Accident Analysis and Prevention 144 (2020) 105613

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Effects of built environment on bicycle wrong Way riding behavior: A data-driven approach

Sen Luan[a], Meng Li[b], Xin Li[c], Xiaolei Ma[a,d,*]

[a] School of Transportation Science and Engineering, Beijing Key Laboratory for Cooperative Vehicle Infrastructure System and Safety Control, Beihang University, Beijing, 100191, China

[b] Department of Civil Engineering, Tsinghua University, Beijing, 100084, China

[c] College of Transportation Engineering, Dalian Maritime University, Dalian, 116026, China

[d] Beijing Advanced Innovation Center for Big Data and Brain Computing, Beihang University, Beijing, 100191, China

ARTICLE INFO

Keywords:
Bicycle wrong way riding
Bike sharing
Built environment
Decision tree

ABSTRACT

Bicycle wrong way riding (WWR) is a dangerous and often neglected behavior that engenders threats to traffic safety. Owing to the lack of exposure data, the detection of WWR and its relationship with the built environment (BE) factors remain unclear. Accordingly, this study fills the research gaps by proposing a WWR detection framework based on bike-sharing trajectories collected from Chengdu, China. Moreover, this study adopts Negative Binomial-based Additive Decision Tree to investigate the impacts of built environment on WWR frequencies. Results reveal that (1) WWR distribution is unaffected by different periods in a day; (2) road length is more influential than road level and road direction in WWR occurrence; (3) company, bus stop, subway station, residence, and catering facility are primary contributors affecting WWR behavior during peak hours, whereas education becomes an emerging influential variable during nonpeak hours; and most importantly, (4) these variables clearly present non-linear effects on the WWR frequencies. Therefore, geographically differentiated policies should be adopted for bicycle safety improvement.

1. Introduction

Cycling is a popular, healthy, and environmentally friendly mode of traveling. It especially benefits the “first-and-last-mile” issues (Lin et al., 2013). In recent years, the shared bike as an alternative mode of transportation have proven its ability to encourage cycling (García-Palomares et al., 2012). However, the sharp increase in bicycle usage also poses other safety risks due to the uncontrolled diverse riding behaviors, particularly the risky ones.

Understanding the circumstances leading to the occurrence of these risky behaviors is significant to traffic safety improvement. Several built environment (BE) factors, such as bike facilities and land use attributes, are often used to explain the relationships between these risky behaviors and bicycle accidents. Moreover, research on bike facilities suggests that local authorities should prioritize bicycle infrastructure investment (e.g., bicycle lane expansion) to improve road safety conditions (Khatri et al., 2016). These studies were conducted at the micro level, with focus on the bike facility conditions on the road segment where the accident frequently occurs. Conversely, research on land use attributes belongs to the macro level, which highlights the overall effects of area-based land use on frequency of accidents (Wang et al., 2017). In sum, these two types of research are mutually supportive in an effort to promote safe riding environments.

Most bicycle accidents are inextricably linked to risky riding behaviors. Among the diverse behaviors, wrong way riding (WWR) is considered a major one, which is closely associated to bicycle safety. However, WWR is easily neglected by the public (Wachtel and Lewiston, 1994). Based on bicycle–vehicle crash records collected from California Statewide Integrated Traffic Records System in 2012, Stimpson et al. (2016) found that the bicyclists are responsible for 63 % of the collisions. Furthermore, findings on bicyclists’ aberrant movements preceding collisions showed that nearly 12 % of the collisions were caused by WWR, followed by 4.7 % resulting from turning left. Hence, WWR is the primary one in the aberrant behaviors, and its negative effects on traffic safety cannot be overlooked.

The lack of exposure data poses challenges in understanding when, where, and on what circumstances WWR behavior occurs frequently. Compared with driving, cycling has been historically considered a non-

* Corresponding author at: School of Transportation Science and Engineering, Beijing Key Laboratory for Cooperative Vehicle Infrastructure System and Safety Control, Beihang University, Beijing, 100191, China.

E-mail address: xiaolei@buaa.edu.cn (X. Ma).

https://doi.org/10.1016/j.aap.2020.105613
Received 27 September 2019; Received in revised form 24 May 2020; Accepted 24 May 2020
Available online 13 June 2020
0001-4575/

dominant transportation mode; most local traffic management authorities have no specific plans to collect bicycle data for riding risk quantification (Chen et al., 2018). However, the recent advent of the trajectory data from bike sharing system (BSS) may fill such gaps due to its wide usage and high-quality data (He et al., 2018). Although the bike-sharing data may not represent the entire population, it still proves to be a unique opportunity for examining certain riding behaviors.

Consequently, the objectives of this research are twofold:

- To develop a WWR behavior detection framework using a data-driven method and account for WWR frequency at segment level.
- To explore the effects of BE factors on the WWR frequency, a Negative Binomial-based Additive Decision Tree (NBADT) model is employed to prioritize influential factors on WWR and quantify their non-linear effects.

The rest of this paper is structured as follows. Section 2 introduces the literature review. Section 3 provides data descriptions including bicycle trajectory, road network, and BE factors. Section 4 introduces the framework of WWR detection. Section 5 describes the NBADT model to explore the relationship between BE factors and WWR frequency. Section 6 presents the model results and discussions. Section 7 summarizes the conclusions.

2. Literature review

2.1. Bicycle safety and built environment (BE)

Recent studies analyzed the relationship between bicycle crash incidents and BE factors to create a safe cycling environment for cyclists (Wei and Lovegrove, 2013; Chen, 2015; Wang et al., 2017). Generally, BE factors are often divided into five categories: (1) roadway design, (2) traffic facilities, (3) traffic controls, (4) socio-demographics, and (5) land use.

In terms of roadway design, positive effects of vehicle lane length, bicycle lane length, and intersection density on bicycle–auto collisions are confirmed (Wei and Lovegrove, 2013; Cai et al., 2016). The implementation of bicycle lanes not only attracts bicycle usage but also decreases the number of bicycle crash incidents (Bacchieri et al., 2010; Chen et al., 2012; Dhakal et al., 2018). As for intersections, high intersection density is significantly associated with bicycle crash frequency (Siddiqui et al., 2012; Strauss et al., 2013; Wei and Lovegrove, 2012; Vandenbulcke et al., 2014). Wang et al. (2019) found that urban bridges were identified as high-risk areas of bicycle-vehicle collisions. Regarding traffic facilities, studies have identified that the bus stop density is significantly related to bicycle crash frequency (Miranda-Moreno et al., 2011a, 2011b; Strauss et al., 2013; Wei and Lovegrove, 2012; Chen, 2015). In addition, subway station is rarely used as an explanatory variable in the bicycle crash analysis due to its scarcity. However, it continues to attract shared bikes for metro or rail access (Weliwitiya et al., 2019).

Among the traffic control variables, roads with speed limit below 15 mph have a negative effect on bicycle crashes, whereas roads with speed limit exceeding 35 mph have a significant effect on the increase in bicycle crash incidents (Chen and Fuller, 2014; Siddiqui et al., 2012). In relation to socio-demographics, the population and employment densities are significantly associated with bicycle crash frequency (Siddiqui et al., 2012), whereas these variables are insignificant in Chen's work (2015). As for land use factors, areas with high percentage of commercial and industrial development increase the likelihood of injuries among cyclists (Narayanamoorthy et al., 2013; Vandenbulcke et al., 2014; Cai et al., 2016). However, Strauss et al. (2012) found that the percentage of commercial land use is not an influential explanatory variable for bicycle crash incidents.

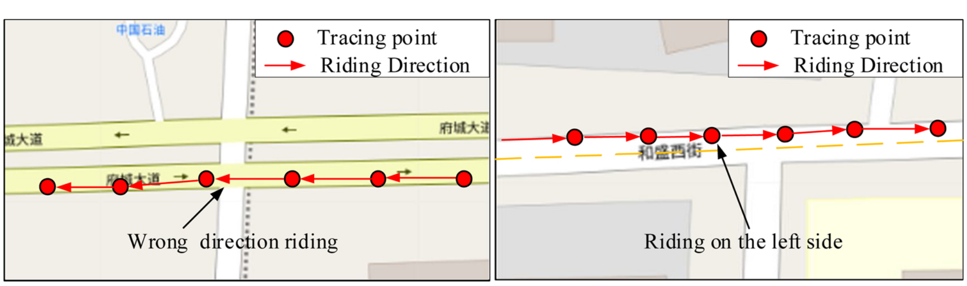
Two significant issues are found in the aforementioned studies. First, using traffic analysis zone (TAZ) as analytical unit hinders fine traffic management for local traffic safety improvement. Second, regression coefficient or "elasticity" (the regression coefficient times the mean value of corresponding variable) is constantly applied to capture the linear relationship between BE factors and bicycle accident frequencies, which cannot insightfully describe the nonlinear effects of BE factors on bicycle accidents.

