Cities 158 (2025) 105712

Contents lists available at ScienceDirect

Cities

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

Spatiotemporal variations of private e-bike trips with explainable data-driven technologies

Peixiao Wang [a,b], Hengcai Zhang [a,b,*], Beibei Zhang [a,b], Shifen Cheng [a,b], Feng Lu [a,b,c], Tong Zhang [d]

[a] State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China
[b] College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China
[c] Fujian Collaborative Innovation Center for Big Data Applications in Governments, Fuzhou 350003, China
[d] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

ARTICLE INFO

Keywords:
Private e-bike trips
Trip patterns
Spatiotemporal variations
Driving mechanisms
Spatiotemporal random forest
Improved SHAP

ABSTRACT

Understanding the trip features and driving mechanisms of e-bikes, particularly their spatiotemporal variations, is essential for improving traffic mobility, reducing pollution, and enhancing road safety. Currently, existing studies have two main gaps: (1) the spatiotemporal variations of private e-bikes remain unclear, and (2) there is a lack of explainable data-driven techniques that can analyze the spatiotemporal variation effects of driving mechanisms, especially considering spatiotemporal heterogeneity. In this study, using the private e-bikes trips in Wuhan, China as a case study, a novel explainable framework is proposed to analyze the spatiotemporal variations in their trip features and driving mechanisms. More specifically, a novel spatiotemporal random forest is presented to build a nonlinear mapping between driving factors and private e-bike trips in the spatiotemporal domain. Then, the classical SHAP method is extended to map Shapley values onto the time and space axes, enabling the exploration of spatiotemporal variations in driving factors. Findings reveal that: (1) private e-bikes are frequently used for short and medium-distance trips, typically exceeding 1 km, and play a crucial role in daily urban commuting; (2) Factors such as Historical trip frequency, Commercial POI Density, and Hospital POI Density are positively correlated with private e-bike trips; (3) the influence of driving factors on private e-bike trips vary significantly across different spatial locations and time windows. This study offers an innovative analytical framework for a more profound comprehension of e-bike trips. Additionally, the findings can aid authorities in crafting more effective policies and planning strategies.

1. Introduction

As motor vehicle ownership continues to rise rapidly, urban traffic congestion and environmental pollution have worsened, prompting human to seek more environmentally friendly modes of transportation (Ren et al., 2023; Zhou et al., 2023), such as bicycles or electric bikes (e-bikes). In this context, understanding the trip features and driving mechanisms of bicycle or e-bike usage is essential, as it provides valuable insights for improving traffic mobility, reducing pollution, and enhancing road safety (Guo et al., 2021; Lv et al., 2021; Schläpfer et al., 2021).

Recently, the growth of shared e-bike and bicycle systems has accelerated the evolution of transportation modes and generated extensive trajectory data on e-bike and bicycle trips, providing an important data source for related research (Chen et al., 2020; Filipe Teixeira et al., 2022). Currently, numerous studies have examined the trip features and driving mechanisms of these trips to help authorities develop more environmentally friendly and sustainable transportation policies (Gao et al., 2021; Lee et al., 2023; Wu et al., 2021). For instance, many studies have analyzed e-bike trip hotspots and cold spots (Xu et al., 2023), examined how bicycle trips connect different transportation modes (Guo et al., 2021), and explored the impact of the urban built

* Corresponding author at: State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China.
E-mail addresses: wpx@lreis.ac.cn (P. Wang), zhanghc@lreis.ac.cn (H. Zhang), zhangbb@lreis.ac.cn (B. Zhang), chengsf@lreis.ac.cn (S. Cheng), luf@lreis.ac.cn (F. Lu), zhangt@whu.edu.cn (T. Zhang).

https://doi.org/10.1016/j.cities.2025.105712
Received 29 May 2024; Received in revised form 17 September 2024; Accepted 31 December 2024
Available online 6 January 2025

environment on e-bike usage (Bi et al., 2022).

Despite extensive research on e-bike and bicycle trips, notable gaps remain. For the empirical analysis, most studies mainly concentrate on shared bicycle and e-bike trips, with less attention given to private e-bike trips (Gan et al., 2021; Yu et al., 2022). Unlike shared bicycle and e-bike trips, private e-bike trips generally reflect more stable trip purposes, such as commuting, rather than temporary ones, like leisure or shopping (Johnson et al., 2023). In many cities, private e-bikes have become essential for daily commuting, and their high numbers on the road have even disrupted normal traffic mobility, particularly during morning and evening rush hours (Liu et al., 2023). Additionally, proliferation of private e-bikes brings various infrastructure challenges, such as the need for adequate parking and charging facilities and planning for non-motorized lanes. Addressing these issues requires analyzing the trip features and driving mechanisms of private e-bike trips, with particular attention to their spatiotemporal variations. However, studies on this topic are currently quite limited. Methodologically, the relationships between trips and driving factors are typically modeled using either statistical methods, such as geographically weighted regression (GWR) (Brunsdon et al., 1996), or machine learning techniques, such as random forests (RF) (Zhang et al., 2023). The former provides clear explanations of model results but struggles to identify non-linear relationships between trips and driving factors. In contrast, the latter can model these non-linear relationships effectively, but its black-box nature makes the results difficult to interpret (Barredo Arrieta et al., 2020). Although machine learning models combined with SHapley Additive exPlanation (SHAP) techniques enhance interpretability, they still cannot match the effectiveness of GWR models in demonstrating the spatiotemporal variation effects of driving factors (Fu et al., 2023). The reason for this is that sample independence assumption in machine learning overlooks the spatiotemporal heterogeneity in trip data.

To address the above challenges, a novel framework is proposed to analyze spatiotemporal variations of private e-bike trips, focusing on how trip features and driving mechanisms vary over time and space, with specific contributions including: (1) the classical random forest is extended to establish a nonlinear mapping between driving factors and trips in the spatiotemporal domain; (2) the classical SHAP method is extended to map Shapley values onto the time and space axes, exploring the spatiotemporal variation effects of driving factors; and (3) sing the private e-bikes trips in Wuhan, China as a case study, the spatiotemporal variations of private e-bike trips are revealed.

2. Literatures

In this section, we first reviewed studies related to e-bike and bicycle trips, and then reviewed the studies related to modeling driving mechanisms.

2.1. Studies related to bicycle and e-bike trips

Recently, many studies have investigated the trip features and driving mechanisms of shared bicycle and e-bike trips across cities worldwide (Hu et al., 2021; Zhang et al., 2021; Zheng et al., 2022). Studies suggest that shared bicycle trips are predominantly for short-distance trips and effectively solve the first and last kilometer challenge in public transportation networks (Zuo et al., 2020). Moreover, urban spatial attributes and daily activities influence the spatiotemporal patterns of these trips, leading to significant spatiotemporal disparities (Tang et al., 2024). Existing findings also show that shared bicycle trips are influenced by weather conditions, socio-economic factors, spatiotemporal attributes, and the built environment (Eren & Uz, 2020). For example, a strong correlation was observed between shared bicycle trips and subway/bus stations, underscoring the critical role of shared bicycles in connecting different modes of transportation (Guo et al., 2021). Similarly, the trip features and driving mechanisms of shared e-bike trips closely mirror those of shared bicycles (Li et al., 2024). For example, shared e-bikes also provide first and last mile connectivity for public transit, showing a positive correlation with the proximity of subway and bus stations (Liu et al., 2023). In general, many studies have been conducted on shared bicycle and e-bike trips, but notable research gaps remain. More specifically, private e-bikes have now become essential for daily commuting, but their trip features and driving mechanisms are still unclear. Gaining insight into these patterns and mechanisms is crucial for effective management of private e-bikes and optimizing road infrastructure, especially with regard to their spatiotemporal variations. However, current studies on private e-bike trips are notably limited.

