Accident Analysis and Prevention 125 (2019) 249–256

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Macro-level traffic safety analysis in Shanghai, China

Xuesong Wang[a,b,*], Qingya Zhou[a,b], Junguang Yang[a,b], Shikai You[a,b], Yang Song[a,b], Meigen Xue[c]

[a] School of Transportation Engineering, Tongji University, Shanghai, 201804, China
[b] The Key Laboratory of Road and Traffic Engineering, Ministry of Education, Shanghai, 201804, China
[c] Shanghai City, Comprehensive Transportation Planning Institute

ARTICLE INFO

Keywords:
Macro-level safety modeling
Traffic analysis zone
Bayesian conditional autoregressive model
Transportation safety planning

ABSTRACT

Continuing rapid growth in Shanghai, China, requires traffic safety to be considered at the earliest possible stage of transport planning. Macro-level traffic safety studies have been carried out extensively in many countries, but to date, few have been conducted in China. This study developed a macro-level safety model for 263 traffic analysis zones (TAZs) within the urban area of Shanghai in order to examine the relationship between traffic crash frequency and road network, traffic, socio-economic characteristics, and land use features. To account for the spatial correlations among TAZs, a Bayesian conditional autoregressive negative binomial model was estimated, linking crash frequencies in each TAZ to several independent variables. Modeling results showed that higher crash frequencies are associated with greater populations, road densities, total length of major and minor arterials, trip frequencies, and with shorter intersection spacing. The results from this study can help transportation planners and managers identify the crash contributing factors, and can lead to the development of improved safety planning and management.

1. Introduction

Shanghai has developed rapidly in recent years. From 2003 to 2008, the modernized area of Shanghai expanded from 1505 km^2 to 2288 km^2 (Shanghai Urban and Rural Construction and Traffic Committee, 2010), while its roadway network increased from 10,451 km to 15,844 km (Shanghai Bureau of Statistics, 2004, 2009). This rapid growth has been accompanied by a sizeable challenge in traffic safety. In 2009, the fatality rate of Shanghai was 4.28 per ten thousand vehicles—a number much higher than in cities of developed countries (Accident Prevention Office of Shanghai Public Security Bureau Traffic Police Corps, 2010; Shanghai Urban Comprehensive Transportation Planning Institute, 2010). For example, in 2009, the rate was 1.37 in New York City (New York State Department of Motor Vehicle, 2009) and 1.19 in London (London's Road Casualties: Collisions, Injuries and Damned Statistics, 2010). In order to improve traffic safety, in 2012 the Shanghai Municipal Government started to compile a Transportation White Paper that identified safety as a primary objective and aimed to reduce the fatality rate by 25% within the next ten years (Shanghai Municipal Government, 2014). The current study was conducted in part to support this initiative.

It is noteworthy that, in developed countries, safety is often one of the most important objectives at the planning stage. For example, the U.S.'s Safe, Accountable, Flexible, Efficient, Transportation Equity Act required state and metropolitan planners to incorporate safety at the design stage of all new transportation projects (U.S. Department of Transportation, 2009). The consideration of safety in planning processes was established most recently in the U.S.'s NCHRP 8–76 (National Cooperative Highway Research Program, 2014). Internationally, 29 of 35 metropolitan areas, including London, Singapore, San Francisco Bay, and Boston, put safety in their comprehensive transportation plans (Peng et al., 2013).

One approach to improving safety that has emerged in recent years in the United States and Europe is the use of macro-level crash prediction models to identify key crash contributing variables. Researchers have focused on various areas, such as defining the appropriate study unit (e.g., traffic analysis zone, census block group, census tract (Abdel-Aty et al., 2013); crash spatial distribution patterns (Levine et al., 1995a; Lee et al., 2015a), the effects of macro-level impact factors on safety (Abdel-Aty et al., 2011; Dong et al., 2015), and implications for community design (Sun and Lovegrove, 2010; Dumbaugh and Zhang, 2013). However, the implementation of safety planning based on macro-level considerations requires much more attention in China.

Previous traffic safety research in Shanghai has been limited by a lack of data. Recently, however, the large amount of high-quality data needed to support macro-level safety analyses in Shanghai has become

* Corresponding author at: School of Transportation Engineering, Tongji University, Shanghai, 201804, China.
E-mail address: wangxs@tongji.edu.cn (X. Wang).

https://doi.org/10.1016/j.aap.2019.02.014
Received 30 March 2014; Received in revised form 19 December 2018; Accepted 11 February 2019
Available online 22 February 2019

available. In 2009, the Shanghai Municipal Government conducted a fourth comprehensive transportation survey and these data were used in this study. In the same year, a WebGIS-based Shanghai Road Traffic Crash Analysis and Warning System was developed by Wang et al. (2012a) for the Shanghai Traffic Police Bureau, and this capability has been applied to this and other safety studies. Based on the crash data of Shanghai in 2009, several studies were conducted that provided a basis for macro-level safety analyses of, for example, urban arterials (Wang et al., 2012a, 2015), suburban arterials (Wang et al., 2013a, 2014; Wang et al., 2016b), signalized intersections (Xie et al., 2013, 2014), and urban expressways (Xu et al., 2012; Yu et al., 2016). A meso-level safety analysis (Li and Wang, 2017) was also conducted based on the high-quality Shanghai crash data.

The main objective of this paper is to identify the macro-level variables that have a significant influence on traffic crashes in Shanghai, the most heavily populated Chinese city. Four categories of independent variables were investigated: roadway, traffic, socio-economic, and land use features.