This study is fundamentally different from the existing literature. Traditional approaches in bicycle safety related research heavily rely on survey methods to obtain individual-level sociodemographic and attitudinal information such as age, gender, cycling frequency, risk compensation and familiarity of the surrounding environment. From the perspective of individual cyclist, the behavioral heterogeneity can be modeled through survey data or mobile app usage data (Dhakal et al., 2018), and may become the contributing factor affecting WWR. However, when multiple WWR events are witnessed on some particular regions, there are reasons to believe that an occurrence of WWR may be no longer an individual cyclist's route choice. The surrounding BE may also yield significant impacts on WWR behavior, which has been verified by other scholars (Chen, 2015; Chen et al., 2018). In this study, we aim to detect the WWR events from massive shared bike trajectories and perform an aggregate level statistical analysis with BE factors on each road segment. The ultimate goal is to prevent potential bike crash happening by adopting proactive countermeasures to regulate WWR behavior.

2.2. WWR analysis

Risky riding behaviors are potential threats to bicycle safety, especially the easily overlooked WWR behaviors. WWR behavior mentioned in this study includes two types: (1) wrong direction riding on one-way road (as shown in Fig. 1a) and (2) riding on the left side of the bidirectional road (as shown in Fig. 1b). Despite our presumption that we are all highly familiar with WWR behaviors, quantifying them remains a difficult task.

In past years, acquiring WWR data relies on survey-based approaches. The survey process is generally divided into two categories: (1) consulting the interviewees to state their riding behavior preferences (Krizek and Johnson, 2006; Sener et al., 2009; Cole-Hunter



(a) WWR on one-way road (b) WWR on bidirectional road

Fig. 1. Diagrams exhibiting two kinds of WWR behavior.

et al., 2015) and (2) requesting the participants to recall riding behavior they had taken in different scenarios (Menghini et al., 2010; Broach et al., 2012; Hood et al. (2011)). However, either the survey or the interview is generally focused on individual cyclists. Consequently, such methods are unsuitable to explore further the riding behavior at a large-scale level, considering the expensive and time-consuming data collection procedure.

Thanks to technological advancement and access to bicycle GPS trajectories in recent years, thoroughly observing riding behaviors of cyclists is now possible. Langford et al. (2015) used ArcGIS to create buffers for each road (i.e., one buffer for one-way road and two sub-buffers for bidirectional road); here, WWR behaviors can be identified through two aspects: (1) on a one-way road, the riding direction is intuitively opposite to the road direction; (2) whereas on bidirectional road, cycling trajectory is located in the sub-buffer on the left side of the riding direction. However, Dhakal et al. (2018) demonstrated that the inherent location drift in GPS tracks might not guarantee the accuracy of the WWR detection on bidirectional roads. Additionally, the coverage of a single buffer and the overlapping adjacent buffers can also affect the detection accuracy. Therefore, accurate WWR detection method that can be applied to large-scale road networks is needed.

In terms of methods for WWR analysis, the relevant literature is very limited. We also reviewed these methodological studies in traffic accident analysis that could be potentially utilized for WWR analysis. Negative binomial (NB) regression as the basic method is applied earlier due to its ability of handling over-dispersion issue (Wei and Lovegrove, 2013; Pulugurtha and Sambhara, 2011). Thereafter, more spatial statistical models have emerged for pedestrian crash analysis with consideration of spatial autocorrelations (Chen, 2015; Vandenbulcke et al., 2014; Chen and Zhou, 2016; Wang and Kockelman, 2013). Considering the randomness of traffic accidents, random effects models with Poisson distribution have become increasingly popular (Aziz et al., 2017; Ukkusuri et al., 2011; Jiang et al., 2014). The abovementioned models use a macroscopic indicator named ‘elasticity’ to characterize the linear relationship between BE factors and traffic accident frequencies, but are difficult to capture their complicated nonlinear effects.

Owing to the high model fitting accuracy and powerful interpretation power, the decision tree-like methods were recently adopted in crash estimation and prediction. Prati et al. (2017) employed a Chi-squared Automatic Interaction Decision Tree model to analyze the severity of 49,621 bicycle crashes. Ding et al. (2018) confirmed the nonlinear effects of BE factors on the pedestrian crashes using Multiple Additive Poisson Regression Tree approach. The advantages of these decision tree based models embody in two aspects: (1) identify the role of variable in affecting the occurrence of traffic incidents based on variable importance; (2) further investigate how variables affect traffic incident frequency according to partial dependence plots. Therefore, the decision tree based method is used as the target model for WWR analysis.

The over-dispersion of WWR is another problem to be solved for model development. It is notably that Poisson distribution is known as the common distribution that have been widely employed in traffic count studies (Joshua and Garber, 1990). However, the WWR frequency distribution may fail to meet the assumption of Poisson distribution due to the over-dispersion phenomenon where the standard deviation of WWR frequency is greater than its mean value. Consequently, as a mixed distribution of Poisson and Gamma, NB distribution is more suitable in describing the over-dispersion traffic accident count data (Wang et al., 2019). Therefore, we aim to integrate decision trees with NB distribution to better model WWR events and identify the associated influential factors.

Collectively, to the best of our knowledge, research on bicycle WWR detection remains highly limited, hence research on the relationship between WWR and BE factors remains inadequate. Therefore, this study seeks to employ data-driven techniques to understand the magnitude of WWR behaviors using massive naturalistic datasets from BSSs. Moreover, we quantify the BE factors including road attributes, traffic facilities, and land use factors to explore their impacts on WWR frequency at segment level. Finally, a decision tree-based model with NB distribution is employed to disentangle the interplays between WWR and BE factors.

3. Data description

Data description aims to pave the way for WWR behavior detection, which contains two parts: describing the trajectory data from BSSs and introducing the road network data used for map matching with road design information extraction.

3.1. Bicycle trajectory data

The Internet-based non-dock BSSs (e.g., “Mobike,” “ofo,” “Bluegogo,” etc.) were extensively implemented in China in late 2016, which was proven effective in promoting cycling. As of July 2017, China has launched approximately 16 million shared bicycles that cater to 106 million users (National Information Center, 2017). Each shared bike can be connected by scanning the QR code via the corresponding mobile app. After a successful connection, the mobile app also records GPS trajectories of each trip while gathering information including user ID, timestamp, and coordinates (i.e., longitude and latitude). Thus, massive bicycle trajectories are accessible, which triggers a myriad of interesting studies in recent years, such as bike lane planning (Greg and Ipek, 2016; Bao et al., 2017), vehicle illegal parking detection (He et al., 2018), and bike usage pattern mining (Liu and Lin, 2019).

Trajectory data used in this study were collected within a week in October 2017 from the BSSs located in Chengdu City, the provincial capital of Sichuan, China. The average daily travel demand is approximately 600,000 times—generating more than 350 million track points. Fig. 2a illustrates the hot map based on the point count of trajectories over a day. It clearly shows that the trajectory points are highly concentrated within the central urban area (colored in red) according to its spatial distribution. Statistical results displayed in Fig. 2b show that 63 % of the distance values between two adjacent points are less than 6 m. Moreover, Fig. 2c shows that more than 80 % of the time interval values are less than 5 s. Therefore, sizeable data amount and high-quality data of shared bike data offer an unprecedented opportunity for bicycle riding behavior detection (Ma et al., 2019).

3.2. Road network data

The target road network is located in Wuhou District in the downtown area of Chengdu, where geographic information data (e.g., road name, road level, road length, road direction, and road nodes) can be downloaded from the open street map (OSM, an open source map platform). By removing highways (bicycles are not allowed on the highway), a total of 3090 edges (including 1505 one-way roads and 1585 bidirectional roads) and 7180 nodes remain as represented by the black lines and yellow dots in Fig. 3, respectively. Using such a road network with only one district aims to narrow down the geographic area and reduce the computational burden. In addition, the selected Wuhou District contains diverse land use attributes (e.g., rich cultural and educational resources, prosperous commerce, convenient transportation, etc.). Such feature consequently leads to a high possibility of WWR occurrence.

4. WWR behavior detection

WWR behavior detection is essentially the identification of riding behavior from massive bicycle trajectories. Therefore, the trajectory preprocessing including data cleaning and map matching is a primary task for accurate WWR detection.