2.2. Studies related to modeling driving mechanisms

Trip data is a classic example of spatiotemporal data. Modeling driving mechanisms essentially involves creating a mapping between spatiotemporal data and their driving factors, typically using regression models in statistical learning (He et al., 2022; Hu et al., 2021; Hu et al., 2022; Zhu et al., 2024). Linear regression (LR) is one of the earliest methods for modeling driving mechanisms and is widely used in the field of economics (Xu et al., 2016). However, its assumption of independently and identically distributed data limits its effectiveness in modeling spatiotemporal data. The reason is that LR model does not consider the inherent spatiotemporal heterogeneity and struggles to capture spatiotemporal correlations. To address these challenges, geographically weighted regression (GWR) (Huang et al., 2010; Wu et al., 2021), multi-scale geographically weighted regression (MGWR) (Qu et al., 2023; Wu et al., 2019), and geographically and temporally weighted regression (GTWR) (Fotheringham et al., 2015; Huang et al., 2010) have been developed for modeling driving mechanisms in spatiotemporal data. Numerous studies have demonstrated that GWR, MGWR, and STGWR can uncover spatiotemporal relationships and explain the impact of driving factors on model outputs, while accounting for spatiotemporal heterogeneity. As research has advanced, many scholars have found that spatiotemporal data not only includes spatiotemporal correlations and heterogeneity but also exhibits strong nonlinear relationships (Dorosan et al., 2024; Yang et al., 2024). While statistical learning approaches like GWR can address spatiotemporal heterogeneity and correlations, they struggle with capturing nonlinear relationships within the data. Recently, machine learning models have been increasingly used for driving mechanism modeling (Zhang et al., 2023, 2024). Although these models achieve satisfactory accuracy, they lack the interpretability of statistical learning methods in explaining model results (Lundberg & Lee, 2017). Even when combined with SHAP techniques to enhance interpretability, machine learning models still fall short compared to GWR models in demonstrating the spatiotemporal variation effects of driving factors. The reason is that, like the LR model, machine learning models operate under the assumption of sample independence and overlook the spatiotemporal correlations and heterogeneity in the data.

2.3. Summary

As mentioned above, existing studies still face two gaps. First, private e-bikes have now become essential for daily commuting, but the current studies are mainly for shared e-bike and bicycle trips, leaving their trip features and driving mechanisms unclear. Second, there is a lack of exploration into the spatiotemporal variations of driving mechanisms, particularly those that model nonlinear relationships with spatiotemporal correlations and heterogeneity. Therefore, a novel framework is proposed to analyze the trip features and driving mechanisms of private e-bike trips, with particular attention given to their spatiotemporal variations.

3. Materials

3.1. Study area

Among cities in China, Wuhan ranks among the top ten cities in the country for private e-bike ownership. By 2023, the total number of private e-bikes in Wuhan has surpassed 2 million, making them one of the primary modes of transportation in the city. The chaos caused by private e-bikes also disrupts normal traffic flow in Wuhan, which is why Wuhan has been chosen as the study area. As depicted in Fig. 1, the study concentrates on the area within the Third Ring Road. Although this area covers only 6 % of the city's total land, it accommodates about 50 % of its population. Referring to Fu et al. (2023), the study region is segmented to a 200 m × 200 m grid.

3.2. Data sources

This study utilizes datasets including private e-bike trip data, traffic road network data, point of interest (POI) data, Gaode congestion data, housing price data, and population data.

The trip data of private e-bikes, extracted from their trajectory data, spans a 9-day period from December 28, 2020, to January 5, 2021. As shown in Table 1, each entry includes a unique trip identifier, a unique user identifier, the departure time, the arrival time, as well as the coordinates of both the origin and destination. In this study, >600,000 trips were used for the experiment.

In addition to the private e-bike trip data, we obtained road network data, POI data, Gaode congestion data, housing price data, and population data using web crawling technology. These datasets are utilized to construct the driving factors for e-bike trips (further discussed in Section 4.2). Among them, the road network, POI, and Gaode congestion data are sourced from the Gaode Open Platform, the housing price data from Fang.com, and the population data from the WorldPop grid. Detailed information about the data and preprocessing can be found in Appendix A and B.

4. Proposed framework

Fig. 2 outlines the proposed framework, structured into three key components: (1) features and spatiotemporal variations of private e-bike trips, (2) driving mechanism modeling in spatiotemporal domain, (3) driving mechanism analysis considering spatiotemporal heterogeneity, as discussed in Sections 4.1–4.3. First, the trip features of private e-bike trips are defined mathematically, and their spatiotemporal variations are analyzed from both macro and micro perspectives. Second, a novel spatiotemporal random forest is presented to build a nonlinear mapping between driving factors and private e-bike trips in the spatiotemporal domain. Finally, the classical SHAP method is extended to map Shapley values onto the time and space axes, enabling the exploration of spatiotemporal variations in driving factors.

Table 1
Examples of private e-bike trip data.

Trip ID	User ID	Departure time	Arrival time	Origin	Destination
1	364,851	2020/12/28 12:15	2020/12/28 12:19	Shape (Point)	Shape (Point)
2	486,214	2020/12/28 08:16	2020/12/28 08:26	Shape (Point)	Shape (Point)
3	364,851	2020/12/29 19:18	2020/12/28 19:32	Shape (Point)	Shape (Point)
……	……	……	……	……	……
684,214	354,785	2021/01/05 14:24	2021/01/05 14:42	Shape (Point)	Shape (Point)

4.1. Features and spatiotemporal variations of private e-bike trips

Trip features are statistical metrics that describe e-bike trips, and we mainly discuss three trip features, including trip frequency, trip distance, and duration time, with their corresponding calculation methods shown in Eqs. (1)–(3).

$$Frequency_i^t = \sum_{k=1}^{N} |\{1 : od_k \in grid_i \wedge od_k \in window_t\}| \tag{1}$$

$$Distance_k = \mathrm{PathDistance}(od_k) \tag{2}$$

$$Duration_k = \mathrm{PathDuration}(od_k) \tag{3}$$

where $grid_i$ represents the ith spatial grid, and $window_t$ represents the tth time window. In this study, the size of the spatial grid is 200 m × 200 m, and the time window is 60 min. od_k represents the kth private e-bike trip, and N represents the total number of trips. According to Eqs. (1)–(3), $Frequency_i^t$ represents the total number of trips in ith spatial grid during tth time window, $Distance_k$ denotes the distance of a trip, referring to the path length from the origin to the destination, and $Duration_k$ represents the duration time of a trip, indicating the time spent from the origin to the destination.

After defining the trip features, we mainly study the spatiotemporal

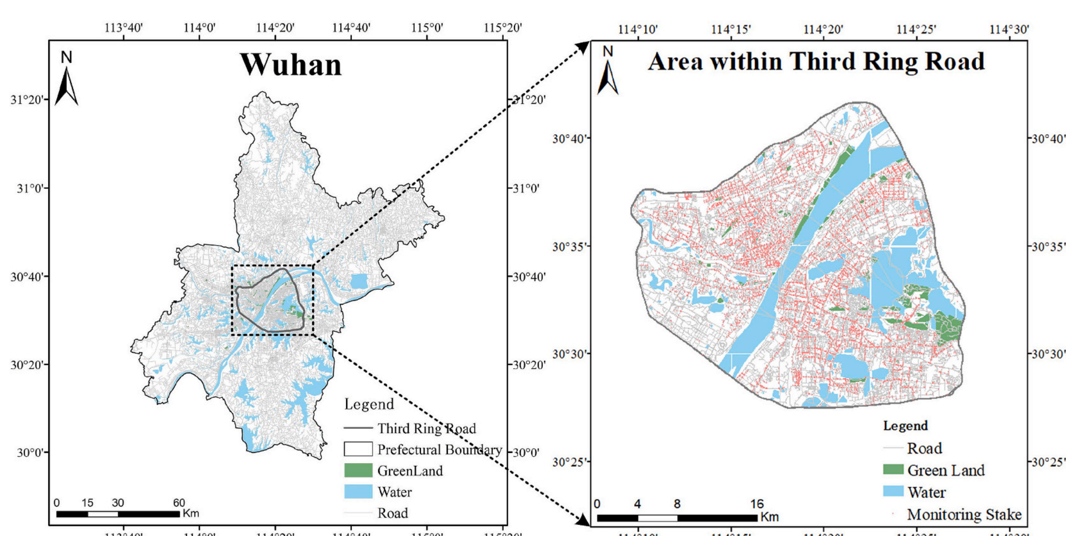

Fig. 1. Illustration of the study area.

