Accident Analysis and Prevention 148 (2020) 105775

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Effects of state-led suburbanization on traffic crash density in China: Evidence from the Chengdu City Proper

Si Qiao, Anthony Gar-On Yeh, Mengzhu Zhang *, Xiang Yan

Department of Urban Planning and Design, The University of Hong Kong, Pokfulam Road, Hong Kong

ARTICLE INFO

Keywords:
Built environment
State-led suburbanization
Neighborhood-level automobile-involved crash density
Safe urban form

ABSTRACT

Road crashes have become a leading cause of death in China. Although enormous efforts have been exerted to determine the factors that affect individual crash incidents, neighborhood-level crash incidence in Chinese cities has not been sufficiently analyzed. This study fills this gap by quantifying the effects of built environment factors on neighborhood-level automobile-involved crash density (NACD) in urban China and identifying its mediators and mediating effects. In American suburbs, urban sprawl is widely recognized to render neighborhoods unsafe for residence, thus leading to a high crash incidence. This study compares the characteristics of built environments between inner-city neighborhoods and the new neighborhoods that have been developed through China's state-led suburbanization since 2008 to reveal how this suburbanization provides a safer neighborhood environment. A structural equation model is used to examine the relationships among suburbanization, built environment factors, and NACD in the city proper of Chengdu, the largest metropolis in southwest China. Thus, this study contributes new empirical evidence to the debates over urban designs that are safest for traffic. Moreover, this study enriches our understanding of different sociospatial consequences between American-style urban sprawl and China's state-led suburbanization.

1. Introduction

In recent decades, China has been undergoing rapid urbanization and motorization alongside its "economic miracle". Accompanying China's great transformation from an underdeveloped agricultural society to a modern motorized urban society is an improvement in living conditions. However, associated with this improvement is a surge in road traffic mortality and injuries. According to Jiang et al. (2017), road injuries have become a leading cause of death in China. Approximately 80 % of accidental deaths in China are caused by traffic crashes, which have killed over 200,000 people and injured more than 500,000 people each year (Wang et al., 2019). Even worse, a traffic crash occurs in China every minute, and a Chinese is killed by traffic crashes every 2 min. Increasing automobile-involved traffic crashes (thereafter crash/-crashes), crash-related injuries, and deaths have become critical detriments to the health and wellbeing of Chinese citizens (Jiang et al., 2017). Therefore, the factors that affect crash incidence in China must be identified and understood to help improve the safety of neighborhood conditions.

Although considerable efforts have been exerted to investigate the risk factors that affect automobile-involved crash incidence in China, the existing literature is replete with gaps concerning this topic. First, most studies focused on factors affecting individual-level crash incidence, especially risky driving/riding behaviors (Wu et al., 2018a; Chu et al., 2019; Wang et al., 2019), traffic climate (Chu et al., 2019; Qu et al., 2019; Wu et al., 2018b), driving skills (Xu et al., 2018; Lyu et al., 2018), and other individual socioeconomic factors (Zhang et al., 2014). However, neighborhood-level automobile-involved crash density (NACD) has not been sufficiently empirically analyzed. Understanding crash incidence that occurs at the neighborhood level, the smallest unit of urban governance, is essential because it enables governmental sectors to identify communities where residents encounter a high risk of traffic crashes. This measure will facilitate the improvement in traffic safety of those communities in a systematic and tar2geted manner through state interventions. This strategy can include geographically targeted interventions to improve the efficiency of traffic management related to traffic crash mitigation. The neighborhood is where people live. Traffic hazards in a neighborhood can prevent the residents from going out freely and safely; thus, they have restricted mobility compared with residents residing in safer communities (Dumbaugh and Rae, 2009;

* Corresponding author.
E-mail addresses: siqiao@hku.hk (S. Qiao), hdxugoy@hkucc.hku.hk (A. Gar-On Yeh), mengzhu.zhang@pku.edu.cn (M. Zhang), shawnyan@hku.hk (X. Yan).

https://doi.org/10.1016/j.aap.2020.105775
Received 21 March 2020; Received in revised form 13 August 2020; Accepted 3 September 2020
Available online 16 October 2020

Ewing and Dumbaugh, 2009). This traffic hazard-induced immobility can lead to few life opportunities and transport-related social exclusion of certain spatially aggregated social groups (Lucas, 2012). Thus, NACD can be a factor for the (re)production of sociospatial inequalities. Needless to say, this phenomenon warrants further investigation.

Second, the existing literature highlights the importance of behavioral interventions (e.g., penalty system and traffic safety awareness training) in reducing crash incidence (Zhang et al., 2014; Wu et al., 2018c). However, the role of built environment characteristics in affecting crash incidence has been largely ignored. Ewing and Dumbaugh (2009) emphasized that the design of built environments is important to improve traffic safety and thus should be given due attention. Accumulating empirical evidence from American cities substantiates the importance of urban form or built environment factors in affecting crash incidence (Chen et al., 2015; Dumbaugh and Rae, 2009; Ewing et al., 2016a). Although Chinese cities differ from American cities in many aspects, especially in terms of built environments and travel behaviors, investigating safe indicators of built environments in China with the goal of reducing traffic crashes is of equal importance. Moreover, empirical findings may potentially advance our understanding of various mechanisms of crash incidence due to built environments formed under different historical and political development paths.

Lastly, China's rapid suburbanization since 2008 has remarkably reshaped suburban built environments and neighborhood forms in various cities. However, little effort has been exerted to understand how this suburbanization process affects traffic safety from the perspective of built environments. In American cities, suburbanization is characterized by urban sprawls that engender heavy reliance on private cars. However, overreliance on private cars is substantially positively related to traffic accidents (Ewing, 1997; Fischer et al., 2013). Thus, given that automobile dependence and heavy automobile use are a widespread consequence of urban sprawls, American suburban neighborhoods are associated with higher crash incidence and considered more unsafe than inner-city neighborhoods (Ewing and Dumbaugh, 2009). By contrast, China's post-2008 suburbanization process is characterized by proactive state interventions in designing built environments. Do China's newly developed suburban neighborhoods tend to be (un)safe due to proactive built environment designs? Answering this question is vital as it can provide invaluable inputs to policy formation regarding traffic safety concerns in urban land developments in other countries.

This study attempts to fill these gaps in research by investigating the relationships among state-led suburbanization, built environments, and NACD in China. Using the city proper of Chengdu as a case study, the largest metropolis in southwest China, this study attempts to answer three research questions: (1) What are the differences in built environment characteristics between inner-city neighborhoods and newly developed suburban neighborhoods? (2) How do built environment factors affect NACD? (3) How does NACD in suburban neighborhoods differ from inner-city neighborhoods from the perspective of built environments? Chinese authorities seldom disclose data on traffic crashes and built environments, particularly at the neighborhood level. To address these limitations, a big data approach is used to collect points of interest (POIs) of built environment characteristics from an open street map (Baidu). A one-year continuous tracking dataset provided by a local insurance company is utilized to calculate NACD. By operating the data through ArcGIS, a dataset on neighborhood-level built environment characteristics and NACD is formed for the final structural equation model (SEM).

The study commences with a review of the relationships among built environments, crash incidence, and suburbanization, which are primarily drawn from empirical evidence from American cities. A brief comparison between American-style urban sprawls and China's post-2008 state-led suburbanization is conducted. Section 3 elaborates on the methodology of the study. Section 4 presents and discusses empirical findings and modeling results. This study culminates with our key conclusions, limitations, and policy implications for mitigating NACD.

2. Literature review

2.1. Linking built environments to crash incidence: mediators, measures, and models

Central to the understanding of the interplay between built environments and NACD is the relationship among exposure, risks, and crash incidence (Merlin et al., 2020; Dumbaugh and Rae, 2009). Although numerous studies have examined the correlations between built environment factors and crash incidence, the lack of insights into the underlying mechanism impedes effective policy formation on (traffic)

Table 1
Measures of built environments and their theoretical relationship to crash incidence.

Category	Example measures	Exposure	Risk	Sample papers
(1) Land use				
Density	Population density, job density, and residential density	(+) Increases exposure, particularly for pedestrians and cyclists	(?) May increase complexity of interactions; may reduce speed	Dumbaugh and Rae (2009), Fischer et al. (2013)
Mixed land use	Entropy as index, the ratio of commercial or residential land use	(?) Increases exposure, particularly for pedestrians and cyclists; may reduce exposure through reductions in VKT	(?) Increases the complexity of interactions; may reduce risk by communicating lower speeds	Ukkusuri et al. (2012), Dumbaugh and Rae (2009), Dumbaugh and Li (2010), Pulugurtha et al. (2013)
(2) Street design				
Road type	Density of roads of different classes (i.e., arterial and expressway)	(−) Wide, fast roads typically have more vehicles but fewer pedestrians and cyclists	(+) Fairly strong evidence that high speed, multilane roads in urban areas increase risk	Dumbaugh and Li (2010), Huang et al. (2010), Quistberg et al. (2015)
Intersection form and number	Density of intersection of different forms (three-leg and four-leg intersections and roundabouts), segment-to-intersection ratio	(?) May increase the number of pedestrians/cyclists but reduce vehicle traffic	(?) May increase complexity of interactions; may reduce speed	Lee et al. (2015), Osama and Sayed (2016), Marshall and Garrick (2011), Dumbaugh and Li (2010)
Pedestrian /bicycling infrastructure	Length of sidewalk, presence of bike lanes	(+) Increases exposure, particularly for pedestrians and cyclists	(?) Findings are mixed	Wei and Lovegrove (2013), Yu (2015), Cai et al. (2017)
Others	Density of bus stop, ratio/density of signalized intersection			

safety-oriented built environment designs (Merlin et al., 2020). Attempts to obtain a firm grasp of the role of traffic **exposure** and **risk** as mediators through which built environments affect crash incidence fill this gap. Theoretically, a built environment affects crash incidence in two indirect ways. First, from the perspective of traffic generation, a built environment is associated with the production of travel origins and destinations, specifically, travel demand and various mobility modes from which residents can choose and the attraction of traffic of different modes in an area. Thus, a built environment is pertinent to local traffic volumes. Several studies have reported that traffic volumes determine the exposure of people to traffic, and high exposure is associated with a high crash incidence, although the relationship may not be linear (Lee et al., 2019a; Merlin et al., 2020). In this respect, a built environment influences NACD by affecting traffic volumes that people experience during daily mobility. Second, a built environment affects the risk or likelihood of encountering crashes when people are involved in traffic. This causality is underpinned by the fact that some segments/neighborhoods with certain features are more dangerous or more likely to induce crashes when traffic volume is the same (Merlin et al., 2020).