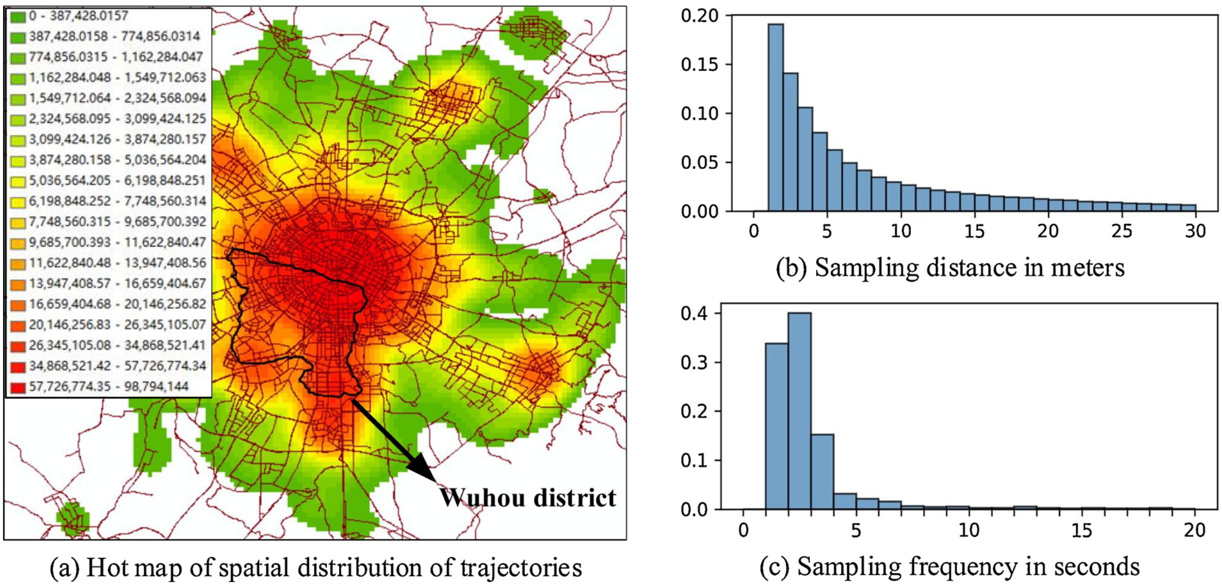


Fig. 2. Trajectory distribution of shared bikes in Chengdu, China, and the trajectory sampling statistics.

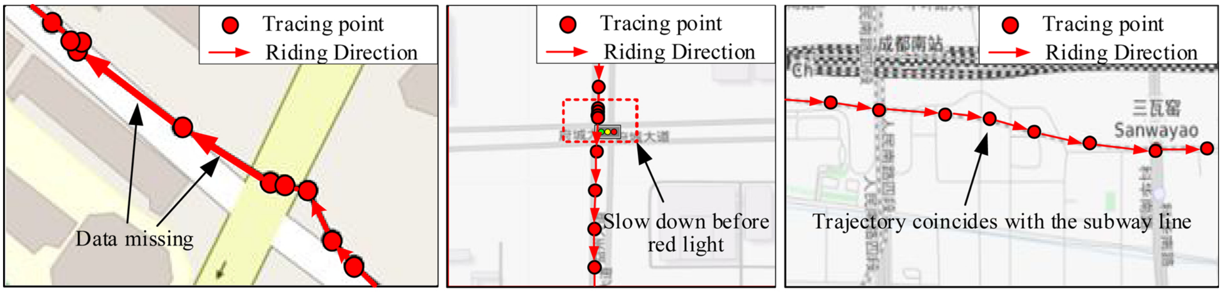
4.1. Trajectory preprocessing

4.1.1. Data cleaning

Noting that bike sharing trajectory data generated by the GPS devices integrated in mobile phones. Data errors affecting the correctness of the WWR detection are inevitable. Thus, data cleaning is necessary. In this study, tracking points that match the following conditions will be removed:

Trajectory with low sampling rate. The trajectory data will be missing when the GPS system is non-functional or the mobile communication network is affected by city canyon such as high buildings. The red lines in Fig. 4a represent a riding trajectory with data missing in some road

Fig. 3. Road network of Wuhou District.



(a) low sampling trajectory (b) abnormal low speed (c) abnormal high speed

Fig. 4. Examples of data errors.

segments.

Trajectory with abnormal speed. Riding speed generally ranges from 5 to 20 kmph, thereby forming the following two types of abnormal speed: (1) Riding speed less than 5 kmph, which may be caused by traffic lights or the cyclist walking to somewhere without locking the bike (as demonstrated in Fig. 4b) and (2) Riding speed greater than 20 kmph, as shown in Fig. 4c, which is typically caused by the GPS error or the cyclist transferring to a bus or subway (red points) without locking the bike.

4.1.2. Map matching

Map matching is used to project the GPS points onto the corresponding road segments on a given digital map, which is crucial for the detection of WWR. The relevant definitions are as follows.

Definition 1. *Trajectory.* A trajectory, $\tau = \{p_1, p_2, ..., p_n\}$, is a sequence of GPS points in chronological order, where each point $p_i = (id, t_i, lng_i, lat_i)$, $1 \le i \le n$, is appended with user identification id, sampling time t_i, longitude lng_i and latitude lat_i coordinates.

Definition 2. *Road Network.* A road network is an essential vector-graph $G = (N, E)$, where $N = \{n_1, n_2, ..., n_{2l}\}$ is a set of road nodes and $E = \{e_1, e_2, ..., e_l\}$ represents the edges or road segments. For each node $n_j \in N$, $1 \le j \le 2l$ contains three elements $n_j = (rid, lng_j, lat_j)$; whereas each road segment $e_k \in E$, $1 \le k \le l$ comprises five elements: (1) *rid* as the identification of each edge; (2) *level* as the road type (e.g., highway, secondary, and tertiary); (3) *one-way* as a binary value used to distinguish the road directions; (4) *azimuth* as a continuous value ranging from 0 to 360 for one-way road, whereas a null value corresponds for bidirectional road; and (5) StartNode and EndNode as the nodes of each edge.

Generally, bicycle trajectories have more flexible features than those of other vehicles: (1) bicycles can be taken in non-road areas such as road shoulders, squares, and so on; (2) bicycles can travel in both directions as long as the road allows them to pass—which is one of the inducements of WWR behavior. In view of the above two special features of bicycle trajectory, this paper performs a map-matching process composed of three steps.

Step 1: Index Building. A trajectory chain may be split into several effective sub-chains after data cleaning. To improve the efficiency of map matching, indexing the retained trajectories in chronological order is necessary.

Step 2: Matching. The map-matching framework borrowed from the work of Yuan et al. (2010) has submitted to four modifications to cater to the features of bicycling: (1) road segments belonging to high-level class are removed due to its exclusive access for vehicles only; (2) road directional information is considered and used to distinguish bidirectional road from a one-way road; (3) "*shift*" which refers to the shortest distance between the GPS point and the road segment, is added; (4) "*offset*" is calculated as the distance between the projected GPS point and the StartNode of the road segment. Finally, as shown in Fig. 5, each GPS point is appended with four elements after map-matching, namely, 1) *rid*, which corresponds to the road ID; 2) *one-way*, which is a binary value to identify whether the road is one way only; 3) *shift* value; and 4) *offset* value.

Step 3: Refinement. This step is significant to remove the matching results with two types of errors. (1) Geometric error means that the *shift* value is greater than a given distance threshold. As shown in Fig. 6a, the road segment with the red trajectories located in the residential area may not exist in the road map library. Thus, the trajectories belonging to them are matched to roads farther away. (2) Direction error, as Fig. 6b demonstrated, represents the riding direction which changes frequently (trajectory in red color) corresponding to the irregularity fluctuation of *offset* under the condition that the *shift* value meets a given distance threshold. Specifically, according to He et al. (2018), this threshold is set as 20 m.