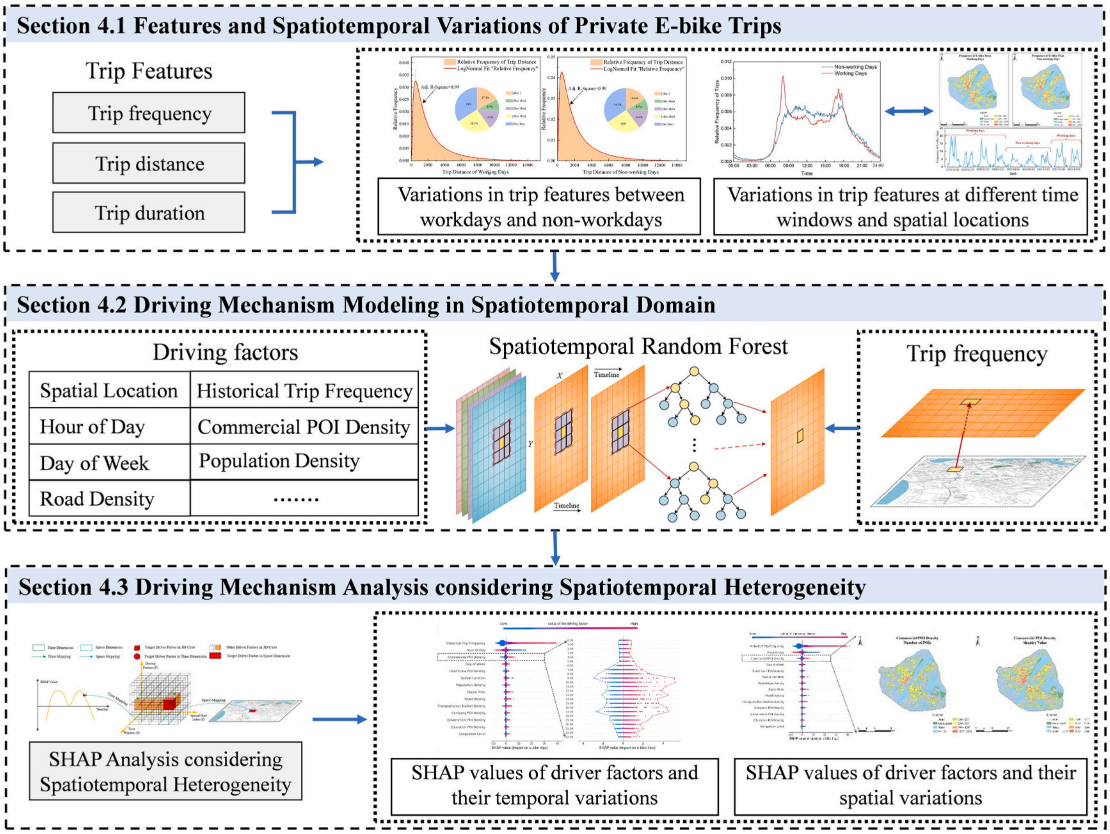

Fig. 2. Overview of proposed framework.

variations of e-bike trips from two perspectives: (1) Exploring the variations of trip features between workdays and non-workdays on a macro-scale; (2) Exploring the variations of trip features in different time windows and spatial positions on a micro-scale.

4.2. *Driving mechanism modeling in spatiotemporal domain*

The RF model is a traditional machine learning approach with strong non-linear fitting capability, making them effective for modeling urban trip mobility (Guidon et al., 2020). Despite achieving satisfactory accuracy, the RF model has certain limitations. Specifically, most existing methods are basic applications of classical random forests and do not account for the spatiotemporal relationships in trip data (Zhang et al., 2023). The assumption of independent and identically distributed samples hinders classical random forests from capturing the complex underlying mechanisms and interactions in the data. Therefore, the classical random forest is extended to propose the spatiotemporal random forest (STRF).

As depicted in Fig. 3, the STRF model extends the classical RF model by incorporating a neighborhood context. It utilizes both internal and external driving factors within the neighborhood to predict future trip frequency. Internal driving factors refer to the historical trip frequencies within the target grid. External driving factors refer to the urban built environment, spatial location, time windows, and other information within the target grid. Table 2 provides a detailed overview of the 14 driving factors used in the STRF model. Additionally, the driving factors

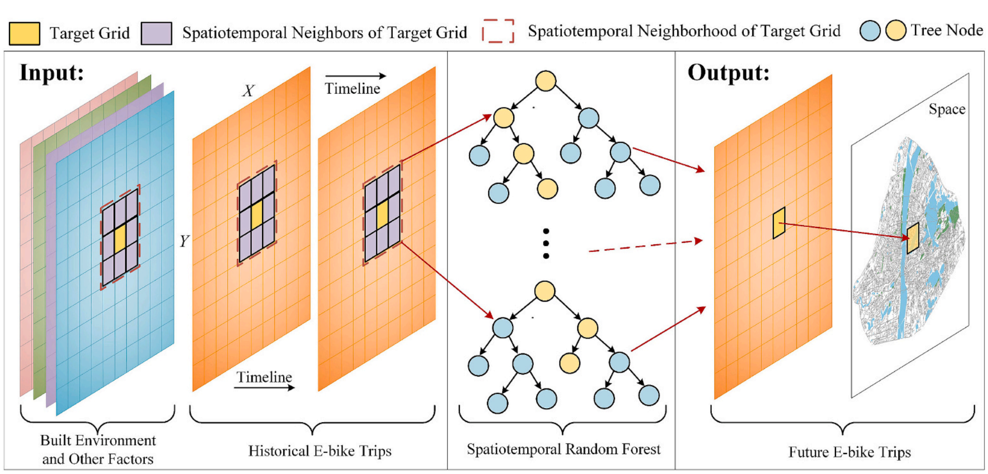

Fig. 3. Illustration of spatiotemporal random forest.

Table 2
Driving factor in spatiotemporal random forests.

Factors	Description
Historical Trip Frequency	Historical trip frequency of the nearest grid to the target grid
Spatial Location	Indexing of the spatial grid
Hour of Day	Hour in a day (0–23)
Day of Week	which day of the week (1–7)
Road Density	Total road length
Population Density	Population calculated through WorldPop data
Commercial POI Density	Including Dining, Shopping, and Service POIs
Education POI Density	Including School, College, and University
Company POI Density	Including Corporate Headquarter and Factory
Healthcare POI Density	Including Clinic, Pharmacy, and Hospital
Government POI Density	Including Townhall and Government Office
Transportation Station Density	Including Bus and Metro Stations
Congestion Level	Congestion duration of the target grid
House Price	Average house price of the target grid

of the target grid are integrated with those from neighboring grids to ensure a one-to-one correspondence between the driving factors and trip frequency. Specifically, the integration process is outlined in Eq. (4).

$$\widehat{f}_i^k = \begin{cases} \sum\limits_{j\in\Omega_i} f_j^k & 5 \le k \le 13 \\ \dfrac{\sum\limits_{j\in\Omega_i} \frac{1}{d_{ij}^2} f_j^k}{\sum\limits_{j\in\Omega_i} \frac{1}{d_{ij}^2}} & k = 14 \end{cases} \tag{4}$$

where f_i^k and $\widehat{f}_i^k$ represent the kth driving factor of the target grid before and after fusion, with the first driving factor being the Historical Trip Frequency in Table 1 and the fourteenth driving factor being the House Price in Table 1; Ω_i indicates the set of spatiotemporal neighbors of target grid; d_{ij}^2 indicates the distance between the target and neighboring grids. According to Eq. (4), for the 5th to 13th driving factors, the fused driving factors are the sum of the driving factors in the neighborhood. For the 14th driver, the fused driver is the weighted average of the driver factors in the neighborhood.

4.3. Driving mechanism analysis considering spatiotemporal heterogeneity

Although the nonlinear relationship between driving factors and trips has been established in the spatiotemporal domain, the specific impact of these factors remains unclear. The SHAP method can be utilized to clarify the contributions of individual driving factors to private e-bike trips (Lundberg & Lee, 2017). However, the current SHAP method primarily reveals the impact weight of driving factors but does not account for variations in these factors across different time windows and spatial locations (Zhang et al., 2024). Therefore, an extension of the classical SHAP approach is proposed to explore the spatiotemporal variation effects of driving factors on private e-bike trips.