Discerning the exact ways by which a built environment affects crash incidence (meditators and mediating effects) is vital in formulating practical policy interventions to improve the safety of our living conditions. This identification requires scholars to improve the methodology previously developed by transport engineering researchers to examine more rigorously the built environment factors affecting crash incidence from two aspects. The first factor is the measurement of a built environment. Notwithstanding the fact that built environment factors are widely considered in accident/collision analysis, they are only roughly measured or merely regarded as controlled variables, far from being systematically categorized and investigated (Quddus, 2008; Pulugurtha et al., 2006, 2013; Osama and Sayed, 2016). Recent efforts from planning researchers have achieved progress in categorizing and measuring built environment factors that could affect NACD either through literature reviews (Merlin et al., 2020; Dumbaugh and Rae, 2009) or empirical investigations (Ewing et al., 2016a; Guerra et al., 2019). According to these researchers, the built environment factors related to traffic safety can be categorized into **(i) land use characteristics** and **(ii) street design characteristics.** Land use characteristics include population density and land-use diversity that are measured at an area level, for instance, census tracts (Fischer et al., 2013) and traffic analysis zones (Pulugurtha et al., 2013). Street design characteristics include the number/density of intersections in different forms (i.e., three-leg/four-leg intersections and roundabouts); the presence/predominance of arteries and other types of roads that are measured at an area level (Guerra et al., 2019); the number of vehicle lanes, bike lanes, and crosswalks; sidewalk length; the signalization of intersection/crosswalks; and the presence of transit stops that are measured either at an area level (by density) or at a segment level (by ratio).

Table 1 provides a summary of several factors that have a positive association with crash incidence and some that are negatively associated. Through a comprehensive review of existing empirical evidence, Merlin et al. (2020) concluded that the effects of most built environment factors on NACD are mixed. Three reasons arguably contribute to the mixed results. First, the association may vary with crash type. Certain studies used NACD at the total level as a dependent variable (Yu and Xu, 2018; Guerra et al., 2019; Lee et al., 2013), whereas others examined a specific crash type, for instance, automobile-involved pedestrian crash (Chen and Shen, 2016; Ding et al., 2018; Lee et al., 2018, 2019b) and automobile-involved bicycle crash (Chen, 2015; Chen and Fuller, 2014; Hamann and Peek-Asa, 2013). Agreeing that different crash types are differently sensitive to built environment conditions is reasonable. Second, the association between built environment factors and NACD may not be linear. The effect (strength and direction) varies with the intensity of built environment conditions, and elementary regression approaches cannot capture this nonlinearity. Two recent studies underpin this argument. Yu and Xu (2018) investigated local variations in the associations between built environment factors and NACD. By implementing a geographical regression approach, they found that the effects of the presence of high-speed roads and the proportion of commercial land use on NACD vary from the periphery to downtown areas. Ding et al. (2018) quantified the nonlinear relationship between built environment factors and NACD by adopting a machine learning approach. Their results showed that the elasticity of household density on NACD decreases when household density is below 0.1 and then increases with household density. Once household density reaches 0.6, the elasticity remains at 2.3. Third, the demographic–socioeconomic composition of an area is not controlled. Numerous studies that applied a disaggregated approach to crash incidence have discovered that some social groups are sensitive to certain crash types (Regev et al., 2018; Rolison et al., 2018). Thus, the population composition of a neighborhood can complicate the relationship between built environments and NACD.

Another aspect that has recently attracted the attention of planning scholars is the quantitative approach. In the literature, the negative binomial model (Hadayeghi et al., 2006; Noland, 2003; Noland and Quddus, 2004) and Poisson regression (Cheng et al., 2013; El-Basyouny and Sayed, 2009; Ye et al., 2018) are commonly used to capture the associations among social, economic, and spatial factors and NACD. Recent studies have gained a deeper understanding of the spatial effect mechanism by using three other quantitative approaches. The first approach is the spatial model; for instance, geographically weighted regression is utilized to determine the spatial variations of associations (Quddus, 2008; Bao et al., 2017). The second approach is SEM, which is employed to disentangle the mediating effects. The mechanism by which an improvement in a specific built environment factor reduces NACD can be determined via this approach (by reducing traffic volume or by lowering the risk) (Ewing et al., 2016a). The third approach is the machine learning approach. For example, the additive Poisson regression trees are used to unravel nonlinear relationships between built environment factors and NACD (Ding et al., 2018). Moreover, a Chi-squared automatic interaction detection decision tree is consulted to rank the relative importance of contributing factors (Prati et al., 2017). With this approach, policymakers can determine which aspect of built environment can be improved to reduce NACD most efficiently in a specific neighborhood. In the present study, we do not intend to contribute to the methodology. Instead, we attempt to select an appropriate model to resolve the empirical questions raised in the next section.

2.2. New town plans and state-led suburbanization in China

Deteriorating traffic safety amid rapid suburbanization has generated intense interest among researchers. Although urban sprawls have long been a critical topic among American academics, scholars have only started paying attention to their effects on NACD. A key concern is that an uncompacted urban form/built environment caused by urban sprawl may not only induce more extensive vehicle use but also increase crash risk (Merlin et al., 2020). Incompactness/sprawl is commonly defined as a form of suburban built environment characterized by (1) low density, (2) low diversity (rigid separation of homes, workplace, and shops), (3) low centrality (of employments and population), and (4) low-accessibility road networks marked by large block sizes (Ewing et al., 2003). Numerous studies have found high vehicle miles traveled and a low share of active travels in suburban neighborhoods as a result of urban sprawls in American suburban areas (Ewing and Hamidi, 2015; Ewing et al., 2016a; Yeo et al., 2015; Ewing et al., 2003; Trowbridge and McDonald, 2008). Thus, urban sprawls increase NACD by increasing the exposure of people to vehicle traffic. Moreover, by examining the relationship between sprawl indices (composite measures of built environments) and city-level crash incidence, Ewing et al. (2016a, 2016b) found that urban sprawls increase crash incidence by increasing risk and traffic volume.

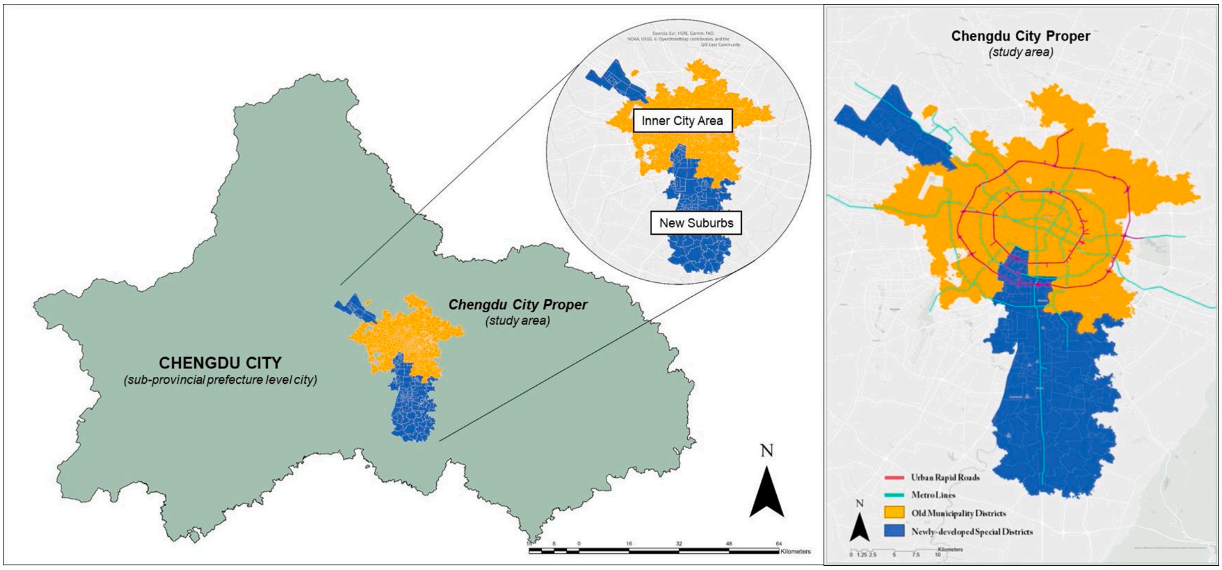

Fig. 1. Geographic location of Chengdu City Proper (left) and, urban rapid roads and metro lines in Chengdu City Proper (right).

Table 2
Variables used in the model and their measures and data sources.

Variables	Variable measures	Data source	Time of data
Dependent variable			
Crash density	Log of number of recorded crashes/area (km^2) of a neighborhood	Local insurance company	2018
Independent variable			
Exposure (endogenous)			
Traffic volume	Relying on dripping massive floating car data	Didi Company	2018
Land use (exogenous)			
Population density	Log of residential population of a neighborhood/area of a neighborhood calculated by residential footprint, building floors, and per capita living area	Local construction dataset	2018
Mixed land use	Entropy index calculated from POI dataset with categories of medical service, educational service, industrial land use, leisure facilities, living facilities, shopping and caterings, firms, and public sectors[a]	Baidu Map	2018
Street design (exogenous)			
Road type			
Expressway density	Log of length of the expressway in a neighborhood/area of a neighborhood	Baidu Map	2018
Arterial density	Log of length of arterial in a neighborhood/area of a neighborhood	Baidu Map	2018
Secondary road density	Log of length of secondary road in a neighborhood/area of a neighborhood	Baidu Map	2018
Branch road density	Log of length of branch road in a neighborhood/area of a neighborhood	Baidu Map	2018
Intersection form			
Three-leg intersection density	Log of number of three-leg intersections in a neighborhood/area of a neighborhood	Baidu Map	2018
Four-leg intersection density	Log of number of four-leg intersections in a neighborhood/area of a neighborhood	Baidu Map	2018
Pedestrian infrastructure			
Sidewalk density	Log of length of sidewalk in a neighborhood/area of a neighborhood	Baidu Map	2018
Others			
Bus stop density	Log of number of bus stops in a neighborhood/area of a neighborhood	Baidu Map	2018
Metro station density	Log of number of metro station in a neighborhood/area of a neighborhood	Baidu Map	2018

[a] $LUM = -[\sum_{i=1}^{n} p_i * \ln(p_i)]/\ln(n)$, where *LUM* is the entropy-based land use mixed score, *Pi* is the proportion of neighborhood covered by land use *i* to the summed area of land use categories of interest, and *n* is the number of land use categories of interest. A land use mix score of 1 indicates the highest mix possible. By contrast, a land use mix score of 0 indicates the area contains a single land use.

The rapid suburbanization in China since 2008 has raised concerns about traffic safety. Since 2008, among 281 prefecture-level cities in China, 272 have launched new suburban town buildings. On average, each city have had 2.5 new suburban towns constructed since 2008 (Chang and Lu, 2017). The median area of new suburban towns is 40 km^2, and a total of 66,300 km^2 of new suburban town areas had been constructed at the national level (Chang and Lu, 2017). However, owing to the leading role of local governments in designing suburban built environments, China's state-led suburbanization may differ from American-style urban sprawls in terms of the resultant suburban built environments and NACD. China's post-2008 suburbanization is the result of state-led channeling of excessive capital from the production sector to the construction of built environments through centralized new suburban town plans (Wu, 2019). In the course of suburbanization of employments and population, municipalities have proactively established comprehensive land use plans for constructing new suburban towns. Affected by new urbanism, Chinese urban planners applied the ideas of transit-oriented development and traditional neighborhood designs to new town planning. In particular, new suburban towns in large cities are characterized by metro-oriented land development. The newly developed employment and residential centers were constructed around newly developed metro lines to promote the compactness of suburban built environments.