2. Literature review

Several researchers in recent years have performed macro-level safety analyses by relating crash frequencies to different characteristics, including roadway, traffic (especially exposure), socio-economic, and land use features. The influence of the various features are assessed at the zone level, into which researchers have aggregated crashes into spatial scales or units. The units commonly used are the block group (Levine et al., 1995b), ward (Noland and Quddus, 2004; Quddus, 2008), census tract (Abdel-Aty et al., 2013), and traffic analysis zone (TAZ) (Ng et al., 2002; Abdel-Aty et al., 2013; Dong et al., 2015; Wang et al., 2016a). Among these analytic units, the TAZ was selected for this study because it offers the most consistent results (Abdel-Aty et al., 2013).

2.1. Zonal features and macro-level safety effects

2.1.1. Roadways and safety

Extensive research has focused on the relationships between roadway characteristics and traffic safety. Roadway characteristics that have been examined include roadway type (e.g., arterial, collector, local), intersection type (e.g., 3-legged, 4-legged, multi-legged) and road network pattern (e.g., cul-de-sac, grid, sparse).

The relationship between roadway type and traffic crashes is commonly described by considering the safety effects of roadway length, density, and the percentage of specific types in a given zone. In previous studies, total road length has been positively associated with crash frequency (Hadayeghi et al., 2010; Wang et al., 2013b). Hadayeghi et al. (2003, 2006) considered the total length of major roads and of minor roads, and found, in both cases, that length was positively correlated with crashes. Length may have somewhat different effects on crashes when speed limit is considered. Abdel-Aty et al. (2013) showed that roadway length was positively associated with crashes throughout the speed limit range of 25–65 mph, while Dong et al. (2014, 2015) showed a positive relationship only for road segments on which the speed limit was higher than 35 mph. Road density has been positively associated with crashes (Rifaat and Tay, 2010), and Huang et al. (2010) observed the same positive relationship specifically on freeways, principle arterials, and minor arterials. The percentage of certain road types within a given zone has also been shown to have a positive correlation with crashes. Wang et al. (2013b) found a positive relationship with the percentage of arterials, and Xu and Huang (2015) found a positive relationship with road segments having a 45-mph speed limit.

In addition to roadway type, researchers have also studied the relationships between crashes and intersections. The total number of intersections within an analytical unit and the intersections' density have both been positively associated with crashes (Hadayeghi et al., 2003, 2006; Huang et al., 2010; Rifaat and Tay, 2010; Abdel-Aty et al., 2011; Abdel-Aty et al., 2013; Dong et al., 2014; Xu et al., 2014; Xu and Huang, 2015), and Wang et al. (2012b, 2013b) showed that, on state roads, the number of signalized intersections was also positively associated with crash occurrence. Although Lovegrove and Sayed (2006) found the percentage of 3-legged intersections to be negatively correlated with crashes, Hadayeghi et al. (2010) looked at both the shape and control type of intersections, and found that the number of 3-legged and 4-legged signalized intersections were positively associated with crashes.

Road network pattern has also been shown to be associated with crashes. Rifaat and Tay (2010) found that both warped parallel and loops & lollipops patterns were safer than grid patterns. This result is supported by Wang et al. (2012b, 2013b), who found that TAZs with grid street patterns were associated with the highest crash frequencies, followed by parallel, mixed, loops and lollipops, and sparse street patterns.

2.1.2. Traffic characteristics and safety

Traffic characteristics have been validated by many existing studies to also have great impact on safety. Vehicle miles traveled (VMT), has been found to be positively associated with traffic crashes in TAZ-level analyses (Hadayeghi et al., 2010; Abdel-Aty et al., 2013; Dong et al., 2014; Lee et al., 2014; Dong et al., 2015; Xu and Huang, 2015), while average volume over capacity (V/C) has been shown to be negatively associated (Hadayeghi et al., 2003, 2006). Researchers have also found that the average annual daily traffic (AADT) for trucks is a statistically significant crash predictor (Huang et al., 2010). In their TAZ-level analysis, Abdel-Aty et al. (2011) demonstrated that the best predictors of total crashes and peak hour crashes were total trip production and total trip attraction.

2.1.3. Socio-economic variables and safety

Socio-economic variables that have been considered as crash predictors include population, employment status, and economic level. Population was found to be positively associated with traffic crashes (Kim et al., 2006; Hadayeghi et al., 2010; Wang et al., 2012b, b), as was population density (Lee et al., 2014; Dong et al., 2014; Xu et al., 2014; Xu and Huang, 2015). Furthermore, vulnerable road users, including younger people (0–15 years) and older people 65 years and older), have been found to be underinvolved in crashes (Abdel-Aty et al., 2013; Wang et al., 2012b), a finding possibly attributable to the decreased physical activity of these vulnerable road users. However, Abdel-Aty et al. (2013) also found that both higher density of children and higher proportion of minority population were associated with increased traffic crashes.

A higher employed population was reported to be positively associated with total crashes (Hadayeghi et al., 2003), while a higher unemployment rate has been positively associated with severe traffic crashes (Huang et al., 2010). The percentage of low-income households has been shown to have a positive correlation with crash frequency (Rifaat and Tay, 2010), as has the proportion of households without vehicles (Lee et al., 2014). However, increased median household income has been associated with reduced traffic crashes (Huang et al., 2010; Dong et al., 2014; Xu et al., 2014; Dong et al., 2015; Xu and Huang, 2015).

2.1.4. Land use and safety

Many researchers have found that commercial land use is associated with greater crash frequency (Ng et al., 2002; Kim et al., 2006; Hadayeghi et al., 2010; Rifaat and Tay, 2010). Resource and industrial land use have also been positively associated with crashes (Hadayeghi et al., 2010; Rifaat and Tay, 2010), and Wang et al. (2012b, 2013b) found that the number of schools in a given analysis unit correlated to increased traffic crashes. These results suggest that land developed for commercial, industrial, and educational use may attract more traffic, thereby leading to more crashes.