As shown in Fig. 4, we established a three-dimensional cube with the spatial grid index as the S-axis, the time window as the T-axis, and the driving factor as the F-axis. For any driving factor in the three-dimensional cube, we can calculate the Shapley value of the driving factor based on game theory principles, and map the Shapley value to the corresponding spatial location and time window. After mapping all driving factors, we can analyze how the driving factors change over different time windows and spatial locations. The Shapley value of driving factors in the three-dimensional cube can be calculated by Eq. (5).

$$STRF\left(\left\{\widehat{f}_i^1, \widehat{f}_i^2, \dots, \widehat{f}_i^{14}\right\}\right) = \frac{\sum_{j=1}^{N} STRF\left(\left\{\widehat{f}_j^1, \widehat{f}_j^2, \dots, \widehat{f}_j^{14}\right\}\right)}{N} + \sum_{k=1}^{14} SHAP\left(\widehat{f}_i^k\right)$$

$$\downarrow$$

$$\downarrow$$

$$STRF\left(\left\{\widehat{f}_{s,t}^1, \widehat{f}_{s,t}^2, \dots, \widehat{f}_{s,t}^{14}\right\}\right) = \frac{\sum_{j=1}^{N_S\times N_T} STRF\left(\left\{\widehat{f}_j^1, \widehat{f}_j^2, \dots, \widehat{f}_j^{14}\right\}\right)}{N_S \times N_T} + \sum_{k=1}^{14} SHAP\left(\widehat{f}_{s,t}^k\right) \tag{5}$$

where $STRF$ stands for spatiotemporal random forest; $\widehat{f}_i^k$ represents the kth driver of the target grid; N represents the total number of time windows and spatial grids, decomposed into $N_S \times N_T$; $\widehat{f}_{s,t}^k$ represents the driving factor in the three-dimensional cube, obtained by the decomposition of $\widehat{f}_i^k$; $SHAP\left(\widehat{f}_{s,t}^i\right)$ represents the Shapley value of the driving factor for sth grid during tth time window. In this study, the impact direction of driving factors on trips is determined by examining the positive and negative Shapley values. Among them, $SHAP\left(\widehat{f}_{s,t}^i\right) > 0$ represents a positive effect, while $SHAP\left(\widehat{f}_{s,t}^i\right) < 0$ signifies a negative effect.

5. Case study

5.1. Spatiotemporal variations of trip features

Fig. 5 illustrates the differences of trip distance and duration on working and non-working days. It reveals a clear lognormal distribution for trip distance and a Hill distribution for trip duration on both types of

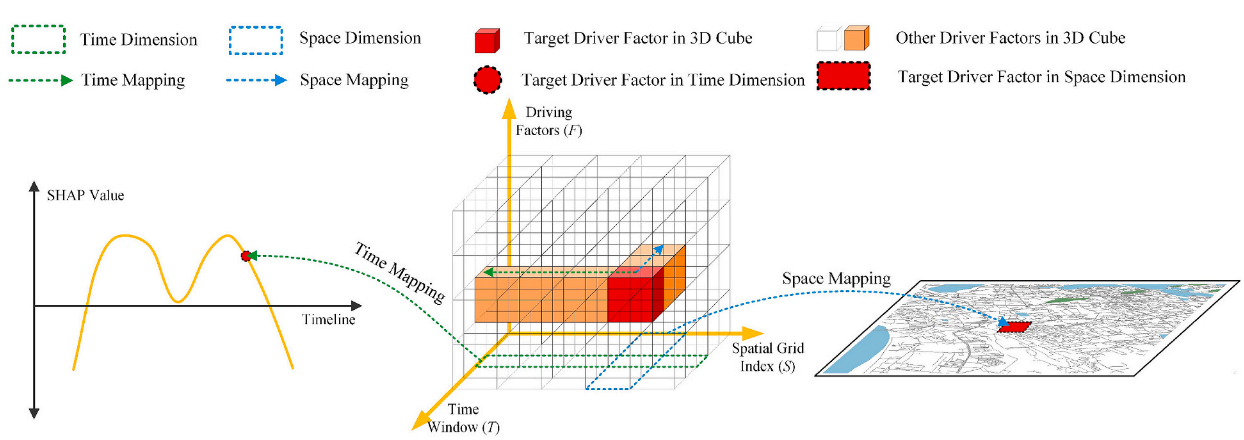

Fig. 4. Spatial mapping and temporal mapping of Shapley values.

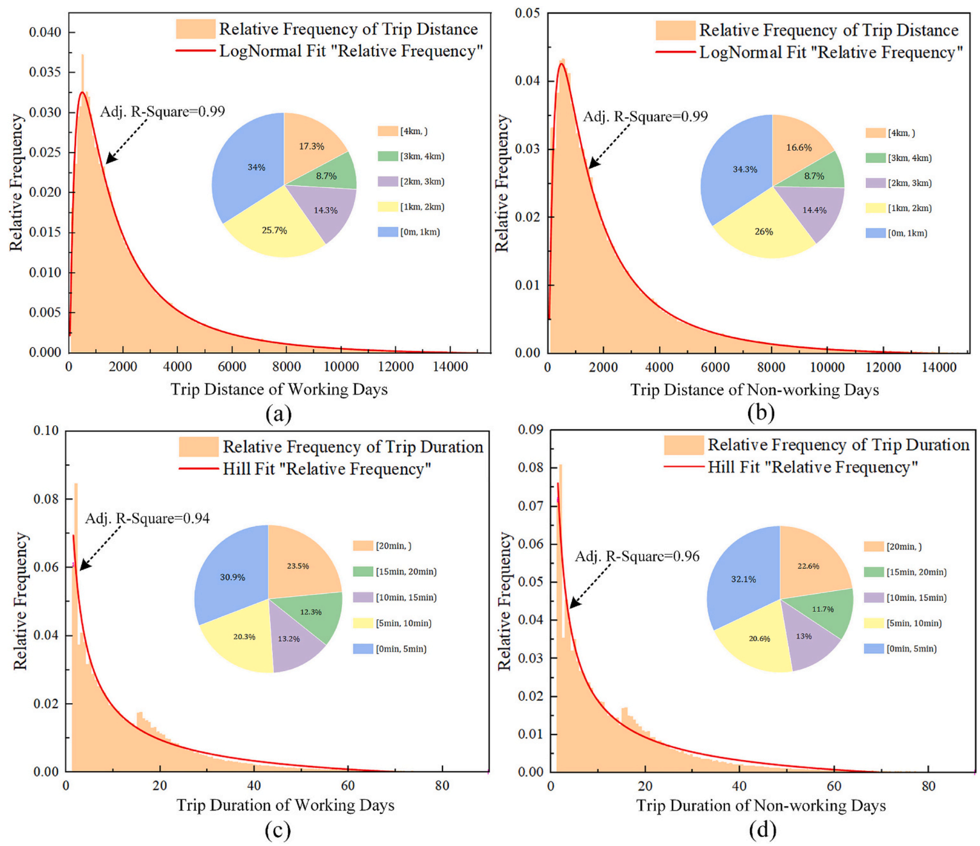

Fig. 5. Illustration of trip distance and duration: (a) trip distance of working days, (b) trip distance of non-working days, (c) trip duration of working days, and (d) trip duration of non-working days.

days. Apart from the statistical distributions, there are subtle differences in trip distance and duration between working and non-working days. Specifically, trips on non-working days tend to be slightly shorter in both distance and duration compared to those on working days. For instance, the proportion of trips with distances <2 km increases from 59.7 % on working days to 60.3 % on non-working days, and the proportion of trips with durations <10 min rises from 51.2 % on working days to 52.7 % on non-working days. Additionally, compared to existing studies on shared e-bikes and bicycle, private e-bike trips tend to have longer trip distances and durations. On working days, 48.7 % of trips have distances between 1 and 4 km, and 45.8 % have durations between 5 and 20 min. On non-working days, 49.1 % of trips have distances between 1 and 4 km, and 45.3 % have durations between 5 and 20 min. The above results emphasize the crucial role that private e-bikes play in short- and medium-distance trips.

Fig. 6 shows the differences in trip frequency between working and non-working days, revealing a clear bimodal distribution on working days, with peaks in the morning (7:30 to 9:00) and evening (16:00 to 18:30). In contrast, non-working days show no distinct morning or evening peaks, indicating that private e-bikes are crucial for urban commuting. Additionally, Fig. 7 presents the differences in trip frequency across different time windows and spatial locations. The results reveal significant heterogeneity in e-bike trips both spatially and temporally. Spatially, the hotspots for private e-bike trips are predominantly located within the city center. Temporally, these hotspots occur during the morning and evening peaks on working days, while on non-working days, they are more evenly distributed throughout the day. The spatiotemporal heterogeneity depicted in Fig. 7 indirectly underscores the necessity of establishing the STRF model.