4.2. Bicycle WWR detection

4.2.1. Detection on one-way road

One-way road means a road with a definite azimuth ranging from 0 to 360 in a digital map as shown in Fig. 7a. The difference between the riding angle and the road azimuth is the key element for WWR detection on a one-way road. The WWR detection strategy for one-way road is given as follows:

$$\begin{cases} \dfrac{\sum_{i=1}^{n_d} |\varphi_i - \omega|}{n_d} \in [d_1, d_2] \\ d_1 = 180 - \dfrac{\sigma_d}{\sqrt{n_d}} z_{a/2} \\ d_2 = 180 + \dfrac{\sigma_d}{\sqrt{n_d}} z_{a/2} \end{cases} \tag{1}$$

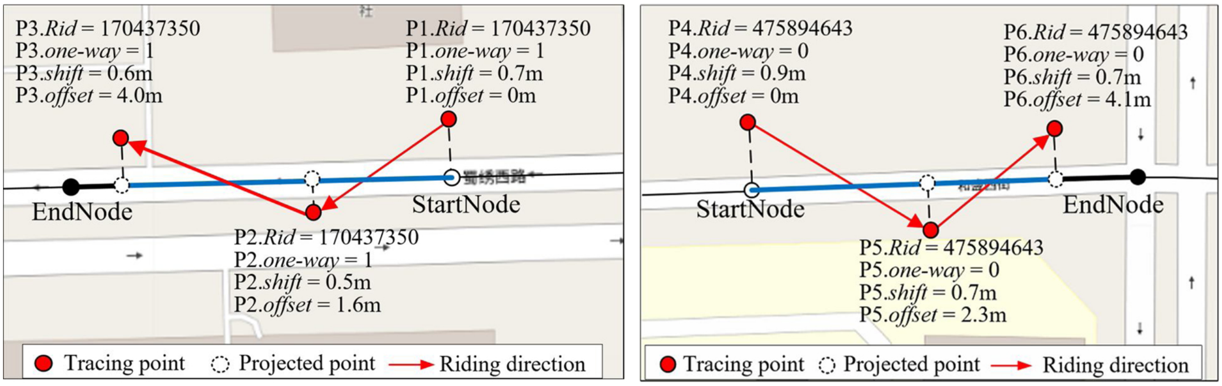
where φ_i represents the riding angle of the GPS point in a trajectory sub-chain; ω is the road azimuth; $[d_1, d_2]$ means the confidence interval of $|\varphi_i - \omega|$ at the significant level α (e.g., $\alpha = 0.05$); 180 is the expected value of $|\varphi_i - \omega|$; n_d and σ_d denote the number of samples and standard deviation, respectively.

4.2.2. Detection on bidirectional road

WWR on a bidirectional road is characterized by the trajectory points located on the left side of the riding direction (i.e., red points demonstrated on Fig. 7b). Thus, a geometric method is applied. First, the riding direction from P_s (StartNode) to P_e (EndNode) is assigned as the temporary azimuth of the road using $\boldsymbol{a} = P_sP_e = (a_x, a_y)$ for a representation. Second, let $\boldsymbol{b}_i = P_sP_i = (b_x, b_y)$. Finally, the WWR detection strategy for bidirectional road can be determined by the following equation:

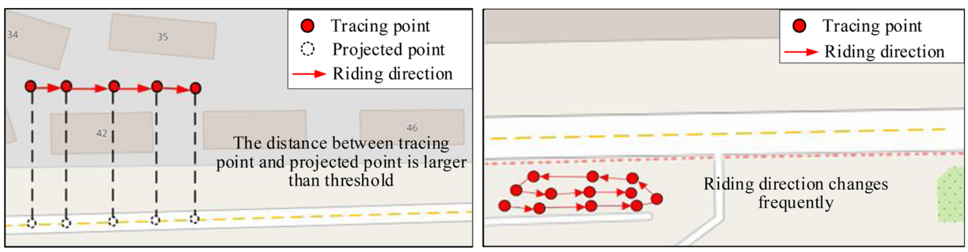
$$\boldsymbol{a} \times \boldsymbol{b_i} = a_x b_y - a_y b_x > 0 \tag{2}$$

The random location drift in bicycle trajectories brings huge disturbance to the WWR detection on bidirectional roads. To improve detection accuracy, we use Eq. (2) to judge each point in a sub-trip. When over 50 % of the track points appear on the left side of the riding



(a) Map-Matching on one-way road (b) Map-Matching on bidirectional road

Fig. 5. Examples of map-matching results.



(a) Distance error (b) Direction error

Fig. 6. Examples of map-matching with errors.

direction, this trip is recorded as a WWR event.

5. Modeling of BE effects on WWR frequency

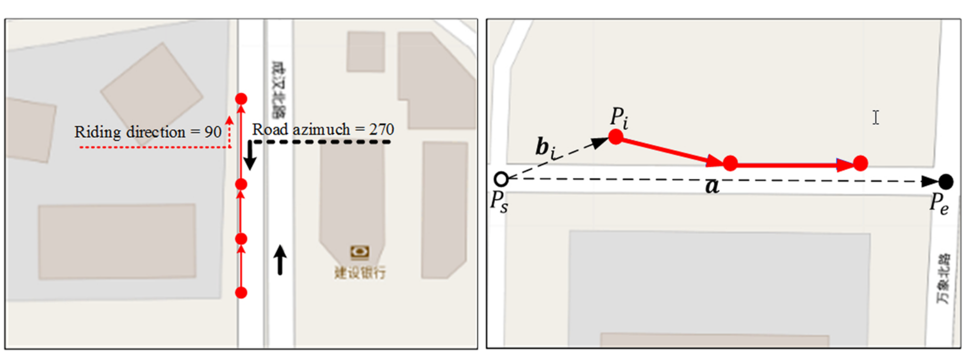
This study employs a machine-learning model, namely, Negative Binomial-based Additive Decision Tree (NBADT), to analyze the built environment effects on WWR frequency. Two reasons explain the decision to use this model: (1) it can accept different types of variables (e.g., discrete or continue, numeric or Boolean) to accommodate diverse BE factors and (2) it can quantify the nonlinear relationship between response and explanatory variables.

5.1. Negative binomial-based additive decision tree model

We assume that $\mathbf{x}$ is a set of explanatory variables (i.e., BE factors) and $F(\mathbf{x})$ is a function of the dependent variable y (i.e., WWR frequency) subjected to a negative binomial (NB), essentially a Poisson-Gamma distribution. Moreover, as demonstrated in Section 6.1, WWR frequencies are hyper variance count variables. Therefore, the proposed NBADT model can be considered an expansion of a basis decision tree $h(\mathbf{x};\alpha_m)$ (Friedman et al., 2001; Chung, 2013; Ma et al., 2017;) as shown in Eq. (3).

$$E\{F[\mathbf{x};(\beta_m,\mathbf{a}_m)_1^M]\} = \exp\left[\sum\nolimits_{m=1}^{M}\beta_m h(\mathbf{x};\mathbf{a}_m)+\varepsilon\right] \quad (3)$$

where the characterized parameters $\mathbf{a}_m$ in a decision tree are the splitting variables, which produce split locations and the terminal node for the individual trees; β_m represents the weight of each tree, and m represents the depth of each tree (also known as the tree complexity); $E(\bullet)$ represents the expectation; and λ denotes the parameter of Poisson distribution, and $\exp(\varepsilon)$ is a random error term subject to standard gamma distribution. To determine the optimal value of β_m and $\mathbf{a}_m$, Friedman (2001) proposed the gradient boosting approach as shown in Algorithm 1. In addition, this algorithm is implemented in the R package "mboost" (Hofner et al., 2014).



(a) WWR on one-way road (b) WWR on bidirectional road

Fig. 7. Interpretation of WWR behavior.

Algorithm 1. Gradient boosting approach.

Require: training data $\{y_i, \mathbf{x}_i\}_1^N$ from the observations	
1.	Initialization: $F_0(\mathbf{x}) = argmin_{\beta_m, \mathbf{a}_m} \sum_{i=1}^{N} L[y_i, (\beta_m, \mathbf{a}_m)]$
2.	**For** $m = 1$ to M, **do**:
3.	Obtain the gradient: $\tilde{y}_i = -\{\frac{\partial L[y_i, F(\mathbf{x}_i)]}{\partial F(\mathbf{x}_i)}\}_{F(\mathbf{x})=F_{M-1}(\mathbf{x})}, i = 1, \cdots, N$
4.	Parameter optimization: $\mathbf{a}_m, \beta_m = \underset{\mathbf{a}_m, \beta_m}{argmin} \sum_{i=1}^{N} L[\tilde{y}_i, F_{M-1}(\mathbf{x}_i) + \beta_m h(\mathbf{x}_i; \mathbf{a}_m)]$
5.	Then, update approximation: $F_m(\mathbf{x}) = F_{m-1}(\mathbf{x}) + \vartheta \cdot \beta_m h(\mathbf{x}_i; \mathbf{a}_m), 0 < \vartheta < 1$
6.	End **For**

The learning rate, ϑ, also known as the shrinkage, is applied in step 5 in Algorithm 1 to regulate the contribution of each decision tree. A small learning rate improves the minimize speed of the loss function but requires a large number of trees in the model. Therefore, we set ϑ to 0.05 in this research.