2.2. Statistical modeling approaches

Researchers have traditionally used Poisson or negative binomial (NB) regression models to develop the statistical relationships between macro-level variables and crash frequency (Ng et al., 2002; Noland and Quddus, 2004). The Poisson distribution is a useful starting point to model crash outcomes. However, the underlying assumption of the Poisson distribution is that the variance is equal to the mean, an assumption that is often violated in the crash count data where the variance can be greater than the mean. To account for this issue, commonly referred to as overdispersion, NB models have been recently been more commonly used. The NB generalized linear model (GLM) framework, however, assumes that the zones selected for study are spatially independent. This assumption, too, is often violated because TAZs near each other are often similar in nature and thus in safety performance. In these cases, spatial autocorrelation or aggregation in crash data preparation might lead to bias in model estimation and statistical inference (MacNab, 2004).

To mitigate the effects of spatial correlation, the Bayesian conditional autoregressive (CAR) model has been successfully used, and has gained in popularity, because it can effectively accommodate the spatial correlations of the analyzed units (MacNab, 2004; Quddus, 2008; Huang et al., 2010; Wang et al., 2012b). For example, Huang et al. (2010) used a Bayesian spatial model with CAR priors, and found a superior ability to accommodate spatial dependencies. Consequently, the Bayesian NB CAR model was used to consider the spatial correlation in this paper.

3. Data preparation

The TAZ was the unit selected for this study for its consistent results. The development of macro-level safety models at the TAZ level required reliable traffic and crash data, quantitatively defined roadway networks, and TAZs with properly delineated boundaries.

3.1. TAZ delineation

The area studied was comprised of 263 TAZs in the 645-km^2 downtown area within Shanghai's Outer Ring. To ensure the TAZ boundaries were compatible with jurisdictional boundaries, the Huangpu River, the urban elevated expressways, and rail lines were not included in the TAZs' road networks. Land use types within each TAZ were checked to ensure that they were homogeneous. Fig. 1 below shows the TAZ boundaries within the Outer Ring.

3.2. Matching data from multiple sources

Crash and road network data were provided by the Shanghai Police Office in GIS format. As is common, such as in a U.S. study (Wang et al., 2012b), the TAZ boundaries and the digital road network did not match perfectly, which can be seen Fig. 2. To ensure accurate assignment of crashes to their proper TAZs, the TAZ boundaries with mismatch problems were manually redrawn based on road alignments, as described by Wang et al. (2012b).

3.3. Description of independent variables

The independent variables were classified into four main categories: roadway features (e.g. total number of 3-legged intersections), traffic features (e.g. car trip generation), land use features (e.g. land use type) and socio-economic feature. Data continues to be limited for socio-economic features such as income level, but population was included as an independent variable. Population data for 2009 for the 263 TAZs were obtained from the Shanghai Urban Planning and Design Research Institute.

Roadway network features were collected for each TAZ, including total number of 3-legged, 4-legged, and multi-legged intersections; average intersection spacing; total number of access points; total length of major arterials, total length of minor arterials, and total length of all other roads. The GIS Spatial Join function in ArcGIS® was used to divide, associate, and aggregate roadway attributes (e.g., length, volume) among the corresponding TAZs. To treat roadway features along TAZ boundaries, a method proposed by Sun and Lovegrove (2010) was applied. Polyline features on the TAZ boundaries were shared by corresponding TAZs by pro-rating those features with a weight equal to the reciprocal of the number of related TAZs. For example, an intersection on the boundary of three TAZs was assigned a weight of 0.333 to each TAZ, and a road on the boundary of two TAZs was assigned a weight of 0.5 to each TAZ. ArcGIS® was then applied to calculate the density of each type of road or intersection within each TAZ.

Collected traffic features included car trip generation and attraction. Peak hour exposure data for trip generation within each TAZ and trip attraction to each TAZ were extracted from the 2009 VISUM® model (which used the same TAZ boundaries as the current study) in the *Research on Driving Condition and Safety of a Typical Mega-city* program by Tongji University and General Motors (2009–2011)Tongji University, 2009Tongji University and General Motors (2009–2011). To facilitate model validation, 24-hour screen line traffic count locations were selected, and it was also arranged to collect Floating Car Data (FCD) provided by GPS-equipped taxis along 17 typical Shanghai roads. These point data were used to validate use of the VISUM® model for this study, by comparing the differences between simulated volumes from the model's traffic assignment and the observed screen line counts. The GEH index was used, which is an empirical formula similar to a chi-squared statistical test that is commonly employed in traffic engineering to compare two sets of data. The lower the GEH, the better the model's goodness-of-fit, with a preferred GEH value of less than 10 (UK Department of Transport, 1996). GEH values for this study were lower than 10 for 69% of all validation points, and were lower than 20 for 90% of the validation points.

Car and truck vehicle kilometers travelled (VKT) and mean speed data were provided by the Shanghai Urban and Rural Construction and Transportation Development Research Institute and the Shanghai City Comprehensive Transportation Planning Institute, generated by running road traffic system simulation models using Emme 4.0® software. This study examined the precision of the Institute simulation results by comparing them with 7 screen lines of Shanghai Comprehensive Transportation Survey data, which included manual survey data from 350 sections, loop detector data from 250 sections, and freeway toll data from 90 sections. Examination showed that, on the whole, the average error between simulation results and survey data was below 15%, which indicates acceptable reliability of the simulation results.

Land use varies considerably in Shanghai by area, making area type a key variable. As shown in Fig. 1, the area west of the Huangpu River is the Puxi District, and the area east of the river is the Pudong District. Puxi is older and more densely developed than Pudong, and its roads are narrower and carry more traffic. Land use type and intensity data were collected from the Shanghai Municipal Project Management Office. For this study, land use intensity was classified into three levels based on the floor area ratio (ratio of a building's total floor area to the size of the piece of land upon which it is built) and the average building density. Land use types (e.g., residential, industrial) were classified into eight categories, and are shown in Table 1.