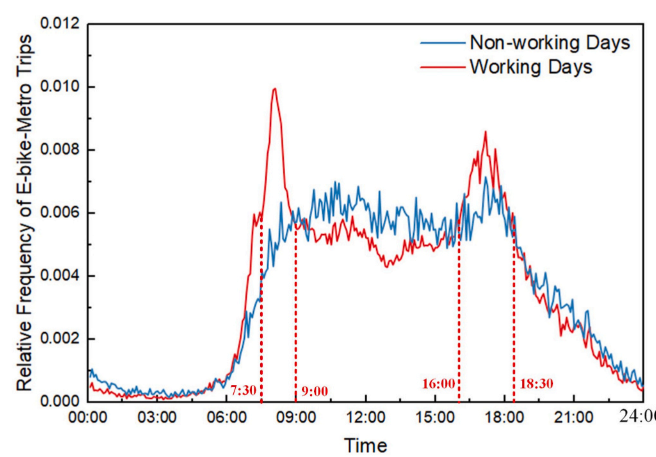

Fig. 6. Relative trip frequency on working and non-working days.

5.2. *Driving mechanism modeling and analysis considering spatiotemporal heterogeneity*

5.2.1. *Fitting precision analysis*

In this subsection, the fitting precision of the STRF model was evaluated by comparing it with baseline models such as LR, XGBoost, and RF models. Additionally, ablation studies were performed to assess the validity of the model design. For instance, temporal random forest (TRF) and spatial random forest (SRF) were employed to evaluate the rationale for incorporating time and spatial relationships. Fig. 8 illustrates the fitting precision of the baselines, as well as the TRF and SRF models. It was found that XGBoost and RF models significantly

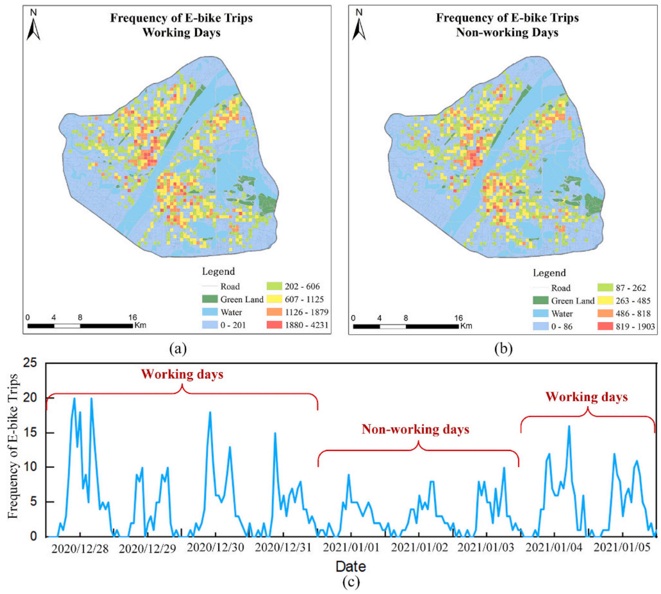

Fig. 7. Spatiotemporal distribution of trip frequency: (a) spatial distribution of trip frequency during working days, (b) spatial distribution of trip frequency during non-working days, and (c) temporal distribution of trip frequency during working days and non-working days.

outperformed the LR model, suggesting the presence of nonlinear relationships between trips and their driving factors. Furthermore, the fitting precision of the SRF and TRF models exceeds that of the classical RF model, highlighting the effectiveness of integrating temporal and spatial relationships into the classical RF approach. Finally, the fitting precision of the STRF outperforms that of the SRF and TRF, highlighting the necessity of simultaneously modeling spatiotemporal relationships. Aside from fitting precision, Table 3 further shows the stability of the STRF across different random seeds. The results indicate that the STRF model not only achieves high fitting precision but also demonstrates consistent stability, further highlighting its advantages.

5.2.2. Spatiotemporal variations of driver factors

Fig. 9(a) illustrates the impact weights of various factors on trips, based on the driving mechanism modeling. Results indicate a strong positive correlation between trips and factors like Commercial POI Density, Healthcare POI Density, Transportation Station Density, Congestion Level, and House Price. Notably, higher congestion levels are associated with more frequent private e-bike trips, which could be attributed to two reasons. First, private e-bikes provide a more convenient transportation option in areas with severe congestion. Second, the increased activity of e-bikes on the roads may displace space that would otherwise be used by other vehicles or pedestrians.

Table 3
Fitting precision (mean ± std) of spatiotemporal random forest and baselines.

Models	R^2	Adj.R^2
LR	0.2200 ± 0.0045	0.2199 ± 0.0045
XGBoost	0.5007 ± 0.0073	0.4989 ± 0.0073
RF	0.5137 ± 0.0066	0.5154 ± 0.0066
SRF	0.5237 ± 0.0073	0.5216 ± 0.0073
TRF	0.8489 ± 0.0052	0.8469 ± 0.0052
STRF	**0.8601** ± 0.0052	**0.8601** ± 0.0052

Bold indicates the results of our proposed model.

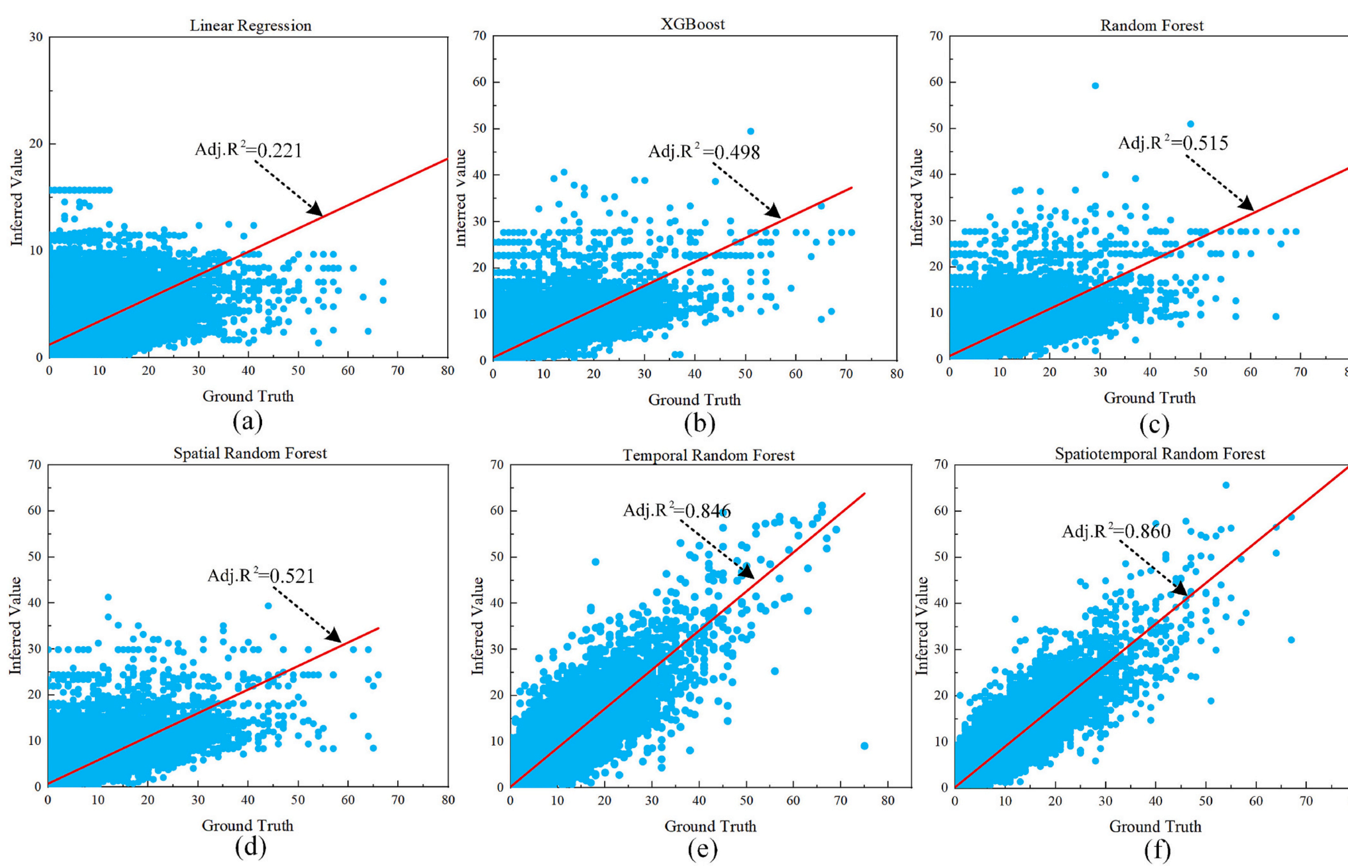

Fig. 8. Comparison results of spatiotemporal random forest with baselines: (a) Linear Regression, (b) XGBoost, (c) Random Forest, (d) Spatial Random Forest, (e) Temporal Random Forest, and (f) Spatiotemporal Random Forest.