5.2. Relative importance

The relative importance, which refers to the contributions of explanatory variables on the variation of dependent variable, is one of the key elements for model interpretation. Explanatory variables have various contributions on the dependent variable. Thus, quantifying their relative importance is helpful for comparison. Given a tree T with J-terminal node in the model, the relative importance $I_i^2(T)$ of the input variable x_i, which is well documented in Breiman et al. (1984), is as follows:

$$I_i^2(T) = \sum_{t=1}^{J-1} \hat{l}_t^2 1(v_t = i) \tag{4}$$

where v_t in the indicator function represents the splitting variable associated with node t, and $\hat{l}_t^2$ is an empirical improvement of the corresponding squared error caused by the split. For a collection of decision trees $\{T_m\}_1^M$, the final relative importance of x_i can be generalized by their averages as follows:

$$I_j^2 = \frac{1}{M} \sum_{m=1}^{M} I_j^2(T_m) \tag{5}$$

5.3. Partial dependence plots[1]

Another measurement for further model interpretation is partial dependence, which provides the dependence between $F(\mathbf{x})$ and the joint explanatory variables (Friedman et al., 2001). Apparently, viewing high-dimensional arguments in plots is difficult. Thus, the partial dependence of $F(\mathbf{x})$ on selected subsets of the explanatory variables is proposed.

Let $\boldsymbol{z}_s$ be a selected target subset of size s in forming the input variables $\mathbf{x}$, that is, $\boldsymbol{z}_s = \{z_1, z_2, ..., z_s\} \subset \{x_1, ..., x_n\}$. Moreover, let $\boldsymbol{z}_c$ be the complement subset, $\boldsymbol{z}_s \cup \boldsymbol{z}_c = \mathbf{x}$. Thus, the approximation $F(\mathbf{x})$ depends on the variables in both subsets, $\widehat{F}(\mathbf{x}) = F(\boldsymbol{z}_s, \boldsymbol{z}_c)$. Then, the partial dependence of $F(\mathbf{x})$ on the selected variable subset $\boldsymbol{z}_s$ can be represented as follows:

$$\overline{F}_s(\boldsymbol{z}_s) = E_{\boldsymbol{z}_c}[\widehat{F}(\boldsymbol{z}_s, \boldsymbol{z}_c)] = \int \widehat{F}(\boldsymbol{z}_s, \boldsymbol{z}_c) p_c(\boldsymbol{z}_c) d\boldsymbol{z}_c \tag{6}$$

where $p_c(\boldsymbol{z}_c)$ is the complement marginal density, which can be estimated through the joint density of all the explanatory variables $\mathbf{x}$, $p(\mathbf{x})$, based on the training data as follows:

$$p_c(\boldsymbol{z}_c) = \int p(\mathbf{x}) d\boldsymbol{z}_s \tag{7}$$

Thus, the partial dependence function can be rewritten as

$$\overline{F}_s(\boldsymbol{z}_s) = \frac{1}{N} \sum_{i=1}^{N} \widehat{F}(\boldsymbol{z}_s, \boldsymbol{z}_{c,i}); \text{ i}= 1, ..., N \tag{8}$$

6. Results and discussions

6.1. Data preparation

In this study, we separately count the WWR frequency of each road segment at three different periods within a one-week time span, which is a temporal aggregation method of reducing the zero inflation in the counting result. Table 1 clearly shows that the mean WWR frequencies are higher during the morning and evening peak hours, achieving 129 and 130 frequencies, respectively. At noon hour, the WWR frequency dropped significantly to 67. These results are consistent with the daily traffic demand confirmed by Bacchieri et al. (2010). Notably, the SD of WWR frequency is greater than its mean value (i.e., over-dispersion), which is why the NB distribution is applied in this paper. The total number of trips are also provided to assess the WWR ratio (calculated as the percentage). It can be seen that, on average, 30 % of trips in the road network belong to WWR, among which 13 % and 46 % are detected as WWR on one-way and bidirectional roads, respectively. These statistics agree with the findings of He et al. (2018)'s study. In this study, we did not deduplicate the over-represented trips brought by the same user. This is because less than 7% of cyclists have two or more WWR events. Most importantly, from the perspective of traffic safety, the occurrence of each WWR event is a potential traffic risk no matter who rides the bicycle. Therefore, when counting the WWR frequency, we only focus on whether WWR behavior has occurred on each road segment and ignore the over-representation effect brought by individual cyclist.

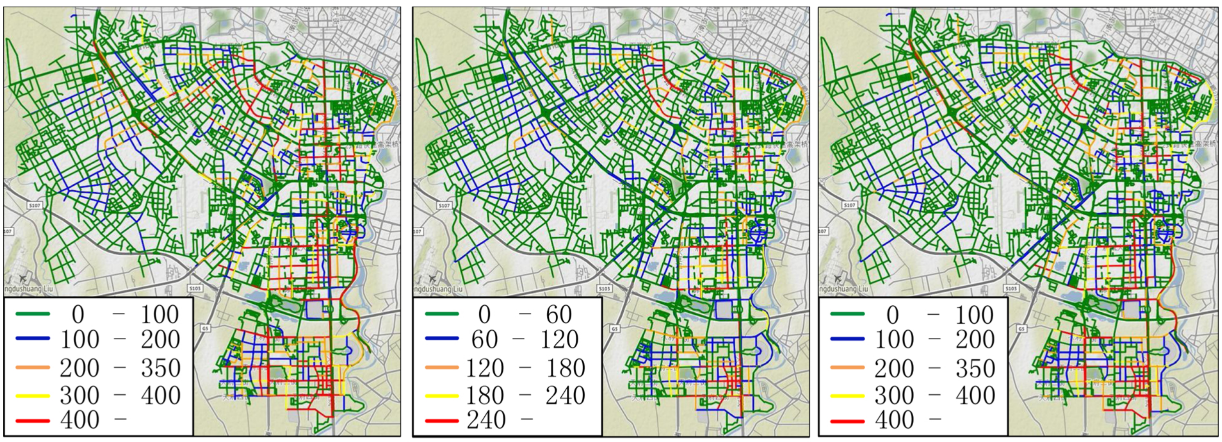
As demonstrated in Fig. 8, we project the WWR frequencies of road segments on the OSM platform to view their spatial distribution. To differentiate the road segments with various levels of WWR frequency, different quantization scales are applied to the three subgraphs representing different typical periods. Evidently, all the three spatial distributions of WWR frequencies show that road segments with high WWR frequency are located in the eastern region of Wuhou District. Thus, a preliminary conclusion can be made that the temporal factor is almost irrelevant to the spatial distribution of WWR frequency.

According to the previous literature, certain BE factors such as

[1] The formulas represented in the first two paragraphs of Section 5.3 are displayed incorrectly, but when entering the editing interface, these formulas are found to be correct. This prevents me from correcting them, even though I entered the correct formulas and SAVED them.

Table 1
General statistics of WWR frequency on road segments (n = 3090).

Time period	Trips with WWR behavior				Total trips	WWR ratio	Unit
	Mean	SD	Min.	Max.			
07:00–09:00	129.23	171.70	0	2135	1,243,990	32.1%	num/week
12:00–14:00	67.35	78.90	0	784	680,103	30.6%	num/week
16:00–18:00	130.10	166.29	0	1822	1,268,167	31.7%	num/week



(a) 7:00 - 9:00 (b) 12:00 - 14:00 (c) 16:00 - 18:00

Fig. 8. WWR frequency plot in different periods.

Table 2
General Statistics of BE factors aggregated in road segments (n = 3090).

Variable	Mean	S.D.	Min.	Max.	Unit
Roadway design					
Road direction	Y: one-way road; N: bidirectional road				Category
Road level	S: secondary level; T: tertiary level				Category
Road length	168.64	155.00	6	1432	Meter
Traffic facilities					
Bus stop	3.22	5.70	0	43	num/road-buffer1
Subway station	0.35	1.30	0	16	num/road-buffer2
Land use attributes					
Company	37.86	79.11	0	959	num/road-buffer2
Education	10.30	15.08	0	167	num/road-buffer2
Entertainment	7.19	10.49	0	105	num/road-buffer2
Finance	7.08	12.60	0	153	num/road-buffer2
Food	83.00	68.38	0	513	num/road-buffer2
Government	4.52	8.65	0	92	num/road-buffer2
Hotel	7.03	13.12	0	125	num/road-buffer2
Life service	40.13	45.25	0	406	num/road-buffer2
Medical	9.70	22.52	0	378	num/road-buffer2
Residential	7.72	9.19	0	86	num/road-buffer2
Shopping	65.51	91.94	0	849	num/road-buffer2
Scenery	0.53	1.23	0	18	num/road-buffer2

Note: buffer1 equals 20 m, and buffer2 equals 50 m.

roadway design, traffic facilities, traffic controls, socio-demographics, and land use attributes are shown to be closely related to bicycle safety. However, considering WWR behavior occurrence on certain road segments, we only select road-related factors as explanatory variables; whereas the intersection-related traffic controls and zone-based socio-demographic factors are excluded in this study. The road buffer is applied as the analytical unit to perform the micro-level analysis. Table 2 illustrates the data summary of these influential variables.

