Transport Department, Hong
11 Kong.

12 Huang, H., Abdel-Aty, M., 2010. Multilevel data and Bayesian analysis in traffic safety. *Accident Analysis and
13 Prevention*, 42(6), 1556–1565.

14 Huang, H., Zeng, Q., Pei, X., Wong, S.C., Xu, P., 2016. Predicting crash frequency using an optimised radial basis
15 function neural network model. *Transportmetrica A: Transport Science*, 12(4), 330–345.

16 Jiang, B., 1999. SimPed: Simulating pedestrian crowds in a virtual environment. *Journal of Geographic Information
17 and Decision Analysis*, 3(1), 21–30.

18 Kim, J.K., Ulfarsson, G.F., Shankar, V.N., Kim, S., 2008. Age and pedestrian injury severity in motor-vehicle
19 crashes: A heteroskedastic logit analysis. *Accident Analysis and Prevention*, 40(5), 1695–1702.

20 Lam, W.H.K., Hung, W.T., Lo, H.K., Lo, H.P., Tong, C.O., Wong, S.C., Yang, H., 2003. Advancement of the annual
21 traffic census in Hong Kong. *Proceedings of the Institution of Civil Engineers, Transport*, 156(2), 103–115.

22 Lam, W.W.Y., Yao, S., Loo, B.P.Y., 2014. Pedestrian exposure measures: A time-space framework. *Travel
23 Behaviour and Society*, 1(1), 22–30.

24 Lee, C., Abdel-Aty, M., 2005. Comprehensive analysis of vehicle–pedestrian crashes at intersections in Florida.
25 *Accident Analysis and Prevention*, 37(4), 775–786.

26 Levym, D.T., Miller, T., 2000. Review: risk compensation literature—the theory and evidence. *Journal of Crash
27 Prevention and Injury Control*, 2(1), 75–86.

28 Liu, X.K., 2007. *Model and analyze Hong Kong street network using space syntax theory*. Shandong: Qingdao
29 University.

30 Loo, B.P.Y., Hung, W.T., Lo, H.K., Wong, S.C., 2005. Road safety strategies: A comparative framework and case
31 studies. *Transport Reviews*, 25(5), 613–639.

32 Loo, B.P.Y., Wong, S.C., Hung, W.T., Lo, H.K., 2007. A review of the road safety strategy in Hong Kong. *Journal
33 of Advanced Transportation*, 41(1), 3–37.

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

36 Lord, D., Miranda-Moreno, L.F., 2008. Effects of low sample mean values and small sample size on the estimation
37 of the fixed dispersion parameter of Poisson-gamma models for modeling motor vehicle crashes: a Bayesian
38 perspective. *Safety Science*, 46, 751–770.

39 Lovegrove, G.R., Sun, J., 2010. Using community-based macrolevel collision prediction models to evaluate safety
40 level of neighborhood road network patterns. Presented at 89th Annual Meeting of the Transportation
41 Research Board, Washington, DC.

42 Lovegrove, G.R., Sun, J., 2013. Comparing the road safety of neighbourhood development patterns: Traditional
43 versus sustainable communities. *Canadian Journal of Civil Engineering*, 40(1), 35–45.

44 Marshall, S., 2004. *Streets and patterns*. New York: Spon Press.

1 Marshall, W.E., Garrick, N.W., 2010a. Street network types and road safety: A study of 24 California cities. *Urban*
2 *Design International*, 15, 133–147.

3 Marshall, W.E., Garrick, N.W., 2010b. Effect of street network design on walking and biking. *Transportation*
4 *Research Record*, 2198, 103–115.

5 Marshall, W.E., Garrick, N.W., 2011. Does street network design affect traffic safety? *Accident Analysis and*
6 *Prevention*, 43(3), 769–781.

7 McCarthy, P., Talley, W., 1999. Evidence on risk compensation and safety behavior. *Economic Letters*, 62, 91–96.

8 Miaou, S.P., Song, J.J., Mallick, B.K., 2003. Roadway traffic crash mapping: A space–time modeling approach.
9 *Journal of Transportation and Statistics*, 6, 33–58.

10 Mitra, S., 2009. Spatial autocorrelation and Bayesian spatial statistical method for analyzing intersections prone to
11 injury crashes. *Transportation Research Record*, 2136, 92–100.

12 Moudon, A.V., Lin, L., Jiao, J., Hurvitz, P., Reeves, P., 2011. The risk of pedestrian injury and fatality in collisions
13 with motor vehicles, a social ecological study of state routes and city streets in King County, Washington.
14 *Accident Analysis and Prevention*, 43(1), 11–24.

15 Naci, H., Chisholm, D., Baker, T.D., 2009. Distribution of road traffic deaths by road user group: A global
16 comparison. *Injury Prevention*, 15(1), 55–59.

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

19 Pei, X., Wong, S.C., Ai, X., Shi, W., 2009. Estimation of speed distribution using floating car survey method.
20 Presented at 8th International Conference on Traffic and Granular Flow, Shanghai, China.

21 Pei, X., Wong, S.C., Sze, N.N., 2011. A joint-probability approach to crash prediction models. *Accident Analysis*
22 *and Prevention*, 43(3), 1160–1166.