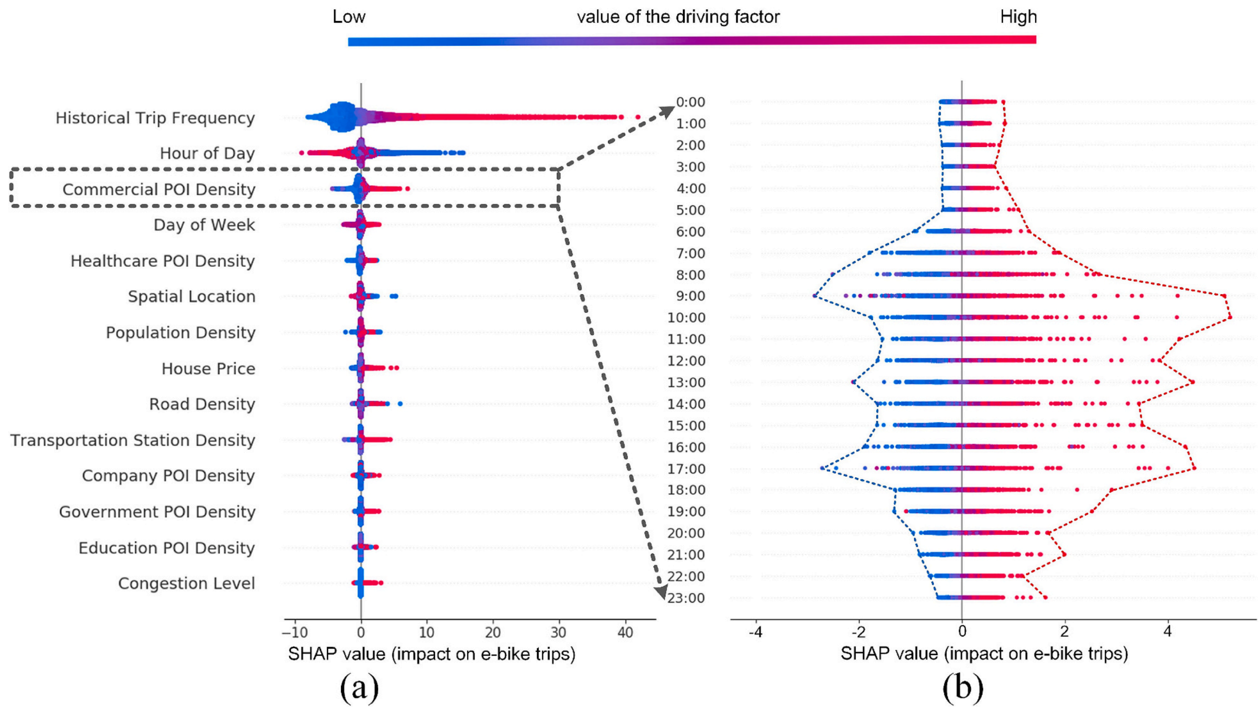

Fig. 9. Impact weights of driver factors and their temporal variations: (a) SHAP values for all driving factors, and (b) SHAP value of Commercial POI Density in temporal dimension.

Compared to the classical SHAP method, the improved SHAP method offers the advantage of revealing spatiotemporal variations in impact weights. In the temporal dimension, as shown in Fig. 9(b), the impact weight of Commercial POI Density on trips exhibits a distinct bimodal distribution in the temporal dimension, correlating with the trip frequency shown in Fig. 6. The result indicates that the effect of driving factors on trips is more pronounced during peak hours compared to non-peak hours.

In the spatial dimension, Fig. 10 shows how these weights vary across different spatial locations. It is worth noting that the improved SHAP method can reveal phenomena that are difficult to detect with the classic SHAP method. For example, when both high and low population densities influence e-bike trips, the classical SHAP method struggles to clarify the relationship between Population Density and private e-bike trips, as illustrated in Fig. 10(a). In contrast, Fig. 10(b) and (c) not only illustrate the spatiotemporal patterns associated with Population Density but also offer a detailed distribution of impact weights across various spatial locations. Furthermore, we observed clustering patterns in impact weights across different spatial areas. These findings suggest that despite the relatively low importance ranking of Spatial Locations in Fig. 10(a), they remain significant for private e-bike trips.

6. Discussion

Understanding the spatiotemporal variations of private e-bike trips is crucial for improving traffic mobility, reducing pollution, and enhancing road safety. In this study, a novel framework is proposed to analyze spatiotemporal variations of private e-bike trips, with a focus on the spatiotemporal variations of trip features and driving mechanisms.

Empirical insights reveal that, unlike shared e-bikes and bicycles used primarily for short-distance or connecting trips (Ji et al., 2022; Xu et al., 2023), private e-bikes have become crucial for daily urban commuting and play a significant role in short- to medium-distance trips. This finding suggests that governments may strategically plan private e-bike parking areas and charging stations in high-density trip zones to enhance commuting convenience and improve public satisfaction, particularly on working days. Additionally, the factors that impact private e-bike trips closely resemble those affecting shared e-bike trips (Eren & Uz, 2020; Wu et al., 2021; Yang et al., 2024). Specifically, private e-bike trips have a positive correlation with Commercial POI Density, Healthcare POI Density, Housing Prices, and Congestion Levels. Notably, higher congestion levels are correlated with more frequent private e-bike trips, implying that while private e-bikes offer convenience in congested areas (Yu et al., 2022), they can also disrupt normal traffic mobility (Liu et al., 2023). To mitigate this, government authorities could conduct field surveys in highly congested areas and designate specific lanes for private e-bikes where they impede traffic, thereby improving overall road efficiency.

Methodological insights highlight the necessity for developing nonlinear, explainable data-driven techniques that account for spatiotemporal heterogeneity. Compared to statistical models such as GWR and GTWR (Brunsdon et al., 1996; Fotheringham et al., 2015), the proposed framework not only uncovers spatiotemporal variations in the impact weights of driving factors but also captures the nonlinear relationships within the data. The framework reveals that the impact of driving factors varies significantly across different time windows and spatial locations. For example, driving factors have a greater influence on e-bike travel during peak hours than off-peak periods, and their influence weights exhibit spatial clustering effects.

This study has some limitations. First, it focuses solely on the OD flow of private e-bikes, leaving the full trip chain of their journeys unestablished. In the future, the full trip chain will be utilized to gain deeper insights into private e-bike trip patterns. Second, the proposed STRF method and the improved SHAP method offer a universal explainable data-driven framework, but our analysis and validation have been limited to private e-bike trips. In the future, the proposed framework will be applied to modeling other trip mechanisms, such as shared bicycle trips, bus journeys, and subway rides.