Knowledge about China's post-2008 state-led suburbanization and its sociospatial consequences is limited. This study attempts to fill this gap from the perspective of built environment characteristics and NACD. We adopt a quantitative approach to answer three key questions:

(1) What are the differences in built environment characteristics between inner-city neighborhoods and newly developed suburban neighborhoods?
(2) How do built environment factors affect NACD?
(3) Compared with traditional inner-city neighborhoods, do newly developed suburban neighborhoods have higher/lower NACD as a result of different built environment characteristics?

3. Methodology

3.1. City context

The city proper of Chengdu is selected as a case study to investigate empirically the relationship between NACD and built environments. As the capital city of Sichuan Province, Chengdu City[1] has a large population of 16 million and a considerable GDP size of 1.5 trillion CNY. As such, the city proper of Chengdu has been able to rapidly suburbanize in the recent decade. From 2008 to 2018, the built-up areas of the city had increased from 368 km^2 to 802 km^2, and this progress was made possible by the municipality's ambitious "Tianfu New Area Plan" and "Hi-Tech Industrial Development Zone Plan" and by establishing several state-owned enterprises to finance and build new town buildings (Wu, 2019). In this respect, Chengdu City is representative of Chinese cities that have been rapidly suburbanizing in the recent decade. Thus, Chengdu City Proper presents an excellent case for examining the relationship between built environments and crash incidence in the context of rapid suburbanization. Second, Chengdu City Proper is a plain urban area characterized by ring road-made leading traffic networks (see Fig. 1). Such an urban form enables a clear identification of the boundary between new suburban town areas and old inner-city areas. Moreover, new development areas are connected to inner cities via the metro to engage mature resources.

3.2. Study unit and area

Neighbourhood (*shequ*), the fundamental administration unit and census zone in Chinese cities, is taken as the study unit. In China, "*shequ*" is regarded as a neighborhood that is officially defined and governed by Community Residents Committee. It is similar to street blocks in Western cities. Chengdu is a sub-provincial and prefecture-level city in China (Fig. 1). It is consisted of a city proper with 5 old municipal districts (i.e., Jinjiang, Wuhou, Qingyang, Jinniu, and Chenghua) and two newly developed special districts since 2008 (i.e., Hi-Tech Industrial Development Zone and Tianfu New Area), 7 outer suburban municipal districts and 8 rural counties (including county-level cities). Chengdu City is a unified urban function area directly governed by the Chengdu City Government. This study selects the Chengdu City Proper as the study area (Chan, 2007). As illustrated in Fig. 1, the five old municipality districts constitute the inner city area of Chengdu City Proper, and the two newly developed special districts comprise the main suburban area of Chengdu City Proper. The inner city has 460 neighborhoods, whereas the newly developed suburban area has 154 neighborhoods. A total of 614 neighborhoods of Chengdu City Proper are selected for this study.

3.3. Data sources and variable specifications

We use multiple data sources for the study (Table 2). Data on crash incidence are obtained from the official record of all claim settlements of automobile-involved crashes in Chengdu City Proper in 2018. This dataset does not cover all crashes that happened because claims were not made for some minor accidents. However, the dataset is excellent in terms of data reliability and comprehensiveness compared with that of the government statistics.[2] Records on crashes from the government statistical report usually exclude single-vehicle crashes that did not lead to injuries, such as a crash between a vehicle and a tree near a road. In such a case, the car owner involved in the accident needs to notify

[1] Chengdu City is one of the sub-provincial prefecture level cities in China. In China, administratively, there are three types of cities: central-administered cities (administered directly by the central government), prefecture-level cities (administered by a province or autonomous region) and county-level cities (administered by a prefecture-level city or its equivalent) cities. A central-administered or prefecture-level city usually comprises a main central urban area (also known as the "city proper") and surrounding rural areas. It may also have one or more subordinate county-level cities. A total of 4 central-administered (namely, Beijing, Tianjin, Shanghai and Chongqing) and 293 prefecture-level cities were established by the end of 2018. Among the 293 prefecture-level cities, 15 cities were granted the status of "sub-provincial cities". The administrative status of a sub-provincial city is above other regular prefecture-level cities but below the central-administered cities.

[2] The recent rapid economic growth over two decades caused the high car penetration in urban China. In 1995, just 1o million private cars were in China, but this number increased to over 200 million in 2018. However, owing to lagging data management and restricted official data disclosure, forming a complete dataset of road accidents at a city level is difficult. As numerous studies have emphasized, statistical methods and scales widely vary among different governmental institutions (e.g., Jiang et al., 2017). Overall, two major official sources of data on road accidents in China are existent. One is the data on traffic injury/death cases recorded by various hospitals and publicized by the health sector. The other is data on traffic accident cases recorded by the police department and published by the municipality. However, the former source lacks geoinformation (e.g., crash location). It does not record uninjured cases, and usually exclusively circulates information within the health sector. Although the latter source does not record unreported cases (in China, some people opt not to report the crash to the police but just to the insurance company if the parties involved arrive at a private agreement and the crash does not involve injuries/deaths) and the municipality merely publishes the number of cases that occurred in a certain year/month instead of specific details of each case. By contrast, the dataset from local insurance companies is the most suitable for our research because they record the geoinformation of each crash and is closest to the real number of automobile-involved crashes that occurred in a city in a certain year. Table A2 in Appendix B presents a comparison of the three data sources in terms of compositions and (dis)advantages.

insurance loss adjusters to record the detailed time and place without having to report the case to the transport department. Therefore, this dataset should come closest to the real number of crashes that occurred in Chengdu City Proper, including vehicle crashes with and without injuries. A total of 24,088 crash cases with the corresponding location and time that occurred in the Chengdu City Proper were recorded in 2018. We locate these cases on a map of Chengdu by using ArcGIS to calculate the crash incidence of each neighborhood. Crash incidence is measured by the ratio of the number of crashes to the size of the neighborhood.

POI data fetched from Baidu Map include eight categories: medical service, educational service, industrial land use, leisure facilities, living facilities, shopping and caterings, firms, and public sectors. This POI dataset has been used to measure indicators of mixed land use on the basis of entropy index method, which is widely adopted in built environment assessment (Cervero, 2003; Ewing et al., 2016b; Yue et al., 2017). Population density is measured by the ratio of the number of residents to the size of the neighborhood. Road density is measured by the ratio of the total length of a specific road type to the size of the neighborhood. According to China's official classification, urban roads are classified into expressways, arterial roads, secondary roads, and branch roads. We include the density of each road type as a variable into the model. Intersection density is measured by the ratio of the number of various intersections (three-leg and four-leg intersections) to the size of the neighborhood. The open street map (Baidu map) does not record roundabouts. Thus, this study does not include the density of roundabouts as a variable. The density of bus stops and sidewalks[3] are also calculated and included as variables in the model. Other variables, including signalization of intersections and bike lanes, are not included because in Chengdu and other Chinese megacities, almost every intersection are signalized, whereas bike lanes are rare. Moreover, the POI of Baidu map has not yet recorded bike lanes. Metro station density is also included as a variable. However, the literature on transport safety seldom mentions it as a factor of crash incidence because many studies have reported that accessibility to metro stations has a positive effect on reducing private car use at the neighborhood level (e.g., Li and Zhao, 2016). Hence, a high metro station density can be reasonably assumed to reduce the travel volume of an area and NACD accordingly. Given that we attempt to disentangle the mediating effects of exposure from those of risk, we include the traffic volume of each neighborhood as a variable. We calculate the average annual traffic volume of each neighborhood by using data from Didi Company, which is China's largest carpooling platform and has a research laboratory specializing in transport analysis. Road speeds and average transit times are calculated by relying on real-time satellite positioning data of massive vehicle volumes via the Didi platform.[4] This study uses the result of traffic volume index that combines the corresponding road width and length and a road occupancy rate derived by trajectories.[5] Table 3 presents the descriptive statistics of the variables used in the study.

We adopt SEM to examine the effects of built environment factors on crash incidence. Unlike the traditional regression approach, SEM enables the disentangling of mediators, mediating effects, and paths of effect. In brief, a mediator refers to the mediating variable between the independent and dependent variables; the former affects the latter by influencing the mediating variable. Mediating effects refer to such effects among the independent variable, the mediator, and the dependent variable. Note that various mediators may exist between the independent and dependent variables. Moreover, submediators may be present between the previous mediator and the dependent variable. Hence, the independent variable may affect the dependent variable via various paths by influencing different mediators and submediators. Path of effect refers to the total effect of a certain path within which the independent variable affects the dependent variable by influencing a certain (sub)mediator (for a discussion, see James and Brett, 1984). As shown in Fig. 2, we assume that (a) inner-city and suburban neighborhoods have different built environment characteristics, (b) variations in these built environment factors could cause variations in exposure (traffic volume) and risk, and (c) variations in exposure and risk produce variations in crash incidence. SEM enables the identification of whether (a), (b), and (c) indeed happened and the determination of the contribution of factors to each path. Exposure can be explicitly measured by traffic volume (Merlin et al., 2020). The risk of built environments at the neighborhood level is unobserved. Therefore, we regard the direct effect of built environment factors on crash incidence as the mediating effect of risk.

The following empirical analysis has three subsections. The first subsection presents the calculation of built environment characteristics at the neighborhood level and compares the differences in built environment factors between suburban and inner-city neighborhoods. We also present the results of an independent sample *t*-test of the disparities in built environment factors between inner-city and newly developed

[3] Note that sidewalk width can be a factor for crash density as observed by previous studies on the relationship between microscale street designs and individual-level crash incidence (Das and Abdel-Aty, 2010). However, the present study is a city-wide examination, and the assessment of built environments is based on the POI and graphic information extracted from the online map. Assessing the width of each city sidewalk is impractical and impossible. First, the online map does not have information on sidewalk width. Second, low-resolution satellite imagery does not support a machine learning approach to automatically identify sidewalks and compute their width.

[4] Didi Company not just provides shared mobility services but also has established a research institution called "Didi Research Institute" in 2016 that specializes on traffic demand prediction, real-time traffic volume monitoring/ estimation, and optimization of ridesharing match by using advanced computing approaches.