To check for multicollinearity, collinearity diagnostics were conducted in SPSS®. Results showed that all variance inflation factors (VIF) were lower than 10, which indicates an absence of multicollinearity issues. Pearson and Spearman correlation tests were also run using SPSS®. Other than mild correlations between car trip generation and length of major arterials (coefficient of 0.544), and between mean speed and truck VKT (coefficient of 0.53), all correlation coefficients were below 0.5, showing overall, no significant correlation between independent variables. Descriptive statistics of the independent

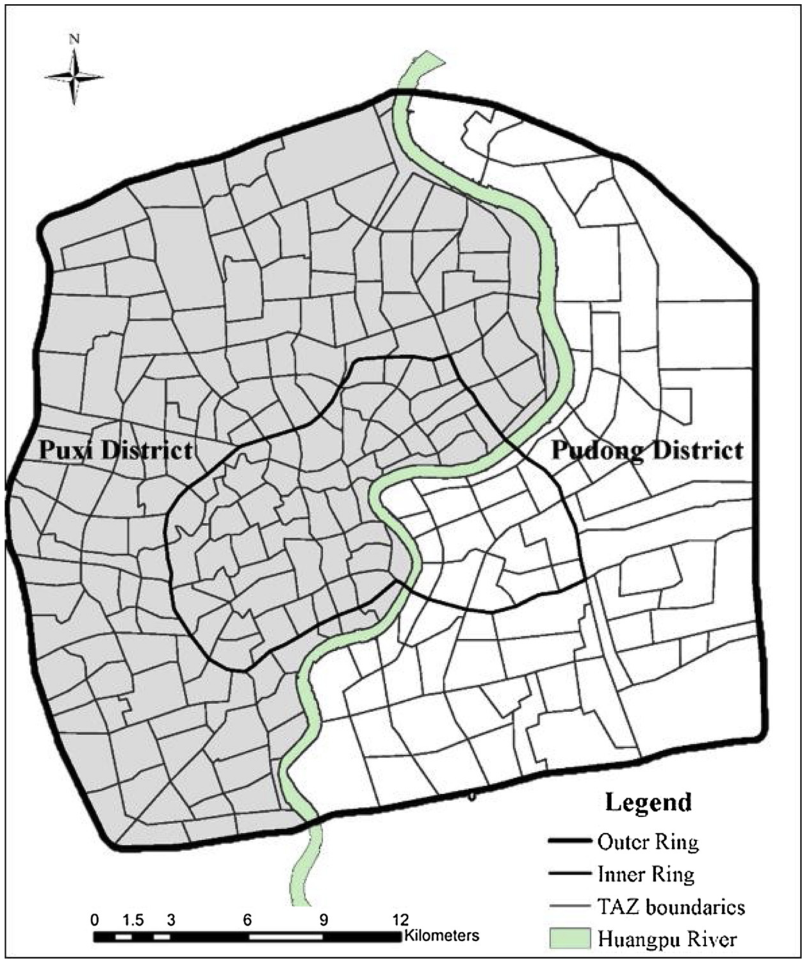

Fig. 1. TAZ boundaries within the Outer Ring.

variables for the 263 TAZs are shown in Table 1 below.

3.4. Description of the dependent variable: crash frequency

High-quality crash data for Chinese roadways had been limited until 2006, when the Chinese Ministry of Public Security unified the road traffic crash collection protocol and developed a standardized form for crash documentation. In the same year, the Shanghai traffic police began to prioritize data quality, and built an electronic database of crash information. In 2009, they collaborated with a Tongji University research team (Wang et al., 2012a) to develop the Shanghai Road Traffic Crash Analysis and Warning System, which uses a Web-GIS based system for crash geocoding. Since then, the police have used the system for analyzing crash data, for example, examining the causes of crashes and identifying hotspots. In 2012, the Shanghai traffic police further standardized the language specifying crash location, and follow strict protocols for acquiring and entering crash data. Five essential elements, including road name, side of the road, and reference point and its direction and distance from the crash, are all required to be recorded for traffic crashes and violations. Before crash data is entered into the system, the police officer checks the accuracy of the location by comparing the crash diagrams with the system's geocoded crash points. In this way, more than 98% of crashes are precisely located.

This study collected police-reported crashes occurring in 2009 within the Shanghai study area. They had been classified into four categories according to severity levels, property damage only (79.86% of total crashes), slight injury (19.93%), severe injury (0.16%), and fatality (0.06%), but only total crashes were considered for this study.

Crashes were assigned to the TAZ containing the road segment in which the crash occurred. A crash occurring on a TAZ boundary was assigned to the closest TAZ if the exact location, e.g., side of the road, of the crash was known; if a precise location could not be confirmed, the crash would be proportionally assigned to the contiguous TAZs using Sun and Lovegrove's (2010) method described above in Section 3.3. As this method sometimes resulted in fractional values for the total number of crashes for a TAZ, the fractional crash frequencies were rounded to the nearest integer. As the median number of crashes with each TAZ was 547, this rounding method had minimal impact.

3.5. Moran's I test for crash spatial correlation

Moran's I, a spatial autocorrelation index, was used in this study to determine whether the observed crashes were spatially correlated among adjacent zones. A positive value of Moran's I indicates positive spatial correlation or clustering within the study area. Moran's I index can be converted to a Z-score and then tested for significance: values greater than 1.96 or lower than -1.96 show significant spatial autocorrelation in the region (Huang et al., 2010). For this study, the Moran's I value was 0.410, which converted to a Z-score of 11.485. This Z-score is significant at the 0.01 level, indicating that crashes were spatially correlated or clustered.