7. Conclusion

In this study, using Wuhan, China as a case study, a novel explainable

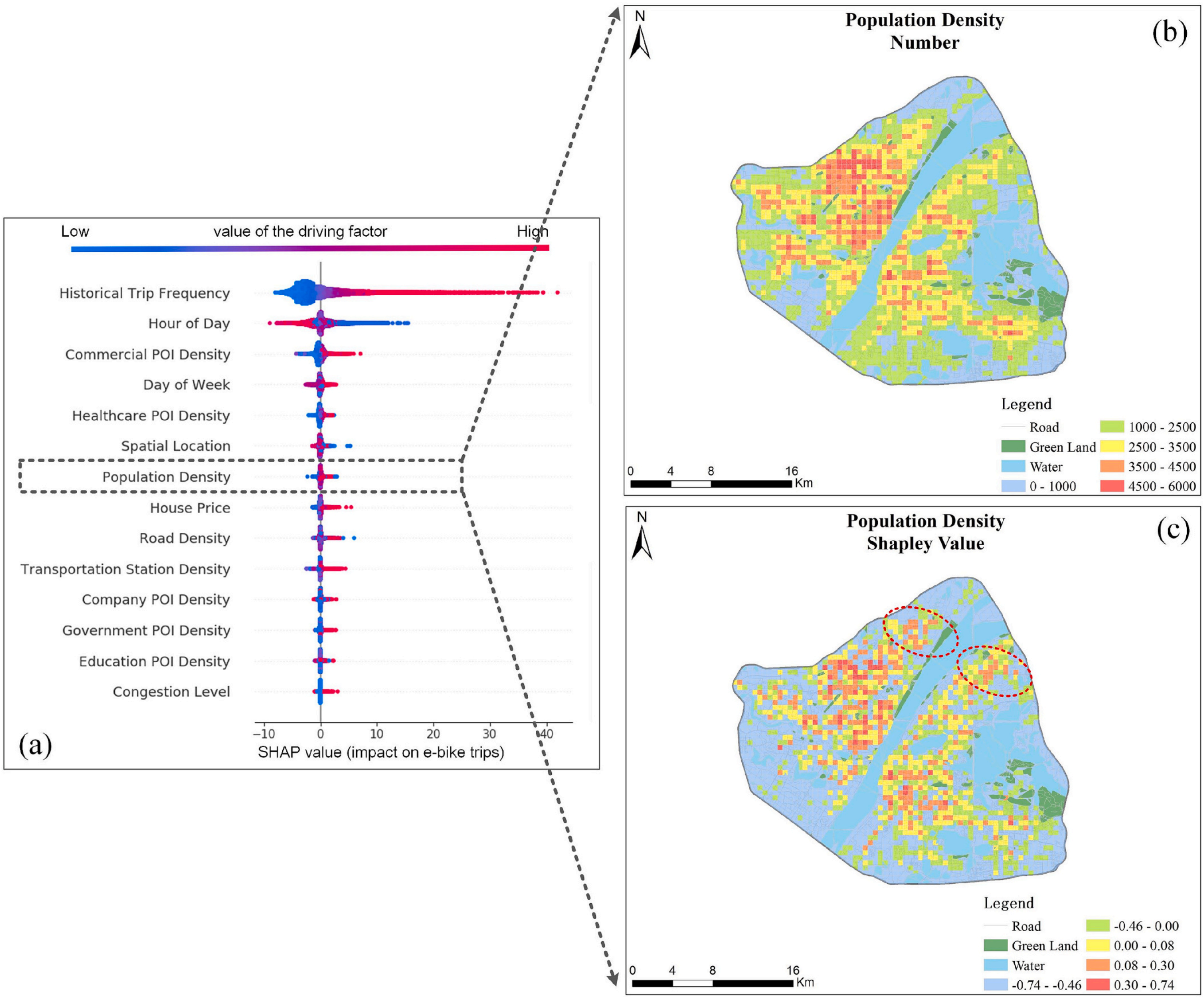

Fig. 10. Impact weights of drivers and their spatial variations: (a) SHAP values for all driving factors, (b) number of Commercial POIs in spatial dimension, and (c) SHAP value of Commercial POI Density in spatial dimension.

framework is proposed to analyze the spatiotemporal variations in trip features and driving mechanisms of private e-bike trips. More specifically, a novel spatiotemporal random forest is presented to build a nonlinear mapping between driving factors and private e-bike trips in the spatiotemporal domain. Then, the classical SHAP method is extended to map Shapley values onto the time and space axes, enabling the exploration of spatiotemporal variations in driving factors. Findings reveal that: (1) private e-bikes are frequently used for short and medium-distance trips, typically exceeding 1 km, and play a crucial role in daily urban commuting; (2) Factors such as Historical trip frequency, Commercial POI Density, and Hospital POI Density exhibit strong positive correlations with private e-bike trips; (3) the influence of driving factors on private e-bike trips vary significantly across different spatial locations and time windows. This study offers an innovative analytical framework for a more profound comprehension of e-bike trips. Additionally, the findings can aid authorities in crafting more effective policies and planning strategies.

CRediT authorship contribution statement

Peixiao Wang: Writing – original draft, Funding acquisition, Formal analysis, Conceptualization. **Hengcai Zhang:** Writing – review & editing, Writing – original draft, Project administration, Methodology, Funding acquisition. **Beibei Zhang:** Writing – review & editing, Formal analysis. **Shifen Cheng:** Writing – review & editing. **Feng Lu:** Writing – review & editing, Project administration. **Tong Zhang:** Writing – review & editing.

Funding

This project was supported by National Key Research and Development Program of China [Grant No. 2023YFB3906804]; National Natural Science Foundation of China [Grant Nos. 42401524, 42371469 and 42371470]; China National Postdoctoral Support Program for Innovative Talents [Grant No. BX20230360]; China Postdoctoral Science Foundation [Grant No. 2023M743454]; Open Fund of State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University [Grant No. 23I03]; Innovation Project of LREIS [Grant No. 08R8A092YA].

Declaration of competing interest

The authors declare no conflicts of interest.

Acknowledgments

The numerical calculations in this paper have been done on the supercomputing system in the Supercomputing Center of Wuhan University.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cities.2025.105712.

Data availability

Data will be made available on request.

References

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Bi, H., Ye, Z., & Zhu, H. (2022). Examining the nonlinear impacts of built environment on ridesourcing usage: Focus on the critical urban sub-regions. *Journal of Cleaner Production, 350*, Article 131314. https://doi.org/10.1016/j.jclepro.2022.131314

Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis, 28*(4), 281–298. https://doi.org/10.1111/j.1538-4632.1996.tb00936.x

Chen, W., Liu, Q., Zhang, C., Mi, Z., Zhu, D., & Liu, G. (2020). Characterizing the stocks, flows, and carbon impact of dockless sharing bikes in China. *Resources, Conservation and Recycling, 162*, Article 105038. https://doi.org/10.1016/j.resconrec.2020.105038

Dorosan, M., Dailisan, D., Valenzuela, J. F., & Monterola, C. (2024). Use of machine learning in understanding transport dynamics of land use and public transportation in a developing city. *Cities, 144*, Article 104587. https://doi.org/10.1016/j.cities.2023.104587

Eren, E., & Uz, V. E. (2020). A review on bike-sharing: The factors affecting bike-sharing demand. *Sustainable Cities and Society, 54*, Article 101882. https://doi.org/10.1016/j.scs.2019.101882

Filipe Teixeira, J., Silva, C., & Moura e Sá, F. (2022). The role of bike sharing during the coronavirus pandemic: An analysis of the mobility patterns and perceptions of Lisbon's *GIRA* users. *Transportation Research Part A: Policy and Practice, 159*, 17–34. https://doi.org/10.1016/j.tra.2022.03.018

Fotheringham, A. S., Crespo, R., & Yao, J. (2015). Geographical and temporal weighted regression (GTWR). *Geographical Analysis, 47*(4), 431–452. https://doi.org/10.1111/gean.12071

Fu, C., Huang, Z., Scheuer, B., Lin, J., & Zhang, Y. (2023). Integration of dockless bike-sharing and metro: Prediction and explanation at origin-destination level. *Sustainable Cities and Society, 99*, Article 104906. https://doi.org/10.1016/j.scs.2023.104906

Gan, Z., Yang, M., Zeng, Q., & Timmermans, H. J. P. (2021). Associations between built environment, perceived walkability/bikeability and metro transfer patterns. *Transportation Research Part A: Policy and Practice, 153*, 171–187. https://doi.org/10.1016/j.tra.2021.09.007

Gao, K., Yang, Y., Li, A., Li, J., & Yu, B. (2021). Quantifying economic benefits from free-floating bike-sharing systems: A trip-level inference approach and city-scale analysis. *Transportation Research Part A: Policy and Practice, 144*, 89–103. https://doi.org/10.1016/j.tra.2020.12.009

Guidon, S., Reck, D. J., & Axhausen, K. (2020). Expanding a(n) (electric) bicycle-sharing system to a new city: Prediction of demand with spatial regression and random forests. *Journal of Transport Geography, 84*, Article 102692. https://doi.org/10.1016/j.jtrangeo.2020.102692