[5] The reliability of Didi online car-hailing service data in estimating traffic volume in China has been widely confirmed (Sun et al., 2018; Luo et al., 2019; Wen et al., 2019; Kuang et al., 2020). The method for estimating traffic volume used in this study combines the method developed by Sun et al. (2018) and Kuang et al. (2020). It has four main steps. The first step involves calculating the number of overall passenger cars of each road on the basis of estimated vehicle density. The second step requires computing the number of other vehicles according to the surveyed traffic composition. The hourly average speed of each road is then assessed by averaging the speeds of all Didi-hailed cars that passed through a road. The mean traffic flow of each road is also calculated. Lastly, the traffic volume of the whole city is estimated by applying the traffic volume inference model to the predicted average velocity and traffic density. Community-level traffic flow is calculated by averaging the traffic volume of roads in a community. Road-level floating car trajectory and speed data are collected between January 1, 2018 and December 31, 2018 to avoid seasonal variations. The entire dataset covers 1643 roads of various types. Traffic volume distribution is visualized at a disaggregated level (Appendix D). The formulas for calculating road traffic volume are as follows: If road $_i$ has two time slices, namely, $t1$ and $t2$, and the road's length is S, then the average speed v of the road is $v = 2{\cdot}S/(t1 + t2)$ during the period from $t1$ to $t2$. Traffic volume Q can be described as $Q = v \times k$, where v is velocity, and k is vehicle density. $S = \{\mathrm{Road\_Link}_1, \mathrm{Road\_Link}_2, \mathrm{Road\_Link}_3, \mathrm{Road\_Link}_4, \ldots\ldots \mathrm{Road\_Link}_5\}$, $\mathrm{v} = \frac{\sum_{i=1}^{N} L_i * \mathrm{W_i}}{\sum_{i=1}^{N} \frac{L_i}{V_i} * \mathrm{W_i}}$, where the size of S is N, the length of road is L_i, the weight of road is W_i, the freeflow of road is V_{free_i}, and the real-time speed of road is V_i. A logarithmic and an exponential model in traffic theory can represent the relationship of these parameters as follows: $\mathrm{v} = \mathrm{v}_c \ln(\frac{k_m}{k})$, $\mathrm{v} = \mathrm{v}_{free\_i} e^{(-\frac{k}{k_c})}$, where v_c is the critical velocity, and k_c is the critical density. When $\mathrm{k} < \mathrm{k}_c$, the traffic is free; when $\mathrm{k} > \mathrm{k}_c$, the traffic is crowded, $\mathrm{k} \in (0, \mathrm{k}_m)$. Through the two velocity–density relationship curves, the final formula can be substituted as $\mathrm{Q} = -\mathrm{v}\frac{k_m}{e}\ln(\frac{v}{v_{free\_i}})$.

Table 3
Descriptive statistics of variables.

Variables	Unit	Min	Max	Mean	Standard error
Crash density	frequency/km2	0.000	2300.07	52.33	157.74
Traffic volume index	index	0.143	1.95	1.22	1.43
Land use					
Population density	residents/km^2	0.000	86358	24333	17211.86
Mixed land use	entropy	0.000	0.958	0.575	0.249
Street design					
Road type					
Expressway density	meter/km^2	0.000	11225.54	779.90	1483.80
Arterial density	meter/km^2	0.000	13101.15	2377.57	2197.64
Secondary road density	meter/km^2	0.000	19019.69	6117.27	3171.40
Branch road density	meter/km^2	0.002	20221.09	5343.73	3498.80
Intersection form					
Three-leg intersection density	number/km^2	0.000	235.71	62.61	43.69
Four-leg intersection density	number/km^2	0.000	107.54	16.64	13.75
Pedestrian infrastructure					
Sidewalk density	meter/km^2	0.000	10303.26	448.50	1129.34
Others					
Bus stop density	number/km^2	0.000	49.52	12.86	9.85
Metro station density	number/km^2	0.000	8.000	0.212	1.04

suburban neighborhoods. A Point-biserial Pearson analysis is also conducted to examine the correlation between built environment characteristics and neighborhood types. In the second subsection, the results of SEM on the relationships between built environment factors and crash incidence are reported and explained. In the third subsection, the results of SEM on the mediating effects and effects of a path between neighborhood types and crash incidence are noted and discussed.

4. Empirical analysis

4.1. *Different built environment characteristics between suburban and inner-city neighborhoods*

As shown in Table 4, *t*-test results indicate substantial disparities between inner-city and newly developed suburban neighborhoods in terms of population density, mixed land use, expressway density, arterial density, branch road density, three-leg intersection density, sidewalk density, bus stop density, and metro station density. However, the disparity between the two neighborhood types in terms of secondary road density and four-leg intersection density are not significant. The results of simple Point-biserial Pearson analysis (Table 5) of the correlation between neighborhood types and built environment factors indicate that suburban neighborhoods are more likely to have a lower population density, expressway density, arterial density, branch road density, three-leg intersection density, sidewalk density, and bus stop density but higher mixed land use and metro station density than inner-city neighborhoods. The results of descriptive analysis (Table 6) also show that the average population density (19,782 residents/km^2), expressway density (693.25 m/km^2), arterial density (1,673.11 m/km^2), branch road density (4,189.26 m/km^2), three-leg intersection density (44.68 intersections/km^2), sidewalk density (161.74 m/km^2), and bus stop density (8.36 stops/km^2) of suburban neighborhoods are considerably lower than those of inner-city neighborhoods. The average mixed land use (0.584) and metro station density (0.380 stations/km^2) of suburban neighborhoods is substantially higher than those of inner-city neighborhoods. Figs. 3 and 4 present the disparities in these built environment characteristics among the neighborhoods in Chengdu City Proper. A salient suburb–inner city discrepancy can be discerned.

Overall, similar to American urban sprawls, Chengdu's suburban built environments at the neighborhood level are characterized by low population density and low-accessibility road networks (low density of different road types). However, Chengdu's suburban built environments differ from American-style urban sprawls by having higher mixed land use and metro station density than inner-city suburbs. Moreover, Fig. 4 shows that suburban neighborhoods have lower traffic volume than inner-city neighborhoods, unlike in American suburban neighborhoods where people rely heavily on private cars, resulting in high traffic volume in suburban areas. Aside from lower population and road densities, another substantial contributor to the lower traffic volume in Chengdu's suburban neighborhoods is Chengdu's municipality-led heavy

Table 4
Results of *t*-test of the disparities in built environment factors between inner-city and suburban neighborhoods in Chengdu City Proper.

Built environment	*t*	Sig.		*t*	Sig.
Population density	4.925	0.000	Three-leg intersection density	6.381	0.000
Mixed land use	3.485	0.017	Four-leg intersection density	0.659	0.511
Expressway density	4.788	0.000	Sidewalk density	5.275	0.000
Arterial density	5.437	0.000	Bus stop density	7.424	0.000
Secondary road density	−0.605	0.545	Metro station density	−2.543	0.011
Branch road density	4.807	0.000			

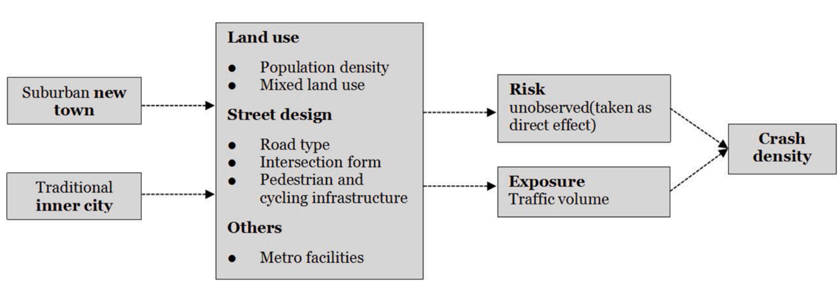

Fig. 2. Conceptual/Analytical framework.

Table 5
Point-biserial Pearson analysis of the correlation between neighborhood types and built environment factors in Chengdu City Proper.

Built environment	Point-biserial correlation coefficient	Sig.		Point-biserial correlation coefficient	Sig.
Population density	−0.178	0.000	Three-leg intersection density	−0.238	0.000
Mixed land use	0.063	0.013	Four-leg intersection density	−0.026	0.517
Expressway density	−0.034	0.005	Sidewalk density	−0.147	0.000
Arterial density	−0.186	0.000	Bus stop density	0.264	0.000
Secondary road density	0.025	0.537	Metro station density	0.180	0.047
Branch road density	−0.191	0.000			

investment into metro construction in new suburban towns. China's post-2008 suburbanization is characterized by state-led built environment designs and construction, a situation that probably resulted in sociospatial consequences different from those of American-style urban sprawls. Fig. 3 illustrates that suburban neighborhoods generally have a lower crash incidence than inner suburban neighborhoods (Dumbaugh and Rae, 2009; Ewing et al., 2016a). In the following sections, SEM is used to examine whether different built environment characteristics lead to low crash incidence in suburban neighborhoods.

4.2. Built environment factors and neighborhood-level crash incidence

Table A1 in Appendix A illustrates the goodness-of-fit statistics for SEM. All statistics, including the root mean square error of approximation, confirmatory fit index (CFI), Tucker–Lewis index, and weighted root mean square residual, indicate a good fit for the model. Tables 6 and 7 present the results of SEM, specifically the estimated relationships (association, effect size, and path of effect) among built environment factors and crash incidence at the neighborhood level. Results indicate a strong effect of built environments on crash incidence. However, the effect size and contribution path vary among different built environment factors. First, traffic volume is positively associated with crash incidence (*0.193****), consistent with the classic account that a high traffic volume determining a high exposure of people to traffic increases the crash incidence of a locality (Merlin et al., 2020). This finding also supports traffic volume as a key meditator on the relationship between exogenous factors and crash incidence (Merlin et al., 2020).

With regard to land use characteristics, population density has a substantial and positive association with crash incidence. This causality is partially attributed to increased traffic volume as a result of high population density (*0.012****). The direct effect of population density on crash incidence is more remarkable (*0.296****) and suggests that a high population density can substantially increase crash incidence in a neighborhood by increasing the risk of local traffic. This trend manifests because a high population density can increase and complicate the interactions of different traffic flows (e.g., vehicle flow, nonvehicle flow, and pedestrian flow), leading to a high possibility of crashes. Moreover, the mixed land use of a neighborhood affects crash incidence. On the one hand, a high mixed land use can attract more inflow of population/goods and generate more traffic in a locality, thus increasing exposure to traffic (*0.014**). On the other hand, mixed land use decreases crash incidence by reducing traffic risk (*−0.012****). An explanation for this decrease is that a high mixed land use can reduce traffic speed and heighten drivers'/pedestrians' sensitivity to their surroundings due to diverse street landscapes. Overall, mixed land use substantially reduces crash incidence at the neighborhood level as the negative direct effect is considerably larger than the positive indirect effect.

A neighborhood with a high expressway density and arterial density tends to have a high traffic volume, thereby leading to a high crash incidence (*0.032**** and *0.014**). However, expressway density reduces crash incidence by lowering traffic risk (*−0.084***), whereas arterial density has an adverse effect (*0.148****). Such discrepancy is attributed

Table 6
Descriptive statistics of SEM variables.