4. Model development

To capture the spatial correlation, or dependence, among TAZs, a Bayesian CAR NB model was developed. Following Wang et al. (2012b),

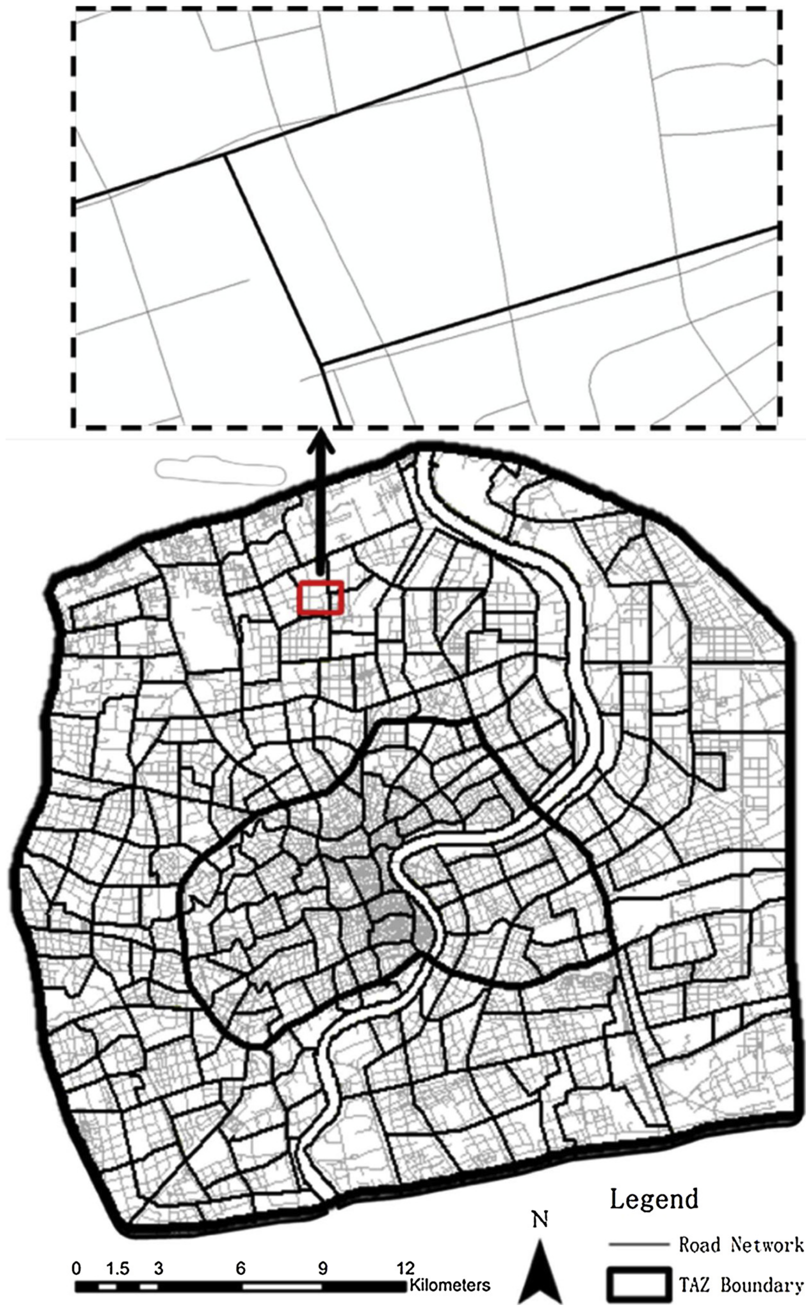

Fig. 2. Area TAZ boundaries and data match problem in Shanghai.

the Bayesian CAR NB model was developed using the frequency of crash occurrence in the studied TAZs as the dependent variable.

y_i represents the number of crashes occurring in TAZ_i, and the Bayesian CAR NB model captures the spatial dependence among TAZs by introducing a random effect variable, ϕ_i into the model. It is defined as (Besag, 1974):

$$\log(\theta_i) = \Psi_i = X'\beta + \varepsilon_i + \phi_i \tag{1}$$

where θ_i is the expectation of y_i, X is the covariate matrix, β is the vector of regression coefficients, and ε_i is the gamma distributed error term.

A reciprocal matrix of the center distance of the analysis unit was adopted in this model to express the spatial relationships among TAZs. The proximity matrix W with entry $w_{i,j}$ indicates the spatial relationship between TAZ_i and TAZ_j:

$$w_{i,j} = 1/d_{i,j} \tag{2}$$

where $d_{i,j}$ is the distance between TAZ_i and TAZ_j.

In this study, the Bayesian method is implemented using a Markov Chain Monte Carlo (MCMC) algorithm, which has recently been used in many Bayesian-related studies, including Quddus (2008), Noland et al. (2013), Wang et al. (2013b), Dong et al. (2014) and Xu and Huang (2015). This study used the open source software WinBUGS® to calibrate the model. Due to the lack of reliable prior information, all the regression coefficients followed non-informative prior distributions, with the assumption that they obeyed normal distribution N (0, 10^5). In the process of developing the model, two MCMC chains of 40,000 iterations were run, and the first 2000 samples were discarded as burn-in. Using Brooks, Gelman and Rubin convergence diagnostics (Smith, 2005) in R®, the convergence of the MCMC chains was tested using criteria proposed by Smith (2005): "If the estimates are approximately equal to one (or, as a rule of thumb, the 0.975 quantile is ≤1.2), the

Table 1
Statistical description of independent variables.