The selected BE factors contain three categories. Roadway design attributes include road direction, road level, and road length. Road direction differentiates one-way road from a bidirectional one, whereas road level determines a secondary or a tertiary road. Lastly, road length refers to the length of road in meters. The traffic facilities include bus stops and subway stations located on the roadside. The land use attributes comprise multiple types of points of interest (POI) including company, education, entertainment, finance, food, government, hotel, medical, residential, shopping, and scenery. Other factors such as car service, car sale, and vehicle repair that are not related to bicycle activities are excluded. When counting the number of bus stops, the width of the road buffer is 20 m (i.e., buffer1 = 20 m) due to its closeness to roadside; whereas the rest of the BE factors applied a buffer of 50 m (i.e., buffer2 = 50 m). In addition, bicycle flow is not considered because of its collinearity with the WWR frequency. In this study, we did not consider the area of each POI due to parcel-level land use data acquisition obstacle. Different from traditional multivariable regression-based models, the developed decision tree-based approach can

Table 3
Relative importance of explanatory variables on WWR frequency.

Category	Variables	7:00–9:00		12:00–14:00		16:00–18:00	
		Rank	Imp. (%)	Rank	Imp. (%)	Rank	Imp. (%)
Road attributes	Road direction	16	0.48	16	0.13	16	0.53
	Road level	15	0.57	17	0.09	15	0.69
	Road length	10	2.09	11	1.76	11	1.51
Traffic facilities	Bus stop	**2**	**27.48**	**4**	**12.43**	**1**	**28.26**
	Subway station	**3**	**11.58**	8	2.70	5	6.53
Land use	Company	**1**	**29.01**	**1**	**19.42**	**2**	**23.72**
	Education	8	2.96	**2**	**16.38**	6	6.08
	Entertainment	9	2.13	13	1.31	10	2.39
	Finance	5	4.20	6	7.26	7	4.90
	Food	6	4.04	5	10.00	**3**	**9.72**
	Government	13	1.23	10	1.91	15	0.99
	Hotel	7	3.66	7	6.18	8	3.94
	Life service	14	0.87	9	2.65	14	1.10
	Medical	12	1.38	14	0.40	12	1.51
	Residence	**4**	**6.64**	**3**	**15.28**	**4**	**8.04**
	Shopping	11	1.39	12	1.75	13	1.29
	Scenery	17	0.29	15	0.34	17	0.29

quantify the effect of changing the number of POIs on WWR events, and this effect is nonlinear. This treatment can alleviate the spatial biases caused by counting the number of POIs without area information to some extent.

6.2. Variable relative importance

As demonstrated in the methodology, the relative importance is a meaningful indicator to measure the contribution of explanatory variables to the variation of response variables. The relative importance is a value ranging from 0 to 1, where the sum of the value of all explanatory variables equals 100 %. Table 3 presents the ranks of the relative importance of the explanatory variables produced by NBADT model during the morning peak, noon, and evening peak periods. In addition, the tree complexity is set to 5, and the learning rate is set to 0.05 to avoid excessive computational burden when building the NBADT model.

During morning peak period, the company ranks first in terms of relative importance, with a contribution of 29.01 % on the variation of WWR frequency. Bus stop, subway station, and residence are the second, third, and fourth most influential variables with a contribution of 27.48 %, 11.58 %, and 6.64 %, respectively. Other land use features contribute less than 5% of effects on WWR frequency. This result is consistent with our intuition, where the four most important variables are closely related to the morning commuting behavior that brings a large demand for cycling, which in turn generates significant impacts on the WWR frequency. On the contrary, factors such as education, finance, government, entertainment, scenery etc., do not attract more cycling demand in morning rush hours. Road level and road direction are weakly correlated with WWR frequency due to their contributions at approximately 0.5 %, and the contribution of road length is 2.09 %, which means that cyclists' choices of WWR behavior have little to do with road attributes, and their motivations of risking WWR may only be how to reach their destinations conveniently.

At noon period, the company remains the most influential variable with a contribution of 19.42 %. By contrast, the education and the residence rank second and third, accounting for 16.38 % and 15.28 % of the overall contributions, respectively. The bus stop accounting for 12.34 % of contribution ranks fourth but is still considered closely associated with WWR frequency. The food ranks fifth, but its contribution is higher than that during the morning peak period. The above results indicate that the WWR frequency is very likely to be affected by the cycling needs from employees and students going out for lunch or participating in extracurricular activities. In terms of road attributes, the contributions of road direction, road level and road length on WWR frequency are still minimal, which once again confirms the marginal effects of road attributes on WWR frequency, and such effects do not change significantly over time.

The relative importance of BE factors in the evening peak period is similar to that in the morning peak period. The bus stop and company remain the two most influential variables to WWR frequency, and the food, residential, subway, and education closely lag behind, demonstrating that transferring to public transit facilities, biking to home, and dining out are primary activities during off-hours. Employees and students are likely to dine out or return home by bike, and thus cause high WWR frequencies around workplaces and educational facilities. In addition, the road and scenery attributes still are irrelevant to the occurrence of WWR, which is consistent with the other two periods.

We can also find from Table 3 that the relative importance of the same variable ranks differently during different times of day. Specifically, the contributions of bus stop, subway station and residence at noon are lower than that during morning/evening peak hours. This is probably due to the issue of "first-and-last-mile": One makes the beginning of his/her trip by public transit from home place, and returns home from workplace to complete his/her last trip. However, most employees and students limit their activities around either workplaces or schools at noon, and thus increase the probability of WWR. That explains why the rankings of education and company at noon are higher than that during morning and evening pick hours.

Collectively, the bus stop, residential, company, and food account for more than 70 % of the total contributions on WWR frequency during the daily peak hours, indicating the key role of diversity of the BE factors. These results mean that the commuter is the main user group of shared bikes, and the vicinity of traffic facilities (i.e., bus and subway) is generally the hot spots where shared bikes arrive or depart to solve the "first-and-last-mile" issue. The total relative importance of education, residential, and food surges during noon period accounts for 41.66 % of the total contributions. Thus, WWR frequency may be stimulated by the increased travel demand for dining out by employees or for going home by students. All of these findings suggest that road segments surrounded by a large number of bus stops, companies, residential buildings, educational institutes, subway stations, and restaurants should be given extra attention to improve cycling safety.

6.3. Nonlinear effects of important variables

The next step is to measure the dependence of the WWR frequency on BE factors via the partial dependence plots. In this study, the four