Variables	Central areas (N = 460)			Suburban new towns (N = 154)		
	Min	Mean	Max	Min	Mean	Max
Crash incidence	0.000	59.09	2300.07	0.000	32.15	705.68
Traffic volume	1.27	1.45	1.95	1.21	1.35	1.67
Land use						
Population density	0.000	26191	86358	0.000	19782	70548
Mixed land use	0.000	0.548	0.958	0.000	0.584	0.917
Street design						
Road type						
Expressway density	0.000	808.38	8022.00	0.000	693.25	11226.00
Arterial density	0.000	2613.41	13101.15	0.000	1673.11	9254.600
Secondary road density	0.000	6071.53	19019.69	0.000	6253.92	13954.092
Branch road density	0.000	5727.63	20221.09	0.000	4189.26	17530.80
Intersection form						
Three-leg intersection density	0.000	68.61	222.25	1.33	44.68	235.71
Four-leg intersection density	0.000	16.85	107.54	0.000	16.02	61.50
Pedestrian infrastructure						
Sidewalk density	0.000	544.496	10303.00	0.000	161.74	3740.00
Others						
Bus stop density	0.000	14.37	49.52	0.000	8.36	37.74
Metro station density	0.000	0.188	8.000	0.000	0.380	5.405

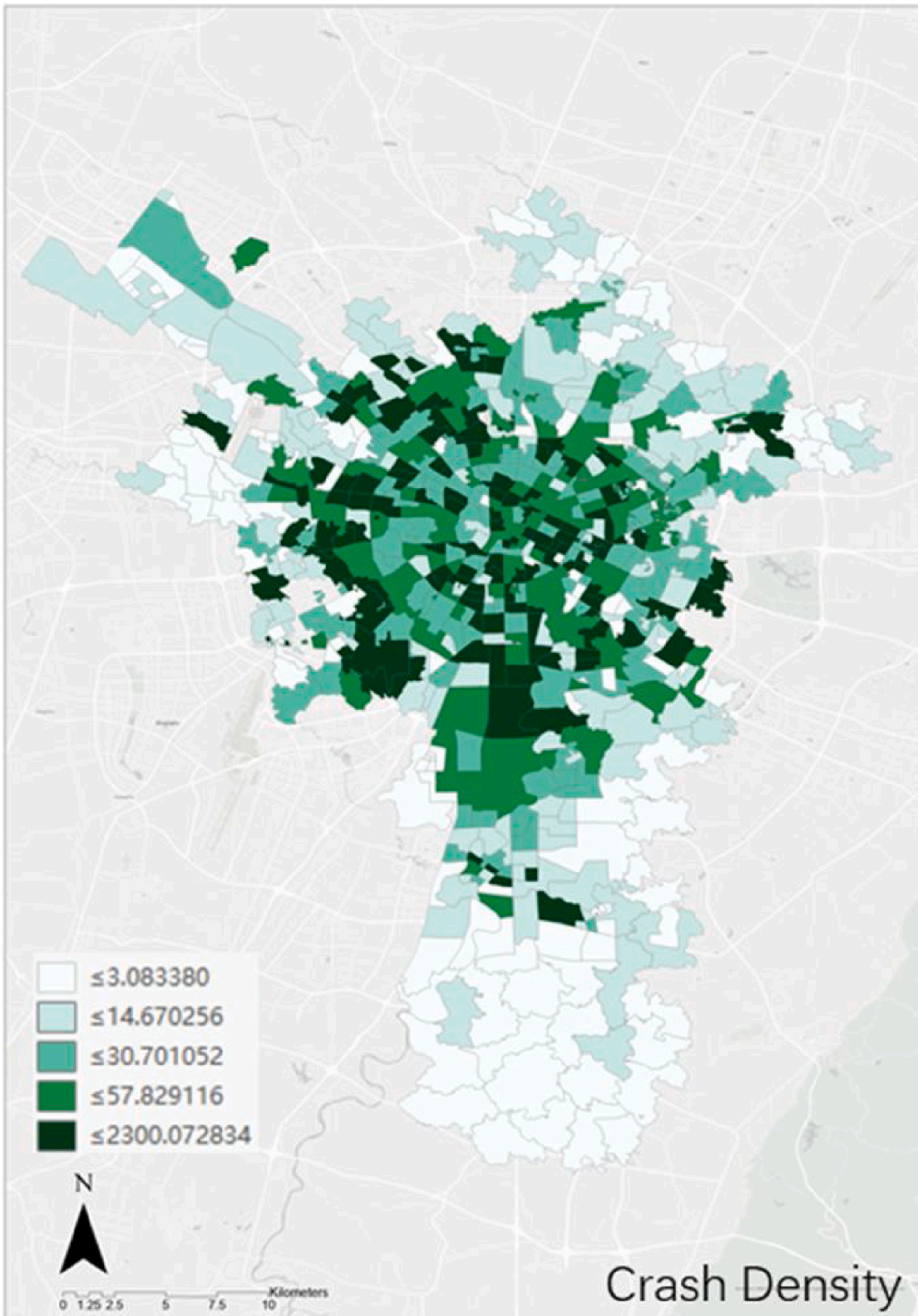

Fig. 3. Crash density at the neighborhood level in Chengdu City Proper.

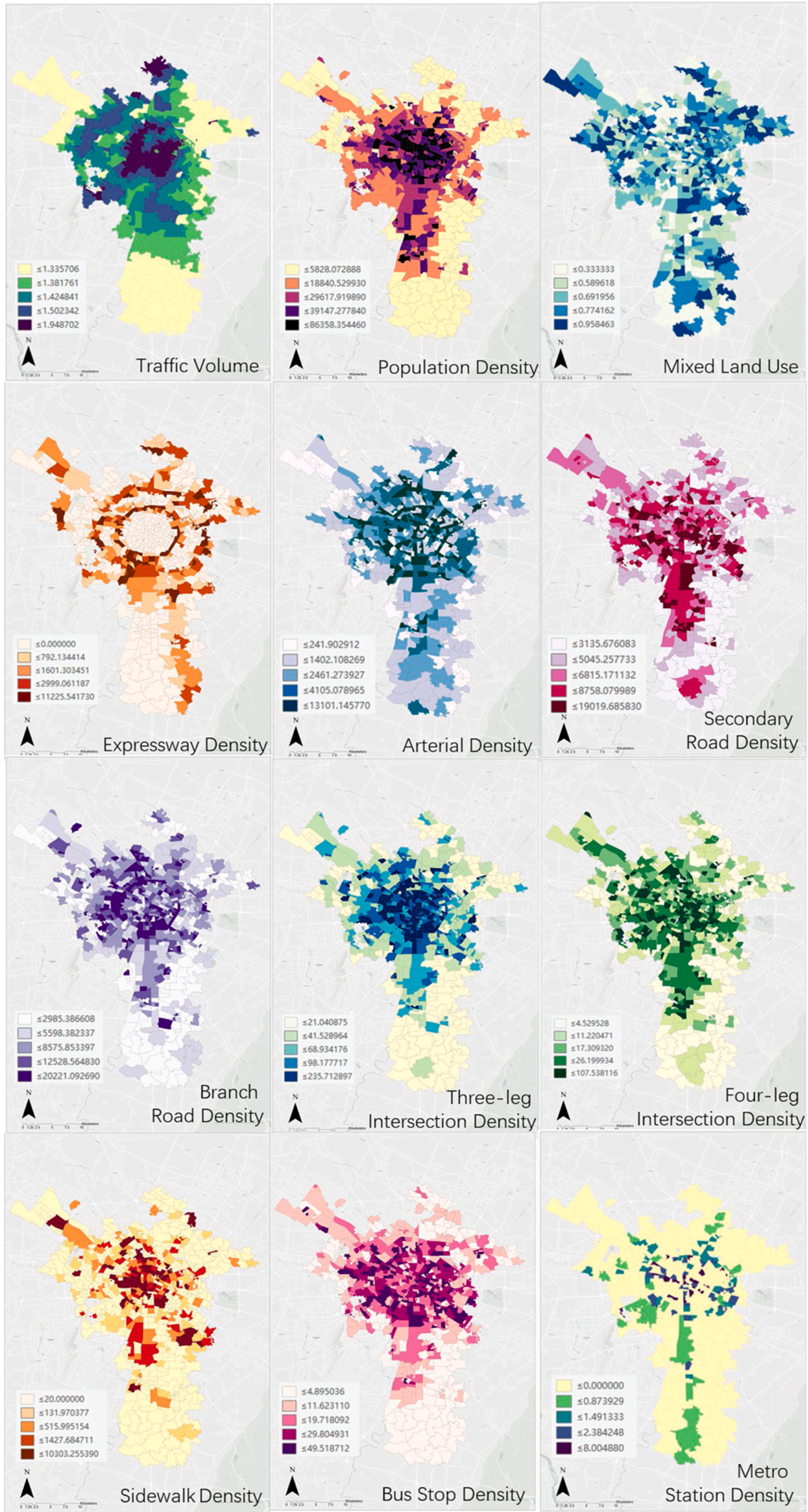

Fig. 4. Built environment characteristics at the neighborhood level of Chengdu City Proper.

Table 7
Direct and indirect effects of variables on crash density.

Variables	Direct effect	Indirect effect (traffic volume as mediator)
Traffic volume	0.193***	
Land use characteristic		
Population density	0.296***	0.012***
Mixed land use	−0.012***	0.014*
Street design characteristic		
Expressway density	−0.084**	0.032***
Arterial density	0.148***	0.014*
Secondary road density	−0.002	−0.002
Branch road density	0.038	0.009
Three-leg intersection density	−0.039	−0.011***
Four-leg intersection density	−0.017	0.069
Sidewalk density	0.085**	0.050
Bus stop density	0.177***	0.003
Metro station density	−0.020	−0.024***
Goodness-of-fit measures	Chi-square test of model fit: Chi-square = 3.374 degrees of freedom = 4 P value = 0.497 RMSER = 0.031 CFI = 0.999 AGFI = 0.981 NFI = 0.995 Standardized RMR = 0.077	

Note: *** significant at 0.01; ** significant at 0.05 * significant at 0.1; effect is standardized.

to the possibility that expressways are fenced and thus isolated from other vehicles, nonvehicles, and pedestrian inflow. Such isolation can markedly reduce the interactions of different traffic flows (type and speed) and thus reduce traffic risk. By contrast, arterial roads are usually the places where different traffic flows intensively interact and have a high traffic speed (compared with secondary/branch road), thereby increasing traffic risk and crashes. Secondary and branch road densities are not found to have a substantial effect on crash incidence.

Three-leg intersection density is found to have a considerable and negative indirect effect on crash incidence with traffic volume as a meditator (*−0.011****). This result suggests that a neighborhood with more three-leg intersections tends to have few crashes as a result of decreased traffic volume. A possible explanation for this observation is that a three-leg intersection reduces the connectivity and accessibility of local road networks, making the neighborhood less attractive for traffic and vehicles. A four-leg intersection density is not found to affect crash incidence.