a) Continuous Variables

Variables	Min	Max	Mean	S.D.
TAZ area (km^2)	0.463	18.055	2.453	2.183
Number of 3-legged intersections	0	184	20.240	23.560
Number of 4-legged intersections	4	124	15.570	15.019
Number of multi-legged intersections	0	9	1.116	0.692
Total number of intersections	7	244	36.926	35.778
Length of major arterials (km)	0.726	4.428	3.600	1.615
Length of minor arterials (km)	0.942	2.452	1.818	0.769
Road density (km/km^2)	0.473	53.726	6.961	6.926
Average intersection spacing (km)	0.226	0.521	0.387	0.235
Car trip generation (10^3pcu)	0.044	105.173	16.349	15.233
Car trip attraction (10^3pcu)	0.024	113.060	16.309	16.936
Car VKT (10^3pcu km)	0.4	640.945	180.720	137.482
Daily Truck VKT (10^3pcu km)	0	439.902	43.291	67.563
Mean Speed(km/h)	6.647	62.988	28.894	9.866
Population (10^3)	1.040	132.440	42.690	24.260

b) Categorical Variables

Variables and Description	Percentage
Land use types	
0: residential	22.4%
1: residential and commercial	28.9%
2: commercial, residential and official	11.0%
3: residential, cultural, educational and commercial	11.0%
4: exhibition, financial and tourism	7.6%
5: residential and industrial	9.1%
6: residential and logistics	7.2%
7: residential and ecological	2.7%
Land use intensity	
0: low	22.4%
1: medium	27.4%
2: high	50.2%
Area type	
0: Puxi District	70.7%
1: Pudong District	29.3%

samples may be considered to have arisen from a stationary distribution". This study's estimation results for all parameters were approximately equal to one, and the 0.975 quantile for all parameters was less than 1.2. Thus, we concluded that the chains converged.

5. Model results

The relationships between crash occurrence and the four types of variables were analyzed: roadway, traffic, socio-economic, and land use. A summary of the Bayesian CAR NB model's estimates is shown below in Table 2. Note that the CAR effect is statistically significant, which confirms the spatial autocorrelation among crashes across TAZs.

5.1. Roadway features

Modeling results show that several roadway features have significant effect on crash frequency. A key finding was that longer major arterials were associated with more crashes than shorter arterials. This finding confirms the results of previous studies such as Levine et al. (1995b) that showed the total length of major arterials in each census block group was positively associated with crash frequency. Similarly, another finding was that longer minor arterials were also associated with increased crashes. One possible explanation is that on many minor arterials there is no bicycle lane, which leads to increased conflicts between motorized and non-motor vehicles (Wang et al., 2012a). Such conflicts between vehicles, especially between heavy vehicles and non-motor vehicles, can lead to more crashes and fatalities (Hadayeghi et al., 2003; Lovegrove, 2007; Quddus, 2008).

Table 2
Posterior Estimates for Bayesian CAR NB model.

Variables	Bayesian CAR NB model			
	Posterior mean	Posterior S.D.	95% BCI	
Intercept	4.66	0.342	4.093	5.203
Population	0.004	0.002	0.001	0.007
Total length of major arterial	0.215	0.03934	0.1503	0.281
Total length of minor arterial	0.547	0.091	0.397	0.694
Average intersection spacing	−1.6	0.488	−2.403	−0.815
Road density	0.582	0.112	0.399	0.768
Car trip generation	0.008	0.005	−0.001	0.016
Area types				
Pudong vs. Puxi	−0.165	0.075	−0.288	−0.0402
Land use intensity				
medium vs. low[a]	0.101	0.107	−0.078	0.276
high vs. low[a]	0.0849	0.097	−0.075	0.243
Car VKT	6.88E-04	2.94E-04	1.97E-04	0.001
Truck VKT	9.28E-04	6.22E-04	−8.86E-05	0.002
Mean Speed	−0.008	0.004	−0.015	−0.002
CAR effect	0.22	0.186	0.041	0.626
DIC Value	3671.480			

[a] Not significant at 95%, 90%, or 75% Bayesian credible interval.

Results also showed that reduction in mean intersection spacing is associated with an increased likelihood of crashes. Xie et al. (2013) found that short intersection spacing had an adverse impact on safety, as shorter intersection spacing leads to more severe conflicts when vehicles attempt to change lanes.

Road density was also positively associated with crashes. Higher road densities tend to have more turning traffic, and more turns mean more conflicts. Aguero-Valverde and Jovanis (2006) reached a similar conclusion in Pennsylvania, where they found counties with higher road density had a significantly higher crash risk. This finding is important for Chinese cities, as many have higher road network densities than cities in other countries, (Xie et al., 2013).

5.2. Traffic, socio-economic, and land use features

The coefficients representing traffic features, including car VKT, truck VKT, and car trip generation within TAZs during peak periods were positively associated with traffic crash frequency, findings consistent with those of Abdel-Aty et al. (2011) and also Siddiqui et al. (2012). This association may be attributable to the increase in crash opportunities as traffic volume increases. Mean speed, however, was found to be negatively related with crash frequency. This was consistent with previous findings by Lovegrove and Sayed (2006).

Larger population was also positively associated crash frequency, concurring with studies by Hadayeghi et al. (2010), Wang et al. (2012b, 2013b), and Lee et al. (2015b). Larger populations tend to have higher levels of traffic activity, leading to increases in crash opportunities. With the exception of population, the above variables all operate at the roadway level, but this study found that the global variable of city district was also linked to differences in crash frequency. A greater number of crashes occurred in the Puxi District than in the Pudong District, possibly because the Pudong District, being newer, has better designed roadways. It is interesting to note in the modeling results, however, that both land use type and land use intensity were not significant at 95%, 90%, or 75% Bayesian credible intervals, which may have been due to the heterogeneous influences of unobserved factors such as the combinations of different land use types in TAZs.

6. Summary and discussion

This study conducted a macro-level traffic safety analysis based on data of 263 TAZs within the Outer Ring Road in Shanghai. Road network, traffic, socio-economic, and land use features for the 263 TAZs were collected, and in order to address the spatial correlation among adjacent TAZs, a Bayesian CAR NB model was developed. Both the model results and values of Moran's I and Z scores confirmed the significant spatial autocorrelation among crashes, i.e., that spatial correlation exists between TAZs.