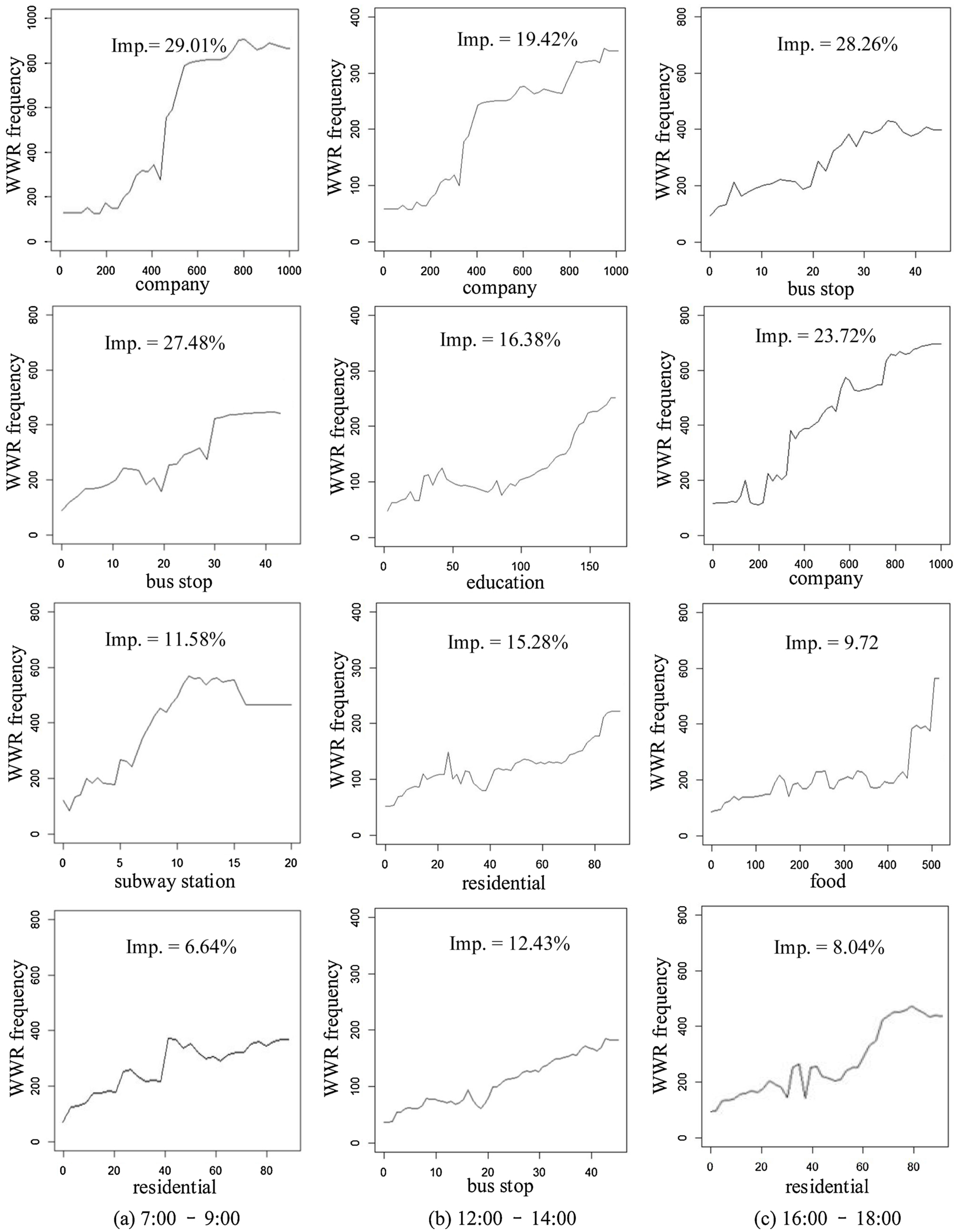


Fig. 9. Non-linear effects of the three most influential variables on WWR frequency.

most important BE factors from each typical period are selected due to their considerable total contributions to WWR frequency as demonstrated in Fig. 9. In each subplot, the horizontal scale refers to the range of the explanatory variables, whereas the vertical scale represents the WWR frequency.

During the morning peak period (Fig. 9a), when the number of the companies near a road segment is less than 550, the WWR frequency increases. The effect remains stable when the number of companies exceeds this range. This is understandable since people do not need to travel by bike between buildings in a highly populated area. The WWR frequency rises slowly when the number of bus stops is below 14. However, when the number of bus stops falls between 14 and 30, the effect substantially fluctuates. After the threshold exceeds 30, the effect remains stable. This implies a positive relationship between bus stop and WWR frequency in those areas with low transit accessibility where passengers have to make a long-distance transfer between two stops. When the number of stops equals to 30, the most WWR events can be observed. However, WWR frequency becomes stable as more convenient transit service is provided. An approximate positive effect of residence on WWR frequency occurs when the number of residential buildings is within a certain threshold (i.e. 42). However, when the number of residential buildings exceeds 42, no obvious effect occurs on WWR frequency. It may be attributed to the fact that most of the shared bikes near residential areas are being taken by commuters, and the remaining available bicycles are highly limited. As for the subway station, a non-linear effect on WWR frequency is also established. More interestingly, when the subway station number is below 12, a positive correlation with WWR frequency can be observed; whereas beyond this range, a weak negative relationship can be observed. This trend is similar with that of bus stop. As more subway entrances are available, passengers are less likely to risk for WWR.

At noon period (Fig. 9b), the company, education, residence, and bus stop all yield significant effects on WWR frequency. The company and education factors attract potential bike users due to employee and students dining out or going home. More WWR events can be found as the number of companies increases. One possible reason is that employees have to return to their workplaces from restaurants on time and thus increase the possibility of WWR behavior. The effect of residence is very similar to that of the morning rush hours, except that it corresponds to a lower WWR frequency since the majority of employees do not need to go home during noon.

During the evening period (Fig. 9c), the dependencies of WWR frequency on bus stop, company, and residence are similar to that during the morning rush hours but only are slightly different in the frequency magnitude. This indicates that road segments near companies and residences contribute a higher WWR frequency. It reflects the role of bike sharing system to address the issue of ‘first-and-last-mile’ for commuters, which triggers a large amount of bicycling activities. Interestingly, when the number of food restaurants is below 450, the WWR frequency fluctuates between 100 and 220. By contrast, the WWR frequency rises sharply when the number exceeds the threshold. This finding indicates that the more concentrated the food services, the more attractive they are to diners, which drives the demand for bicycling. Thus, road segments near the area with catering service yield high traffic risks during evening rush period. One possible reason is that people may go for dinner appointments on time by taking the task of WWR.

In the perspective of microscopic analysis, the non-linear effects of these influential factors on WWR frequency are confirmed. When the number of the aforementioned POIs (i.e., company, bus stop, and subway station, residence, food) reaches a certain value, the WWR frequency no longer increases but remains steady or even decreases. Interestingly, according to Fig. 9a, subway stations incur more WWR events than bus stops, which may be attributed to the fact that the city bicyclists prefer the subway for commuting than buses during rush hours. However, the relative importance of subway stations shown in the Table 3 is smaller than that of bus stops. This result may be caused by the sparseness of the subway station in the road network compared with other POIs.

Collectively, to improve safety of the bicycling environment, road segments surrounded by multiple companies, educational institutes, catering services, and residences should be given extra attention. Reasonable POI distribution as well as accessible transportation facilities are helpful to reduce WWR frequency, thereby reducing potential traffic risk. For example, properly decentralizing catering services and adding subway entrances (at least one on each side of the road) may be efficient strategies to improve cycling safety. In the road network level, geographically differentiated policies should be adopted.

7. Conclusions

Cyclists riding with WWR behavior collide easily with the opposite bikes or vehicles. Despite the lack of attention from the majority of bicyclists, WWR behavior is actually a hazard to urban traffic safety with the surge of bike usage. To improve the urban bicycling environment, this study develops a WWR detection framework and explores the effects of BE factors on WWR frequency.

Thanks to the massive accessible trajectories from BSSs, bicyclists' frequent WWR behavior is identified via data-driven methods. The WWR frequency is then counted on the basis of data collected in a span of a week from Wuhou District located in Chengdu, China. With the OSM platform, WWR frequency plots from three different periods evidently show that road segments with high WWR frequency are concentrated in the eastern region of Wuhou District, where spatial distributions of the three periods are very similar. By considering the hyper variance or over-dispersion phenomenon in WWR frequency samples, the NB-based model is applied. As an expansion of Poisson distribution, NB distribution is suitable to describing the WWR frequency. According to the relative importance, bus stops and subway stations from traffic attributes as well as the land use of company, residential, education, and food are identified as the most influential variables. Among the road attributes, road level and road direction contribute relatively less to WWR, whereas road length exhibits mediocre contribution. Most importantly, partial dependence plots strongly prove the non-linear relationship between WWR frequencies and the influential BE factors. The results suggest that road segments surrounded by a range of companies, educational facilities, residences, public transit hubs and catering services deserve more attention from transportation agencies because WWR behavior associated with these places is likely to be observed. In addition, increasing the accessibility of public transit facilities such as adding more subway station entrances or bus stops could help reduce WWR frequency. In the road network level, geofencing strategies should be adopted to regulate cyclists to park their shared bikes in designated areas or rebalance bikes more effectively to prevent WWR for cycling safety.

There are certain limitations of this study. For instance, spatial biases in POI data are not fully eliminated. To address this issue, the parcel-level land use data are desired to consider the area of each POI. In addition, this study did not incorporate individual cyclist's attitudinal, utility-related and familiarity factors. These unobserved factors play vital roles in explaining WWR behavior, and should be collected through either surveys or interviews as explanatory variables.

Frequent risky riding behaviors are not limited to WWR. They also include the preemption of vehicle lanes, sudden lane changing, and beating the red light. To reduce urban traffic hazard and promote the friendly development of bicycle transportation mode, continual research on cycling safety is an urgent need. In addition, other advanced methods such as zero-inflated regression and random effect model can also be explored to accommodate the data variety or consider both temporal and spatial random effects. In terms of data availability, the area of each POI should be considered as an explanatory variable to quantify the attractiveness of bicycle ridership in the future work.

Author statement

Sen Luan: Data analysis, model development. Meng Li.: Raw data processing. Xin Li: Visualization, result interpretation. Xiaolei Ma: Conceptualization, paper editing.

Declaration of Competing Interest

None.

Acknowledgements

This paper is supported by National Key R&D Program of China (2018YFB1601600). The authors also thank Joint Laboratory for Future Transport and Urban Computing of Amap for data support.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.aap.2020.105613.

References

Aziz, H.M.A., Nagle, N.N., Morton, A.M., Hilliard, M.R., White, D.A., Stewart, R.N., 2017. Exploring the impact of walk-bike infrastructure, safety perception, and built-environment on active transportation mode choice: a random parameter model using New York City commuter data. Transportation 1–23.

Bacchieri, G., Barros, A.J.D., Santos, J.V.D., Gigante, D.P., 2010. Cycling to work in Brazil: users profile, risk behaviors, and traffic accident occurrence. Accid. Anal. Prev. 42 (4), 1025–1030.