Sidewalk density and bus stop density are found to have a substantial and positive direct effect on crash incidence (*0.085*** and *0.177****), suggesting that a neighborhood with more sidewalks and bus stops tends to have more crashes as a result of increased traffic risks. Many field experiment-based studies have found that bus stops and sidewalks complicate the interaction of road traffic and particularly induce emergency encountered by drivers, leading to a high crash incidence (Wei and Lovegrove, 2013; Yu, 2015; Cai et al., 2017). Therefore, our results can be supported and explained by their argument. Lastly, metro

Table 8
Estimates of SEM.

Variables	Traffic volume		Crash incidence	
	Estimates	S.E	Estimates	S.E
Traffic volume			4.058***	0.868
Land use				
Population density	0.002	0.001	0.178***	0.028
Mixed land use	0.010**	0.004	0.034***	0.095
Street design				
Road type				
Expressway density	0.004***	0.001	−0.039**	0.016
Arterial density	0.002*	0.001	0.081***	0.019
Secondary road density	−0.001	0.004	−0.004	0.087
Branch road density	0.003	0.003	0.049	0.054
Intersection form				
Three-leg intersection density	−0.031***	0.006	−0.070	0.126
Four-leg intersection density	−0.004	0.003	−0.026	0.075
Pedestrian infrastructure				
Sidewalk density	0.003	0.001	0.020**	0.009
Others				
Bus stop density	0.001	0.003	0.273***	0.069
Metro station density	−0.011***	0.001	−0.040	0.066

Note: *** significant at 0.01; ** significant at 0.05 * significant at 0.1.

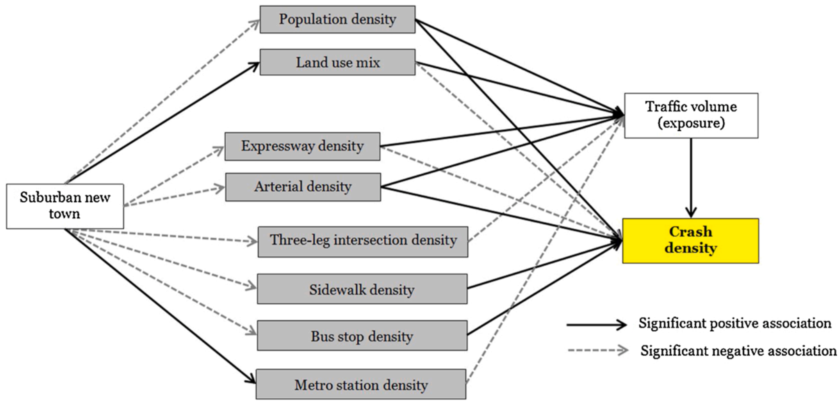

Fig. 5. Summary of estimated relationships among neighborhood types, built environments, and crash incidence (source: the authors).

station density is found to negatively associate with NACD with traffic volume as the mediator (−*0.024****), thus supporting our previous assumption that a high metro station density can reduce travel volume and thus reduce crash incidence in a neighborhood.

Overall, SEM results suggest that neighborhoods with high population density, expressway and arterial density, and sidewalk and bus stop density tend to have a high crash incidence. By contrast, neighborhoods with high mixed land use and metro station density tend to have a low crash incidence. In this regard, expressways, arterials, and bus stops produce an unsafe mobility environment for residents despite enhancing the accessibility of a neighborhood to public services and jobs.

4.3. New suburban towns, built environments, and crash density

SEM results indicate a significant and positive association (−*0.272***) between new suburban towns and crash incidence at the neighborhood level. As illustrated by Table 8 and summarized in Fig. 5, suburban neighborhoods tend to have a considerably lower crash incidence than inner-city neighborhoods because of certain built environment characteristics. According to path analysis on mediation (Table 9), the tendency of having less population density, expressway density, arterial density, sidewalk density, bus stop density, and metro station density reduces crash incidence in suburban neighborhoods, among which the mediating effect of population density is the largest (−*0.133****). By contrast, the mediating effect of metro station density is the smallest (−*0.003****). The tendency of having less three-leg intersection density and expressway density increases crash incidence in suburban neighborhoods (*0.009*** and *0.003*****, respectively). Moreover, the total mediating effect of less traffic volume on less crash incidence (−*0.003***) in suburban neighborhoods is substantially smaller than that of less traffic risk (−*0.269***). This

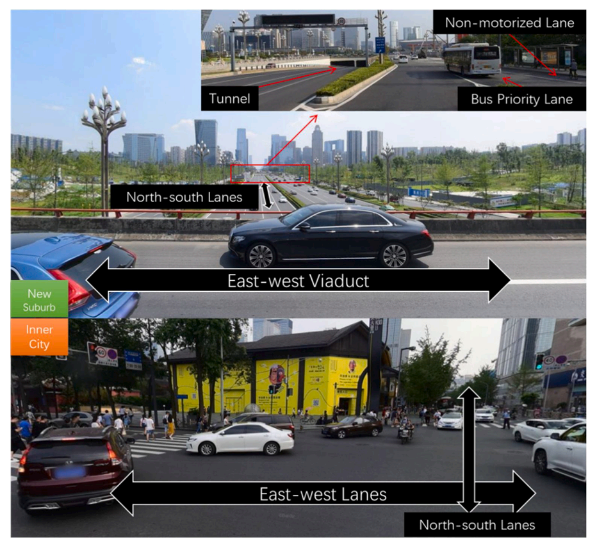

Fig. 6. Comparison of road systems in new suburban and inner-city neighborhoods of Chengdu (Source: Baidu Street Map).

Table 9
Effects of being suburban on crash density.

Meditator	Via exposure	Via risk	Standardized total effect
Land use characteristic			
Population density	−0.005	−0.128	−0.133
mixed land use	0.001	−0.016	−0.015
Street design characteristic			
Expressway density	−0.002	0.004	0.003
Arterial density	−0.002	−0.024	−0.026
Secondary road density	null	null	null
Branch road density	null	null	null
Three-leg intersection density	0.009	null	0.009
Four-leg intersection density	null	null	null
Sidewalk density	null	−0.021	−0.021
Bus stop density	null	−0.085	−0.085
Metro station density	−0.003	null	−0.003
Total effect	−0.003	−0.269	−0.272

finding suggests that suburban neighborhoods tend to have a higher traffic safety than inner-city neighborhoods mainly because of the less-risky built environment instead of less exposure of residents to traffic.

5. Conclusion

This study addresses three critical issues of state-led suburbanization, built environments, and neighborhood-level crash incidence in China: (1) differences in built environment characteristics between suburban and inner-city neighborhoods in the context of China's post-2008 rapid suburbanization, (2) associations between built environment factors and crash incidence at the neighborhood level in China, and (3) reasons suburban neighborhoods tend to have lesser crash incidence than inner-city neighborhoods in China from the perspective of built environment characteristics. The city proper of Chengdu, the largest metropolis in southwestern China, is selected as a case study, and SEM is adopted to a dataset formed mainly by using a big data approach. This study quantifies the causality among neighborhood types (suburban versus inner city), built environment factors, and NACD. Three important conclusions can be drawn from the modeling results.

First, compared with inner-city neighborhoods, suburban neighborhoods in Chengdu tend to have less population density, road density (all different types), intersection density of different types, sidewalk density, bus stop density, and metro station density but a high land use mix. Traffic volume in suburban neighborhoods also tends to be lower than that in inner-city neighborhoods. These features are due to two reasons. First, new suburbs advocate an integration between industrial development and urbanization initiatives to reduce the commuting needs of residents. The second reason is the construction of public transport infrastructures so that travelers do not need to rely heavily on private cars for cross-regional mobility. This suburbanization process is different from American-style suburban sprawl process, which is characterized by heavy reliance on automobiles and interstate highways. By contrast, American suburbs have a larger population and higher traffic volume than inner city (Hamidi and Ewing, 2014). Thus, our findings support the narrative that China's post-2008 suburbanization differs from American-style, market-led urban sprawl process in terms of strong state interventions in promoting public transit-oriented (metro line) suburban developments and mixed land use.

Moreover, consistent with the findings in American cities (Ukkusuri et al., 2012; Wier et al., 2009; Dumbaugh and Rae, 2009; Dumbaugh and Li, 2010; Huang et al., 2010; Quistberg et al., 2015), built environment factors are found to have remarkable effects on crash incidence at the neighborhood level in Chinese cities. Our findings also support the premise that mediating effects exist among built environment factors and crash incidence (Ewing and Dumbaugh, 2009; Merlin et al., 2020). Our modeling results reveal that high population density and arterial density increase crash incidence by compounding traffic volume (increasing exposure) and increasing traffic risk. High sidewalk density and bus stop density increase crash incidence by raising traffic risk. Furthermore, three-leg intersection density decreases crash incidence by reducing the traffic volume of a neighborhood. However, expressway density and mixed land use have a contradictory effect on crash incidence by increasing the traffic volume of a neighborhood while decreasing traffic risk. Moreover, the negative effect of mixed land use and expressway density in the city proper of Chengdu on crash incidence is considerably stronger than their positive effect. Thus, neighborhoods with high mixed land use and expressway density tend to have a lower crash incidence.

Furthermore, suburban neighborhoods in Chengdu tend to have lower crash incidence than inner-city neighborhoods by 27.2 % primarily because of the tendency of suburban built environment characteristics to lower traffic risk instead of to lower traffic volume. Compared with alleys that are generally narrow in inner-city neighborhoods but have to accommodate high traffic volume, the unified planning and construction of arterial road networks and slender branches in suburbs have substantially reduced the risk of both humans and vehicles occupying available road spaces (Fig. 6). Although this design for promoting a more pedestrian-friendly city is controversial from the perspective of people-oriented urban concept (Henderson, 2013), its potential in creating a safer transport environment has to be recognized. In this regard, unlike urban sprawls, which are a risk factor in car incidence and pedestrian fatalities in American suburban neighborhoods (Ewing and Dumbaugh, 2009; Ewing and Hamidi, 2015), China's suburban neighborhoods developed by state-led suburbanization initiative tend to have better traffic safety than inner-city neighborhoods. Less population density and bus stop density are the major contributors to the better road safety of suburban neighborhoods.

Our findings have three critical implications to policy formulation on traffic safety concerns in urban land (re)development. First, transit-oriented development should be pursued given its merits in producing a safe suburban built environment. As demonstrated by the city proper of Chengdu, state-led proactive investments into metro line constructions have primarily mitigated the demand for road facilities (expressways and arterials), curbed traffic volume increase, and reduced crash incidence in new suburban towns. By contrast, highway-oriented American-style urban sprawls produce unsafe suburban neighborhoods and dangerous living conditions for children (Dumbaugh and Rae, 2009). Second, secondary roads and branches should be prioritized if the accessibility of a neighborhood is sought to be improved without exacerbating local traffic/street safety. This implication is particularly relevant in redeveloping aging neighborhoods and neighborhoods with numerous children. Third, bus stops and sidewalks should be designed in such a way that they prevent complicating the interaction of traffic flow and thus increasing traffic risks.