The model estimates show that, from a macro-level perspective, zonal level variables including population, arterial (both major and minor) length, road density, car trip generation, car and truck VKT, and short intersection spacing are negatively associated with traffic safety. These findings, reported in the context of a Chinese metropolis, confirm the conclusions of existing studies worldwide (Hadayeghi et al., 2010; Levine et al., 1995b; Wang et al., 2012a; Abdel-Aty et al., 2011; Aguero-Valverde and Jovanis, 2006; Xie et al., 2013). Apart from zonal level safety influence factors, the global variable of city district also had a demonstrated impact on crash occurrence. Likely attributable in part to roadway network age difference, the possibility of uninvestigated design factors is raised.

This study's findings on safety influence factors can be used to provide suggestions and references for transportation planners and managers. Based on the associations between different factors and traffic crashes reported in this paper, integrated approaches to designing future low crash networks can be adopted. For example, the macro-level crash prediction model from this study can be used to identify TAZs with higher-than-expected crash levels, and area-based engineering, enforcement, and educational strategies and programs could set priorities for improving these TAZs. Results can also be used to predict TAZ-level changes in traffic crashes associated with socio-economic and land use development. Countermeasures could then be implemented by manipulating other variables such as road mileage, road network density, and intersection design.

Although a significant number of macro-level data were collected to support this study, several factors were not considered in the analysis, such as the effect of street pattern on traffic safety, which may have been an unidentified factor in the Pudong/Puxi district type influence. Also possibly contained in the district influence are additional socio-economic factors. More TAZ-level socioeconomic data, including income level and car ownership, should be considered in future macro-level research in order to better explore the crash contributing factors. Finally, this study only considered total crash numbers. Future research might also investigate crashes by severity and traffic mix. The abundance of non–motorized vehicles (mostly bicycles and electric bicycles) on China's roads may be an important influence at the macro level.

Acknowledgements

This study was jointly sponsored by the National Natural Science Foundation of China (51878498), the Science and Technology Commission of Shanghai Municipality (18DZ1200200), and the 111 Project (B17032).

References

Abdel-Aty, M., Siddiqui, C., Huang, H., Wang, X., 2011. Integrating trip and roadway characteristics in managing safety at traffic analysis zones. Transp. Res. Rec.: J. Transp. Res. Board 2213, 20–28.
Abdel-Aty, M., Lee, J., Siddiqui, C., 2013. Geographical unit based analysis in the context of transportation safety planning. Transp. Res. Part A: Policy Pract. 49, 62–75.
Accident Prevention Office of Shanghai Public Security Bureau Traffic Police Corps, 2010. A brief introduction of road traffic accidents in Shanghai in 2009. Traffic Transp. 26 (1), 72–73.
Aguero-Valverde, J., Jovanis, P.P., 2006. Spatial analysis of fatal and injury crashes in Pennsylvania. Accid. Anal. Prev. 38 (3), 618–625.
Besag, J., 1974. Spatial interaction and the statistical analysis of lattice systems. J. R. Stat. Soc. Ser. B (Methodol.) 192–236.
Dong, N., Huang, H., Xu, P., Ding, Z., Wang, D., 2014. Evaluating spatial proximity structures in TAZ-level crash prediction models. Transp. Res. Rec.: J. Transp. Res. Board 2432, 46–52.
Dong, N., Huang, H., Zheng, L., 2015. Support vector machine in crash prediction at the level of traffic analysis zones: assessing the spatial proximity effects. Accid. Anal. Prev. 82, 192–198.
Dumbaugh, E., Zhang, Y., 2013. The relationship between community design and crashes involving older drivers and pedestrians. J. Plann. Educ. Res. 33 (1), 83–95.
Hadayeghi, A., Shalaby, A.S., Persaud, B.N., 2003. Macro level accident prediction models for evaluating safety of urban transportation systems. Transp. Res. Rec.: J. Transp. Res. Board 1840, 87–95.
Hadayeghi, A., Shalaby, A.S., Persaud, B.N., Cheung, C., 2006. Temporal transferability and updating of zonal level accident prediction models. Accid. Anal. Prev. 38 (3), 579–589.
Hadayeghi, A., Shalaby, A.S., Persaud, B.N., 2010. Development of planning level transportation safety models using full Bayesian Semiparametric Additive Techniques. Transportation Research Board 88th Annual Meeting.
Huang, H., Abdel-Aty, M., Darwiche, A.L., 2010. County-level crash risk analysis in Florida: Bayesian spatial modeling. Transportation Research Board 89th Annual Meeting.
Kim, K., Brunner, I.M., Yamashita, E., 2006. Influence of land use, population, employment, and economic activity on accidents. Transp. Res. Rec.: J. Transp. Res. Board 1953, 56–64.
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.
Levine, N., Kim, K.E., Nitz, L.H., 1995a. Spatial analysis of Honolulu motor vehicle crashes: I. Spatial patterns. Accid. Anal. Prev. 27 (5), 663–674.
Levine, N., Kim, K.E., Nitz, L.H., 1995b. Spatial analysis of Honolulu motor vehicle crashes: II. Zonal generators. Accid. Anal. Prev. 27 (5), 675–685.
Li, J., Wang, X., 2017. Safety analysis of urban arterials at the meso level. Accid. Anal. Prev. 108, 100–111.
London's road casualties: collisions, injuries and damned statistics, 2010. http://www.theguardian.com/uk/davehillblog/2010/jun/22/transport-for-london-road-safety-statistics-crap-cycling-waltham-forst-blog (Accessed 10.6.22).
Lovegrove, G.R., 2007. Road Safety Planning: New Tools for Sustainable Road Safety and Community Development. VDM Verlag Dr. Müller E K, Berlin, Germany.
Lovegrove, G.R., Sayed, T., 2006. Macro-level collision prediction models for evaluating neighborhood traffic safety. Can. J. Civ. Eng. 33 (5), 609–621.
MacNab, Y.C., 2004. Bayesian spatial and ecological models for small-area accident and injury analysis. Accid. Anal. Prev. 36 (6), 1019–1028.
National Cooperative Highway Research Program, 2014. Institutionalizing Safety in Transportation Planning Processes: Techniques, Tactics and Strategies. Washington, D.C.. .
New York State Department of Motor Vehicle, 2009. Summary of New York City Motor Vehicle Accidents. https://dmv.ny.gov/statistic/2009nycsummary.pdf.
Ng, K., Hung, W., Wong, W., 2002. An algorithm for assessing the risk of traffic accident. J. Saf. Res. 33 (3), 387–410.
Noland, R.B., Quddus, M.A., 2004. A spatially disaggregate analysis of road casualties in England. Accid. Anal. Prev. 36 (6), 973–984.
Noland, R.B., Klein, N.J., Tulach, N.K., 2013. Do lower-income areas have more pedestrian casualties? Transportation Research Board 92nd Annual Meeting.
Peng, J., Wang, X., Yang, J., 2013. Transportation problems and metropolitan planning in China: insights from latest metropolitan planning abroad. Transportation Research Board 92nd Annual Meeting.
Quddus, M.A., 2008. Modeling area-wide count outcomes with spatial correlation and heterogeneity: an analysis of London crash data. Accid. Anal. Prev. 40 (4), 1486–1497.
Rifaat, S.M., Tay, R., 2010. Effect of street pattern on road safety: are policy recommendations sensitive to different aggregations of crashes by severity? Transportation Research Board 89th Annual Meeting.
Shanghai Bureau of Statistics, 2004. Shanghai Statistical Yearbook 2003. China Statistics Press, Beijing.
Shanghai Bureau of Statistics, 2009. Shanghai Statistical Yearbook 2008. China Statistics Press, Beijing.
Shanghai Municipal Government, 2014. Shanghai Transportation Development White Paper 2020. Shanghai Municipal Government, Shanghai.
Shanghai Urban and Rural Construction and Traffic Committee, 2010. Shanghai Urban Comprehensive Transportation Planning Institute, Shanghai Fourth Comprehensive Transportation Office. Shanghai Fourth Comprehensive Traffic Survey Report.
Shanghai Urban Comprehensive Transportation Planning Institute, 2010. Shanghai comprehensive transportation annual report 2010 (Abstract). Traffic Transp. 26 (5), 1–4.
Siddiqui, C., Abdel-Aty, M., Huang, H., 2012. Aggregate nonparametric safety analysis of traffic zones. Accid. Anal. Prev. 45, 317–325.
Smith, B.J., 2005. Bayesian Output Analysis Program (BOA) Version 1.1 User's Manual. http://www.public-health.uiowa.edu/boa/BOA.pdf.
Sun, J., Lovegrove, G.R., 2010. Using community-based macro-level collision prediction models to evaluate the safety level of neighborhood road network patterns. Transportation Research Board 89th Annual Meeting.
Tongji University, General Motors, 2009–2011. Research on Driving Condition and Safety of Typical Mega-city. Shanghai, China.