23 Pei, X., Wong, S.C., Sze, N.N., 2012. The roles of exposure and speed in road safety analysis. *Accident Analysis*
24 *and Prevention*, 48, 464–471.

25 Quddus, M.A., 2008. Modeling area-wide count outcomes with spatial correlation and heterogeneity: An analysis
26 of London crash data. *Accident Analysis and Prevention*, 40(4), 1486–1497.

27 Raftery, N., Chiaradia, A., Gil, J., 2007. Space syntax: The role of urban form in cyclist route choice in central
28 London. Institute of Transportation Studies: UC Berkeley Traffic Safety Center.

29 Ratti, C., 2004. Space syntax: Some inconsistencies. *Environment and Planning B: Planning and Design*, 31(4),
30 487–499.

31 Rifaat, S., Tay, R., De Barros, A., 2010. Effect of street pattern on road safety are policy recommendations sensitive
32 to aggregations of crashes by severity? *Transportation Research Record*, 2147, 58–65.

33 Rifaat, S., Tay, R., De Barros, A., 2011. Effect of street pattern on the severity of crashes involving vulnerable road
34 users. *Accident Analysis and Prevention*, 43(1), 276–283.

35 Rifaat, S., Tay, R., De Barros, A., 2012. Urban street pattern and pedestrian traffic safety. *Journal of Urban Design*,
36 17(3), 337–352.

37 Siddiqui, C., Abdel-Aty, M., Choi, K., 2012. Macroscopic spatial analysis of pedestrian and bicycle crashes.
38 *Accident Analysis and Prevention*, 45, 382–391.

39 Siddiqui, C., Abdel-Aty, M., Choi, K., 2014. Implications of pedestrian safety planning factors in areas with
40 minority and low-income populations. *International Journal of Sustainable Transportation*, 8, 360–381.

41 Southworth, M., Ben-Joseph, E., 2003. *Streets and the shaping of towns and cities*. Washington, DC: Island Press.

42 Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van Der Linde, A., 2002. Bayesian measures of model complexity and
43 fit. *Journal of the Royal Statistical Society Series B*, 64(4), 583–639.

44 Tong, C.O., Hung, W.T., Lam, W.H.K., Lo, H.K., Lo, H.P., Wong, S.C., Yang, H., 2003. A new survey methodology

1 for the annual traffic census in Hong Kong. *Traffic Engineering and Control*, 44(6), 214–218.

2 Vuong, Q., 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57, 307-334.

3 Wakefield, J.C., Best, N.G., Wallgen, L., 2000. Bayesian approaches to disease mapping. In: P. Elliott, J. Wakefield,
4 N. Best, & D. Briggs (Eds.), *Spatial epidemiology, methods and applications*. UK: Oxford University Press.

5 Wang, X., Jin, Y., Abdel-Aty, M., Tremont, P.J., Chen, X., 2012. Macrolevel model development for safety
6 assessment of road network structures. *Transportation Research Record*, 2280, 100–109.

7 Wang, X., Wu, X., Abdel-Aty, M., Tremont, P.J., 2013. Investigation of road network features and safety
8 performance. *Accident Analysis and Prevention*, 56, 22–31.

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

11 Washington, S.P., Karlaftis, M.G., Mannering, F.L., 2011. *Statistical and econometric methods for transportation*
12 *data analysis*, Second edition. Boca Raton, FL, USA: Chapman and Hall/CRC.

13 Wei, V.F., Lovegrove, G., 2012. Sustainable road safety: A new (?) neighbourhood road pattern that saves VRU
14 lives. *Accident Analysis and Prevention*, 44(1), 140–148.

15 Wier, M., Weintraub, J., Humphreys, E.H., Seto, E., Bhatia, R., 2009. An area-level model of vehicle–pedestrian
16 injury collisions with implications for land use and transportation planning. *Accident Analysis and Prevention*,
17 41(1), 137–145.

18 Wong, S.C., Hung, W.T., Lo, H.K., 2002. *Road safety strategy and implementation*. Shenzhen, China: China Public
19 Security Publisher.

20 Wong, S.C., Sze, N.N., Li, Y.C., 2007. Contributory factors to traffic crashes at signalized intersections in Hong
21 Kong. *Accident Analysis and Prevention*, 39(6), 1107–1113.

22 Wong, W., Wong, S.C., 2015. Systematic bias in transport model calibration arising from the variability of linear
23 data projection. *Transportation Research Part B: Methodological*, 75, 1–18.

24 Wong, W., Wong, S.C., 2016. Network topological effects on the macroscopic bureau of public roads function.
25 *Transportmetrica A: Transport Science*, 12(3), 272–296.

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

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

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

33 Zhang, Y., Bigham, J., Li, Z., Ragland, D., Chen, X., 2012. Association between road network connectivity and
34 pedestrian-bicyclist accidents. Presented at 91st Annual Meeting of the Transportation Research Board,
35 Washington, DC.

36 Zhang, Y., Bigham, J., Ragland, D., Chen, X., 2015. Investigating the associations between road network structure
37 and non-motorist accidents. *Journal of Transport Geography*, 42, 34–47.