The limitations of this study should be acknowledged. First, we use a dataset developed from a single city. Hence, our results may not be applicable to other Chinese cities. The relationship (association, mediator, and effect size) between built environments and crash incidence can be complicated by sociocultural conditions that vary among cities (e.g., local people's traffic safety awareness) and efficiency of traffic management. Therefore, a national-level analysis and comparative studies among cities are necessary to improve our understanding of the relationships among suburbanization, built environments, and NACD. Second, we do not make a classification of crashes but reckon crash incidence at a comprehensive level. Hence, the causality between built environment factors and NACD might have been overgeneralized. The sensitivity of different crash types (e.g., pedestrian-involved crash, bicyclist-involved crash, and crashes involving only automobiles) to built environment factors can vary (Merlin et al., 2020). However, the present study cannot consider crash types because of the lack of data. Third, the traffic safety of a neighborhood should be further assessed according to fatality rates and severity of injury suffered by drivers and pedestrians (Guerra et al., 2019; Dumbaugh and Rae, 2009). However, the present study cannot consider these factors because of the lack of relevant data. Lastly, the demographic composition of a neighborhood has an effect on crash incidence (Guerra et al., 2019). In China, population census at the neighborhood level is conducted every decade. Thus, data on the demographic composition of newly developed suburban neighborhoods are lacking.[6] Moreover, the flows of other trip modes (e.g., bicycling and electric biking) in a neighborhood can have an effect on NACD. However, these traffic flows are currently out of monitoring,

[6] The most recent census at the neighborhood level was conducted in 2010 when new suburban towns had not been developed yet. Moreover, the recent decade has witnessed great population changes in the inner-city neighborhoods of Chengdu City Proper. Thus, applying the old demographic data to this research is unreasonable. A new census will be conducted this year, and it will take one or two years before the dataset is published. Owing to the inaccessibility to reliable demographic data, and because the focus of this study is on built environment factors, we did not include demographic composition as a variable in the model.

counting, and estimation either by the government or by technology companies in China. We will improve our model by reflecting various demographic compositions and flows of different trip modes as soon as relevant data become available in the future.

CRediT authorship contribution statement

Si Qiao: Conceptualization, Data curation, Methodology, Software, Writing - original draft. **Anthony Gar-On Yeh:** Supervision, Writing - review & editing. **Mengzhu Zhang:** Conceptualization, Methodology, Software, Writing - original draft. **Xiang Yan:** Data curation.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgement

This research is funded by the Chan To-Haan Endowed Professorship Fund of the University of Hong Kong.

Appendix A

Table A1
Goodness-of-fit statistics of SEM.

Goodness-of-fit measures	**Reference value**	Value of our model
Chi-Square/degree of freedom	**< 3.00**	0.843
RMSEA	**<0.10**	0.031
AGFI	**>0.90**	0.981
NNFI	**>0.09**	0.210
CFI	**>0.09**	0.999
Standardized RMR	**<0.08**	0.077

Appendix B

Table A2
Comparison of three sources of data on traffic crashes in China.

Source	Data	Subject of declaration	Composition#	(dis)advantage
Health sector	Injuries and deaths caused by crashes	Patients in the hospital	Excludes crashes with no injuries	• Has details on injury severity • Has no geoinformation • Details on each case are not publicized
Police sector	Crash cases	Interested parties of a crash	Excludes majority of crashes without injuries and property damages	• Has details on injury severity • Records geoinformation • Details on each case are not publicized
Insurance companies	Crash cases with a claim for compensation	Car drivers involved in a crash	Excludes minority of crashes with no/ small property damages*	• Records geoinformation • Details on injury severity and personal information of the driver/pedestrian/passenger are confidential

Note: * In China, property damage due to car crash is small in some cases. The car driver/owner may opt not to claim a compensation because records will increase the insurance premium in the next year. # Overall, all three sources include crashes involving injuries/deaths. The main difference is on records of uninjured cases.

Appendix C

See Table A3.

Table A3
Severity statistics of traffic crash data in 2018.

	Number	Ratio
Local Insurance Company Dataset		
Chengdu City Proper (study area, 657 km2)		
Crash cases used in this study	24,088	100%
Involving death/injury	1,606	6.67%
Involving no death/injury	22,482	93.33%
Involving minor crash	20,820	92.61%
Involving moderate crash	1,628	7.24%
Involving serious crash	35	0.15%
Whole Chengdu City (14,335 km2)		
Crash cases	128,717	
Involving death/injury	6,940	
Involving no death/injury	121,777	
Government Statistical Report		
Whole Chengdu City (14,335 km2)		
Crash cases	1,805	
Number of injuries and deaths	2,001	

Note: * Date source is the *Chengdu Statistical Year Book*: http://www.cdstats.chengdu.gov.cn/htm/detail_145407.html.

Appendix D

See Fig. D1.

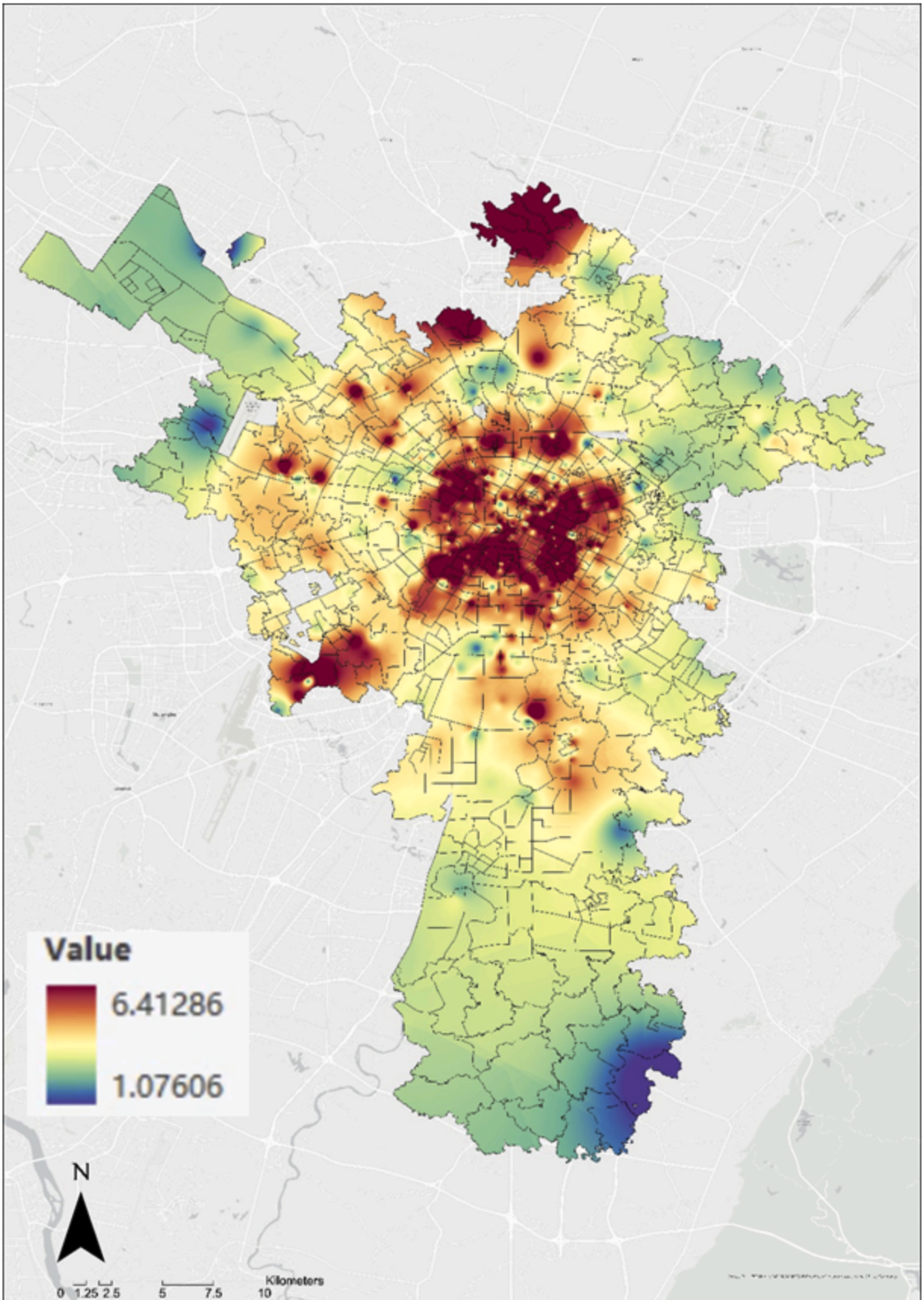

Fig. D1. Visualization of traffic volume distribution at a disaggregated level.

References

Bao, J., Liu, P., Yu, H., et al., 2017. Incorporating twitter-based human activity information in spatial analysis of crashes in urban areas. Accid. Anal. Prev. 106, 358–369.

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

Cervero, R., 2003. Growing smart by linking transportation and land use: perspectives from California. Built Environ. 29 (1), 66–78.

Chan, K.W., 2007. Misconceptions and complexities in the study of China's cities: definitions, statistics, and implications. Euras. Geogr. Econ. 48 (4), 383–412.

Chang, C., Lu, M., 2017. Misery of new town——density, distance and debt. China Econ. Q. 16 (4), 1621–1642 (in Chinese).

Chen, P., 2015. Built environment factors in explaining the automobile-involved bicycle crash frequencies: a spatial statistic approach. Saf. Sci. 79, 336–343.

Chen, D.A., Fuller, D., 2014. Analyzing road surface conditions, collision time, and road structural factors associated with bicycle collisions from 2000 to 2010 in Saskatoon, Saskatchewan. J. Transp. Health 1, 40–44.

Chen, P., Shen, Q., 2016. Built environment effects on cyclist injury severity in automobile-involved bicycle crashes. Accid. Anal. Prev. 86, 239–246.

Chen, C., Zhang, G., Tarefder, R., et al., 2015. A multinomial logit model-Bayesian network hybrid approach for driver injury severity analyses in rear-end crashes. Accid. Anal. Prev. 80, 76–88.

Cheng, L., Geedipally, S.R., Lord, D., 2013. The Poisson–Weibull generalized linear model for analyzing motor vehicle crash data. Saf. Sci. 54, 38–42.

Chu, W., Wu, C., Atombo, C., et al., 2019. Traffic climate, driver behaviour, and accidents involvement in China. Accid. Anal. Prev. 122, 119–126.