Bao, J., He, T., Ruan, S., Li, Y., Yu, Z., 2017. Planning bike lanes based on sharing-bikes' trajectories. Acm Sigkdd International Conference on Knowledge Discovery & Data Mining.

Breiman, L., Friedman, J., Olshen, R.A., Stone, C.J., 1984. Classification and Regression Trees.

Broach, J., Dill, J., Gliebe, J., 2012. Where do cyclists ride? A route choice model developed with revealed preference GPS data. Transp. Res. Part A Policy Pract. 46 (10), 1730–1740.

Cai, Q., Lee, J., Eluru, N., Abdel-Aty, M., 2016. Macro-level pedestrian and bicycle crash analysis: incorporating spatial spillover effects in dual state count models. Accid. Anal. Prev. 93, 14–22.

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

Chen, L., Chen, C., Srinivasan, R., McKnight, C.E., Ewing, R., Roe, M., 2012. Evaluating the safety effects of bicycle lanes in New York City. Am. J. Public Health 102 (6), 1120–1127.

Chen, P., Sun, F., Wang, Z., Gao, X., Jiao, J., Tao, Z., 2018. Built environment effects on bike crash frequency and risk in Beijing. J. Safety Res. 64, 135–143.

Chen, P., Zhou, J., 2016. Effects of the built environment on automobile-involved pedestrian crash frequency and risk. J. Transp. Health 3 (4), 448–456.

Chung, Y.S., 2013. Factor complexity of crash occurrence: an empirical demonstration using boosted regression trees. Accid. Anal. Prev. 61 (8), 107–118.

Cole-Hunter, T., Donaire-Gonzalez, D., Curto, A., Ambros, A., Valentin, A., Garcia-Aymerich, J., Martinez, D., Braun, L.M., Mendez, M., Jerrett, M., 2015. Objective correlates and determinants of bicycle commuting propensity in an urban environment. Transp. Res. D Transp. Environ. 40, 132–143.

Dhakal, N., Cherry, C.R., Ling, Z., Azad, M., 2018. Using CyclePhilly data to assess wrong-way riding of cyclists in Philadelphia. J. Safety Res. 67, 145–153.

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.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 1189–1232.

Friedman, J., Hastie, T., Tibshirani, R., 2001. The Elements of Statistical Learning. Springer series in statistics New York.

García-Palomares, J.C., Gutiérrez, J., Latorre, M., 2012. Optimizing the location of stations in bike-sharing programs: a GIS approach. Appl. Geogr. 35 (1-2), 235–246.

Greg, G., Ipek, S., 2016. Planning for bike share connectivity to rail transit. J. Public Trans. 19 (2), 1.

He, T., Bao, J., Li, R., Ruan, S., Li, Y., Tian, C., Zheng, Y., 2018. Detecting vehicle illegal parking events using sharing bikes' trajectories. KDD 340–349.

Hofner, B., Mayr, A., Robinzonov, N., Schmid, M., 2014. Model-based boosting in R: a hands-on tutorial using the R package mboost. Comput. Stat. 29 (1-2), 3–35.

Hood, J., Sall, E., Charlton, B., 2011. A GPS-based bicycle route choice model for San Francisco, California. Transport. Letters 3 (1), 63–75.

Jiang, X., Abdel-Aty, M., Alamili, S., 2014. Application of Poisson random effect models for highway network screening. Accid. Anal. Prev. 63 (1), 74–82.

Joshua, S.C., Garber, N.J., 1990. Estimating truck accident rate and involvements using linear and Poisson regression models. Transp. Plan. Technol. 15 (1), 41–58.

Khatri, R., Cherry, C.R., Nambisan, S.S., Han, L.D., 2016. Modeling route choice of utilitarian bikeshare users with GPS data. Transp. Res. Rec. 2587 (1), 141–149.

Krizek, K.J., Johnson, P.J., 2006. Proximity to trails and retail: effects on urban cycling and walking. J. Am. Plan. Assoc. 72 (1), 33–42.

Langford, B.C., Chen, J., Cherry, C.R., 2015. Risky riding: naturalistic methods comparing safety behavior from conventional bicycle riders and electric bike riders. Accid. Anal. Prev. 82, 220–226.

Lin, J.R., Yang, T.H., Chang, Y.C., 2013. A hub location inventory model for bicycle sharing system design: formulation and solution. Comput. Ind. Eng. 65 (1), 77–86.

Liu, H.C., Lin, J.J., 2019. Associations of built environments with spatiotemporal patterns of public bicycle use. J. Transp. Geogr. 74, 299–312.

Ma, X., Ding, C., Luan, S., Yong, W., Wang, Y., 2017. Prioritizing influential factors for freeway incident clearance time prediction using the gradient boosting decision trees method. Ieee Trans. Intell. Transp. Syst. 18 (9), 2303–2310.

Ma, X., Zhang, X., Li, X., Wang, X., Zhao, X., 2019. Impacts of free-floating bikesharing system on public transit ridership. Transp. Res. D Transp. Environ. 76, 100–110.

Menghini, G., Carrasco, N., Schüssler, N., Axhausen, K.W., 2010. Route choice of cyclists in Zurich. Transp. Res. Part A Policy Pract. 44 (9), 754–765.

Miranda-Moreno, L.F., Morency, P., El-Geneidy, A.M., 2011a. The link between built environment, pedestrian activity and pedestrian-vehicle collision occurrence at signalized intersections. Accid. Anal. Prev. 43 (5), 1624–1634.

Miranda-Moreno, L.F., Strauss, J., Morency, P., 2011b. Disaggregate exposure measures and injury frequency models of cyclist safety at signalized intersections. Transp. Res. Rec. 2236 (1), 74–82.

Narayanamoorthy, S., Paleti, R., Bhat, C.R., 2013. On accommodating spatial dependence in bicycle and pedestrian injury counts by severity level. Transp. Res. Part B Methodol. 55 (3), 245–264.

National Information Center, 2017. Shared Bicycle Industry Employment Research Report. . (accessed Aug 14, 2019; in Chinese). http://www.sic.gov.cn/News/250/8452.htm.

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., Sambhara, V.R., 2011. Pedestrian crash estimation models for signalized intersections. Accid. Anal. Prev. 43 (1), 439–446.

Sener, I.N., Eluru, N., Bhat, C.R., 2009. An analysis of bicycle route choice preferences in Texas. US. Transportation 36 (5), 511–539.

Siddiqui, C., Abdelaty, M., Choi, K., 2012. Macroscopic spatial analysis of pedestrian and bicycle crashes. Accid. Anal. Prev. 45 (3), 382–391.

Stimpson, J.P., Zhu, H., Wilson, F.A., 2016. Bicyclists found at fault for bicycle crashes in California. Am. J. Emerg. Med. 34 (8), 1699–1701.

Strauss, J., Miranda-Moreno, L.F., Morency, P., 2013. Cyclist activity and injury risk analysis at signalized intersections: a Bayesian modelling approach. Accid. Anal. Prev. 59 (5), 9–17.

Ukkusuri, S., Hasan, S., Aziz, H.M.A., 2011. Random parameter model used to explain effects of built-environment characteristics on pedestrian crash frequency. J. Transport. Res. Record.

Vandenbulcke, G., Thomas, I., Panis, L.I., 2014. Predicting cycling accident risk in Brussels: a spatial case-control approach. Accid. Anal. Prev. 62 (1), 341–357.

Wachtel, A., Lewiston, D., 1994. Risk factors for bicycle-motor vehicle collisions at intersections. ITE J. (Inst. Transport. Eng.) 64 (9), 30–35.

Wang, Y., Kockelman, K.M., 2013. A Poisson-lognormal conditional-autoregressive model for multivariate spatial analysis of pedestrian crash counts across neighborhoods. Accid. Anal. Prev. 60 (12), 71–84.

Wang, J., Lindsey, G., Hankey, S., 2017. Exposure to Risk and the Built Environment, an Empirical Study of Bicycle Crashes in Minneapolis.

Wang, H., De Backer, H., Lauwers, D., Chang, S.J., 2019. A spatio-temporal mapping to assess bicycle collision risks on high-risk areas (Bridges)-A case study from Taipei (Taiwan). J. Transp. Geogr. 75, 94–109.

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 (12), 129–137.

Weliwitiya, H., Rose, G., Johnson, M., 2019. Bicycle train intermodality: effects of demography, station characteristics and the built environment. J. Transp. Geogr. 74, 395–404.

Yuan, J., Zheng, Y., Zhang, C., Xie, X., Sun, G., 2010. An interactive-voting based map matching algorithm. In Eleventh International Conference on Mobile Data Management.