Guo, Y., Yang, L., Lu, Y., & Zhao, R. (2021). Dockless bike-sharing as a feeder mode of metro commute? The role of the feeder-related built environment: Analytical framework and empirical evidence. *Sustainable Cities and Society, 65*, Article 102594. https://doi.org/10.1016/j.scs.2020.102594

He, M., Ma, X., Wang, J., & Zhu, M. (2022). Geographically weighted multinomial logit models for modelling the spatial heterogeneity in the bike-sharing renting-returning imbalance: A case study on Nanjing, China. *Sustainable Cities and Society, 83*, Article 103967. https://doi.org/10.1016/j.scs.2022.103967

Hu, B., Zhong, Z., Zhang, Y., Sun, Y., Jiang, L., Dong, X., & Sun, H. (2022). Understanding the influencing factors of bicycle-sharing demand based on residents' trips. *Physica A: Statistical Mechanics and its Applications, 586*, Article 126472. https://doi.org/10.1016/j.physa.2021.126472

Hu, H., Chen, Q., Qian, Q., Lin, C., Chen, Y., & Tian, W. (2021). Impacts of traffic and street characteristics on the exposure of cycling commuters to PM2.5 and PM10 in urban street environments. *Building and Environment, 188*, Article 107476. https://doi.org/10.1016/j.buildenv.2020.107476

Huang, B., Wu, B., & Barry, M. (2010). Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *International Journal of Geographical Information Science, 24*(3), 383–401. https://doi.org/10.1080/13658810802672469

Ji, S., Wang, X., Lyu, T., Liu, X., Wang, Y., Heinen, E., & Sun, Z. (2022). Understanding cycling distance according to the prediction of the XGBoost and the interpretation of SHAP: A non-linear and interaction effect analysis. *Journal of Transport Geography, 103*, Article 103414. https://doi.org/10.1016/j.jtrangeo.2022.103414

Johnson, N., Fitch-Polse, D. T., & Handy, S. L. (2023). Impacts of e-bike ownership on travel behavior: Evidence from three northern California rebate programs. *Transport Policy, 140*, 163–174. https://doi.org/10.1016/j.tranpol.2023.06.014

Lee, S., Ko, E., Jang, K., & Kim, S. (2023). Understanding individual-level travel behavior changes due to COVID-19: Trip frequency, trip regularity, and trip distance. *Cities, 135*, Article 104223. https://doi.org/10.1016/j.cities.2023.104223

Li, Q., Zhang, E., Luca, D., & Fuerst, F. (2024). The travel pattern difference in dockless micro-mobility: Shared e-bikes versus shared bikes. *Transportation Research Part D: Transport and Environment, 130*, Article 104179. https://doi.org/10.1016/j.trd.2024.104179

Liu, S., Zhang, F., Ji, Y., Ma, X., Liu, Y., Li, S., & Zhou, X. (2023). Understanding spatial-temporal travel demand of private and shared e-bikes as a feeder mode of metro stations. *Journal of Cleaner Production, 398*, Article 136602. https://doi.org/10.1016/j.jclepro.2023.136602

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*. https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html.

Lv, Y., Zhou, L., Yao, G., & Zheng, X. (2021). Detecting the true urban polycentric pattern of Chinese cities in morphological dimensions: A multiscale analysis based on geospatial big data. *Cities, 116*, Article 103298. https://doi.org/10.1016/j.cities.2021.103298

Qu, X., Xiao, X., Zhu, X., Wu, H., Gong, J., & Li, D. (2023). 6Ds-HLM: Hierarchical models of road traffic crash influencing factor analysis considering the urban built environment of intersection visibility. *International Journal of Applied Earth Observation and Geoinformation, 123*, Article 103472. https://doi.org/10.1016/j.jag.2023.103472

Ren, F., Qiu, Z., Liu, Z., Bai, H., & Gao, H. O. (2023). Trees help reduce street-side air pollution: A focus on cyclist and pedestrian exposure risk. *Building and Environment, 229*, Article 109923. https://doi.org/10.1016/j.buildenv.2022.109923

Schläpfer, M., Dong, L., O'Keeffe, K., Santi, P., Szell, M., Salat, H., … West, G. B. (2021). The universal visitation law of human mobility. *Nature, 593*(7860), Article 7860. https://doi.org/10.1038/s41586-021-03480-9

Tang, J. H. C. G., Huang, Y., Zhu, Y., Yang, X., & Zhuge, C. (2024). The association between travel demand of docked bike-sharing and the built environment: Evidence from seven US cities. *Sustainable Cities and Society, 105325*. https://doi.org/10.1016/j.scs.2024.105325

Wu, C., Kim, I., & Chung, H. (2021). The effects of built environment spatial variation on bike-sharing usage: A case study of Suzhou, China. *Cities, 110*, Article 103063. https://doi.org/10.1016/j.cities.2020.103063

Wu, C., Ren, F., Hu, W., & Du, Q. (2019). Multiscale geographically and temporally weighted regression: Exploring the spatiotemporal determinants of housing prices. *International Journal of Geographical Information Science, 33*(3), 489–511. https://doi.org/10.1080/13658816.2018.1545158

Xu, B., Luo, L., & Lin, B. (2016). A dynamic analysis of air pollution emissions in China: Evidence from nonparametric additive regression models. *Ecological Indicators, 63*, 346–358. https://doi.org/10.1016/j.ecolind.2015.11.012

Xu, X., Wang, J., Poslad, S., Rui, X., Zhang, G., & Fan, Y. (2023). Exploring intra-urban human mobility and daily activity patterns from the lens of dockless bike-sharing: A case study of Beijing, China. *International Journal of Applied Earth Observation and Geoinformation, 122*, Article 103442. https://doi.org/10.1016/j.jag.2023.103442

Yang, W., Fei, J., Li, Y., Chen, H., & Liu, Y. (2024). Unraveling nonlinear and interaction effects of multilevel built environment features on outdoor jogging with explainable machine learning. *Cities, 147*, Article 104813. https://doi.org/10.1016/j.cities.2024.104813

Yu, Y., Jiang, Y., Qiu, N., Guo, H., Han, X., & Guo, Y. (2022). Exploring built environment factors on e-bike travel behavior in urban China: A case study of Jinan. *Frontiers in Public Health, 10*. https://www.frontiersin.org/articles/10.3389/fpubh.2022.1013421.

Zhang, B., Cheng, S., Wang, P., & Lu, F. (2024). Inferring freeway traffic volume with spatial interaction enhanced betweenness centrality. *International Journal of Applied Earth Observation and Geoinformation, 129*, Article 103818. https://doi.org/10.1016/j.jag.2024.103818

Zhang, B., Cheng, S., Zhao, Y., & Lu, F. (2023). Inferring intercity freeway truck volume from the perspective of the potential destination city attractiveness. *Sustainable Cities and Society, 98*, Article 104834. https://doi.org/10.1016/j.scs.2023.104834

Zhang, H., Zhuge, C., Jia, J., Shi, B., & Wang, W. (2021). Green travel mobility of dockless bike-sharing based on trip data in big cities: A spatial network analysis. *Journal of Cleaner Production, 313*, Article 127930. https://doi.org/10.1016/j.jclepro.2021.127930

Zheng, L., Meng, F., Ding, T., Yang, Q., Xie, Z., & Jiang, Z. (2022). The effect of traffic status on dockless bicycle-sharing: Evidence from Shanghai, China. *Journal of Cleaner Production, 381*, Article 135207. https://doi.org/10.1016/j.jclepro.2022.135207

Zhou, Y., Yu, Y., Wang, Y., He, B., & Yang, L. (2023). Mode substitution and carbon emission impacts of electric bike sharing systems. *Sustainable Cities and Society, 89*, Article 104312. https://doi.org/10.1016/j.scs.2022.104312

Zhu, B., Hu, S., Kaparias, I., Zhou, W., Ochieng, W., & Lee, D.-H. (2024). Revealing the driving factors and mobility patterns of bike-sharing commuting demands for integrated public transport systems. *Sustainable Cities and Society, 104*, Article 105323. https://doi.org/10.1016/j.scs.2024.105323

Zuo, T., Wei, H., Chen, N., & Zhang, C. (2020). First-and-last mile solution via bicycling to improving transit accessibility and advancing transportation equity. *Cities, 99*, Article 102614. https://doi.org/10.1016/j.cities.2020.102614