U.S. Department of Transportation, 2009. Safe, Accountable, Flexible, Efficient, Transportation Equity Act: A Legacy for Users (SAFETEA-LU). Washington, D.C.. .
UK Department of Transport, 1996. Traffic Appraisal in Urban Areas, Volume 12 Section 1 of the Design Manual for Roads and Bridges (DMRB v12s1).
Wang, X., Chen, M., Liu, H., 2012a. Crash estimation for urban arterials considering their operational conditions. Transportation Research Board 91st Annual Meeting.
Wang, X., Jin, Y., Tremont, P., Abdel-Aty, M., 2012b. Macro level model development for safety assessment of road network structures. Transp. Res. Rec.: J. Transp. Res. Board 2280, 100–109.
Wang, X., Song, Y., Huang, H., 2013a. Safety analysis of suburban arterials in Shanghai. Transportation Research Board 92nd Annual Meeting.
Wang, X., Wu, X., Abdel-Aty, M., Tremont, P.J., 2013b. Investigation of road network features and safety performance. Accid. Anal. Prev. 56, 22–31.
Wang, X., Song, Yang, Yu, R., Schultz, G., 2014. Safety analysis of suburban arterials in Shanghai. Accid. Anal. Prev. 70, 215–224.
Wang, X., Fan, T., Chen, M., Deng, B., Wu, B., Tremont, P., 2015. Safety modeling of urban arterials in Shanghai, China. Accid. Anal. Prev. 83, 57–66.
Wang, X., Yang, J., Lee, C., You, S., 2016a. Macro-level safety analysis of pedestrian crashes in Shanghai, China. Accid. Anal. Prev. 96, 12–21.
Wang, X., Yuan, J., Schultz, G., Meng, W., 2016b. Investigating the safety impacts of the roadway network features of suburban arterials in Shanghai. Transportation Research Board 95th Annual Meeting.
Xie, K., Wang, X., Chen, X., Huang, H., 2013. Corridor-level signalized intersection safety analysis in Shanghai, China using Bayesian Hierarchical Models. Accid. Anal. Prev. 50, 25–33.
Xie, K., Wang, X., Ozbay, K., Yang, H., 2014. Crash frequency modeling for signalized intersections in a high-density urban road network. Anal. Methods Accid. Res. 2, 39–51.
Xu, P., Huang, H., 2015. Modeling crash spatial heterogeneity: random parameter versus geographically weighting. Accid. Anal. Prev. 75, 16–25.
Xu, C., Wang, X., Chen, X., 2012. Urban expressway speed spatial inconsistency and its effect on safety. Transportation Research Board 91st Annual Meeting.
Xu, P., Huang, H., Dong, N., Abdel-Aty, M., 2014. Sensitivity analysis in the context of regional safety modeling: identifying and assessing the modifiable areal unit problem. Accid. Anal. Prev. 70, 110–120.
Yu, R., Wang, X., Yang, K., Abdel-Aty, M., 2016. Crash risk analysis for shanghai urban expressways: a Bayesian semi-parametric modelling approach. Accid. Anal. Prev. 95, 495–502.