Das, A., Abdel-Aty, M., 2010. A genetic programming approach to explore the crash severity on multi-lane roads. Accid. Anal. Prev. 42 (2), 548–557.

Ding, C., Chen, P., Jiao, J., 2018. Non-linear effects of the built environment on automobile-involved pedestrian crash frequency: a machine learning approach. Accid. Anal. Prev. 112, 116–126.

Dumbaugh, E., Li, W., 2010. Designing for the safety of pedestrians, cyclists, and motorists in urban environments. J. Am. Plan. Assoc. 77, 69–88.

Dumbaugh, E., Rae, R., 2009. Safe urban form: revisiting the relationship between community design and traffic safety. J. Am. Plan. Assoc. 75, 309–329.

El-Basyouny, K., Sayed, T., 2009. Collision prediction models using multivariate Poisson-lognormal regression. Accid. Anal. Prev. 41, 820–828.

Ewing, R., 1997. Is Los Angeles-style sprawl desirable? J. Am. Plan. Assoc. 63, 107–126.

Ewing, R., Dumbaugh, E., 2009. The built environment and traffic safety: a review of empirical evidence. J. Plan. Lit. 23, 347–367.

Ewing, R., Hamidi, S., 2015. Urban sprawl as a risk factor in motor vehicle occupant and pedestrian fatalities: update and refinement. Transp. Res. Rec. 2513, 40–47.

Ewing, R., Schieber, R.A., Zegeer, C.V., 2003. Urban sprawl as a risk factor in motor vehicle occupant and pedestrian fatalities. Am. J. Publ. Health 93, 1541–1545.

Ewing, R., Hamidi, S., Grace, J.B., 2016a. Urban sprawl as a risk factor in motor vehicle crashes. Urban Stud. 53, 247–266.

Ewing, R., Hamidi, S., Grace, J.B., et al., 2016b. Does urban sprawl hold down upward mobility? Landsc. Urban Plan. 148, 80–88.

Fischer, K., Sternfeld, I., Melnick, D.S., 2013. Impact of population density on collision rates in a rapidly developing rural, exurban area of Los Angeles County. Inj. Prev. 19, 85–91.

Guerra, E., Dong, X., Kondo, M., 2019. Do denser neighborhoods have safer streets? Population density and traffic safety in the Philadelphia region. J. Plan. Educ. Res. 0739456X19845043.

Hadayeghi, A., Shalaby, A.S., Persaud, B.N., et al., 2006. Temporal transferability and updating of zonal level accident prediction models. Accid. Anal. Prev. 38, 579–589.

Hamann, C., Peek-Asa, C., 2013. On-road bicycle facilities and bicycle crashes in Iowa, 2007–2010. Accid. Anal. Prev. 56, 103–109.

Hamidi, S., Ewing, R., 2014. A longitudinal study of changes in urban sprawl between 2000 and 2010 in the United States. Landsc. Urban Plan. 128, 72–82.

Henderson, J., 2013. Street Fight: The Politics of Mobility in San Francisco. University of Massachusetts Press, Amherst, MA.

Huang, H., Abdel-Aty, M.A., Darwiche, A.L., 2010. County-level crash risk analysis in Florida: Bayesian spatial modeling. Transp. Res. Rec. 2148, 27–37.

Jiang, B., Liang, S., Peng, Z.-R., et al., 2017. Transport and public health in China: the road to a healthy future. Lancet 390, 1781–1791.

Kuang, L., Hua, C., Wu, J., Yin, Y., Gao, H., 2020. Traffic volume prediction based on multi-sources gps trajectory data by temporal convolutional network. Mob. Netw. Appl. 1–13.

Lee, J.S., Zegras, P.C., Ben-Joseph, E., 2013. Safely active mobility for urban baby boomers: the role of neighborhood design. Accid. Anal. Prev. 61, 153–166.

Lee, J., Nam, B., Abdel-Aty, M., 2015. Effects of pavement surface conditions on traffic crash severity. J. Transp. Eng. 141, 04015020.

Lee, J., Abdel-Aty, M., Cai, Q., Wang, L., Huang, H., 2018. Integrated modeling approach for non-motorized mode trips and fatal crashes in the framework of transportation safety planning. Landsc. Urban Plan. 2672 (32), 49–60.

Lee, J., Abdel-Aty, M., Shah, I., 2019a. Evaluation of surrogate measures for pedestrian trips at intersections and crash modeling. Accid. Anal. Prev. 130, 91–98.

Lee, J., Abdel-Aty, M., Huang, H., Cai, Q., 2019b. Transportation safety planning approach for pedestrians: an integrated framework of modeling walking duration and pedestrian fatalities. Transp. Res. Rec. 2673 (4), 898–906.

Lucas, K., 2012. Transport and social exclusion: where are we now? Transp. Policy 20, 105–113.

Luo, X., Liu, B., Jin, P.J., Cao, Y., Hu, W., 2019. Arterial traffic flow estimation based on vehicle-to-cloud vehicle trajectory data considering multi-intersection interaction and coordination. Transp. Res. Rec. 2673 (6), 68–83.

Lyu, N., Cao, Y., Wu, C., et al., 2018. The effect of gender, occupation and experience on behavior while driving on a freeway deceleration lane based on field operational test data. Accid. Anal. Prev. 121, 82–93.

Marshall, W.E., Garrick, N.W., 2011. Does street network design affect traffic safety? Accid. Anal. Prev. 43, 769–781.

Merlin, L.A., Guerra, E., Dumbaugh, E., 2020. Crash risk, crash exposure, and the built environment: a conceptual review. Accid. Anal. Prev. 134, 105244.

Noland, R.B., 2003. Traffic fatalities and injuries: the effect of changes in infrastructure and other trends. Accid. Anal. Prev. 35, 599–611.

Noland, R.B., Quddus, M.A., 2004. A spatially disaggregate analysis of road casualties in England. Accid. Anal. Prev. 36, 973–984.

Osama, A., Sayed, T., 2016. Evaluating the impact of bike network indicators on cyclist safety using macro-level collision prediction models. Accid. Anal. Prev. 97, 28–37.

Prati, G., Pietrantoni, L., Fraboni, F., 2017. Using data mining techniques to predict the severity of bicycle crashes. Accid. Anal. Prev. 101, 44–54.

Pulugurtha, S.S., Nambisan, S.S., Maheshwari, P., 2006. Estimating pedestrian counts in urban areas for transportation planning and safety analyses. Proceedings of 9th International Conference on Applications of Advanced Technology in Transportation. Chicago, Illinois, August 13–15, 2006 257–262.

Pulugurtha, S.S., Duddu, V.R., Kotagiri, Y., 2013. Traffic analysis zone level crash estimation models based on land use characteristics. Accid. Anal. Prev. 50, 678–687.

Qu, W., Xu, J., Ge, Y., et al., 2019. Development and validation of a questionnaire to assess public receptivity toward autonomous vehicles and its relation with the traffic safety climate in China. Accid. Anal. Prev. 128, 78–86.

Quddus, M.A., 2008. Modelling area-wide count outcomes with spatial correlation and heterogeneity: an analysis of London crash data. Accid. Anal. Prev. 40, 1486–1497.

Quistberg, D.A., Howard, E.J., Ebel, B.E., et al., 2015. Multilevel models for evaluating the risk of pedestrian–motor vehicle collisions at intersections and mid-blocks. Accid. Anal. Prev. 84, 99–111.

Regev, S., Rolison, J.J., Moutari, S., 2018. Crash risk by driver age, gender, and time of day using a new exposure methodology. J. Saf. Res. 66, 131–140.

Rolison, J.J., Regev, S., Moutari, S., et al., 2018. What are the factors that contribute to road accidents? An assessment of law enforcement views, ordinary drivers' opinions, and road accident records. Accid. Anal. Prev. 115, 11–24.

Sun, D.J., Zhang, K., Shen, S., 2018. Analyzing spatiotemporal traffic line source emissions based on massive didi online car-hailing service data. Transp. Res. D Transp. Environ. 62, 699–714.

Trowbridge, M.J., McDonald, N.C., 2008. Urban sprawl and miles driven daily by teenagers in the United States. Am. J. Prev. Med. 34, 202–206.

Ukkusuri, S., Miranda-Moreno, L.F., Ramadurai, G., et al., 2012. The role of built environment on pedestrian crash frequency. Saf. Sci. 50, 1141–1151.

Wang, Y., Li, L., Prato, C.G., 2019. The relation between working conditions, aberrant driving behaviour and crash propensity among taxi drivers in China. Accid. Anal. Prev. 126, 17–24.

Wei, F., Lovegrove, G., 2013. An empirical tool to evaluate the safety of cyclists: community based, macro-level collision prediction models using negative binomial regression. Accid. Anal. Prev. 61, 129–137.

Wen, Y., Chen, H., Wei, Y., Liu, Z., 2019. An approach of vehicle speed estimation based on didi GPS data. Proceedings of 19th COTA International Conference of Transportation. Nanjing, China, July 6–8, 2019 4207–4215.

Wu, F., 2019. Land financialisation and the financing of urban development in China. Land Use Policy 104412.

Wu, J., Xu, H., Zheng, Y., et al., 2018a. A novel method of vehicle-pedestrian near-crash identification with roadside LiDAR data. Accid. Anal. Prev. 121, 238–249.

Wu, Y., Abdel-Aty, M., Lee, J., 2018b. Crash risk analysis during fog conditions using real-time traffic data. Accid. Anal. Prev. 114, 4–11.

Wu, Y., Abdel-Aty, M., Park, J., et al., 2018c. Effects of crash warning systems on rear-end crash avoidance behavior under fog conditions. Transp. Res. Part C Emerg. Technol. 95, 481–492.

Xu, J., Liu, J., Sun, X., et al., 2018. The relationship between driving skill and driving behavior: psychometric adaptation of the Driver Skill Inventory in China. Accid. Anal. Prev. 120, 92–100.

Ye, X., Wang, K., Zou, Y., et al., 2018. A semi-nonparametric Poisson regression model for analyzing motor vehicle crash data. PLoS One 13. https://doi.org/10.1371/journal.pone.0197338.

Yeo, J., Park, S., Jang, K., 2015. Effects of urban sprawl and vehicle miles traveled on traffic fatalities. Traffic Inj. Prev. 16, 397–403.

Yu, C.-Y., 2015. How differences in roadways affect school travel safety. J. Am. Plan. Assoc. 81, 203–220.

Yu, C.-Y., Xu, M., 2018. Local variations in the impacts of built environments on traffic safety. J. Plan. Educ. Res. 38, 314–328.

Yue, Y., Zhuang, Y., Yeh, A.G., et al., 2017. Measurements of POI-based mixed use and their relationships with neighbourhood vibrancy. Int. J. Geogr. Inf. Sci. 31, 658–675.

Zhang, G., Yau, K.K., Zhang, X., 2014. Analyzing fault and severity in pedestrian–motor vehicle accidents in China. Accid. Anal. Prev. 73, 141–150.