The Journal of Transport and Land Use jtlu.org
Vol. 19 No. 1 [2026] pp. 207-248
https://doi.org/10.5198/jtlu.2026.2859

How to promote the combined use of metro and shared bicycles under different land-use patterns around metro stations: A case study of Shanghai

Yantang Zhang (corresponding author), School of Civil Engineering, Heilongjiang University, 2025076@hlju.edu.cn
Xiaowei Hu, School of Transportation Science and Engineering, Harbin Institute of Technology, xiaowei_hu@hit.edu.cn

Abstract: The combined use of metro and dockless bike-sharing (M-DBS) is an emerging strategy for improving public-transport efficiency and tackling first-/last-mile gaps, and it is essential for sustainable urban mobility. However, existing studies often overlook land-use pattern heterogeneity and the nonlinear impacts of the built environment on M-DBS use. This study, using Shanghai as a case, employs K-Means clustering to identify four land-use patterns around metro stations—low-density, mixed-use, transit hub, and commercial core—and applies the XGBoost algorithm to model the nonlinear relationship between the built environment and M-DBS use across these patterns. We find that the built environment factors influencing M-DBS use vary by land-use pattern. In particularly, bus accessibility, population density, the proportion of shopping facilities, and subway accessibility play key roles in the four land-use patterns, respectively. Furthermore, the relationship between built environment variables and M-DBS use is nonlinear, with varying thresholds and even reversed correlations in different patterns. For example, in the transit hub pattern, M-DBS use negatively correlates with the proportion of dining facilities, while in the commercial core pattern, the correlation is positive. This study also reaffirms the differing impacts of the built environment on the two feeder modes of M-DBS use. Based on these findings, we provide strategic recommendations to optimize the integration of dockless bike-sharing and metro systems, improving public transportation efficiency and supporting sustainable urban development.

Keywords: Dockless bike-sharing, combined usage, built environment, land-use patterns, nonlinear relationship, machine learning

1 Introduction

Dockless Bike Sharing (DBS), as an innovative urban transportation mode, has gained widespread acceptance and application globally since 2016, particularly in countries like China (Guo & He, 2020). Within the framework of urban transportation systems, the proliferation of dockless bike-sharing systems has significantly reduced short-distance motor vehicle trips, for instance, trips previously made by car for first-mile and last-mile connections, markedly enhancing energy conservation and emission reduction effects in urban transportation (Cheng et al., 2022; Zhao & Li, 2017). More importantly, its unique dockless system endows shared bicycles with greater flexibility, enabling them to form a close integration with public transportation modes. The most common integrated usage pattern is the combination of metro travel and dockless bike sharing, typically referring to the travel behavior of transferring from the subway to dockless bike sharing or vice versa.

Received January 18, 2026; Revised July 13, 2026; Accepted July 13, 2026; Available online August 6, 2026

The *Journal of Transport and Land Use* is the official journal of the World Society for Transport and Land Use (WSTLUR) and is published and sponsored by the University of Minnesota Center for Transportation Studies.

In this study, "integrated travel mode" specifically refers to the physical integration of metro and dockless bike-sharing trips, i.e., using these two modes within a single trip chain to complete a journey that neither mode could serve independently, rather than payment integration or trip-planning integration. This integrated travel mode promotes seamless integration of public transportation modes, greatly improving public transportation efficiency and providing effective solutions to first-mile and last-mile problems (Zhao & Li, 2017). It has been recognized by numerous urban planners and policymakers as an important means of optimizing urban transportation structures (Böcker et al., 2020; Cheng et al., 2022; Guo et al., 2021; Zhao & Li, 2017;). Therefore, how to effectively enhance the combined use of M-DBS is of critical importance for improving urban public transport network efficiency, reducing transport-related carbon emissions, and advancing sustainable urban mobility goals.

The built environment around subway stations, encompassing socio-economic attributes, land-use patterns, and street design, is widely recognized as a critical determinant of Metro and Dockless Bike Sharing (M-DBS) combined use (Böcker et al., 2020; Cheng et al., 2022; Guo et al., 2021; Zhao & Li, 2017). Accordingly, various strategies have been adopted and implemented to attract more residents to choose the combined use of M-DBS by adjusting the attributes of the built environment around subway stations. However, planners' understanding is hindered by at least two research gaps.

First, existing studies have overlooked the heterogeneity of land-use patterns around subway stations (Chen, Wang, et al., 2020). Simply put, land-use patterns reflect the general function of land-use within a region, and their heterogeneity can be understood as the differences in the general function of land-use in different specific regions on a larger research scale. Previous studies have shown that land-use patterns vary significantly around different subway stations (Liu, Fan, et al., 2022), due to this, the impact of the built environment in different land-use modes is not consistent (Zhou et al., 2023). Tao and Cao (2024) found that the effects of the same built environment interventions vary significantly between suburban and Central Business District (CBD) areas. However, existing research exploring the relationship between the built environment and M-DBS combined use rarely considers the heterogeneous impact of land-use patterns around subway stations, which likely results in existing built environment interventions aimed at enhancing M-DBS combined use failing to achieve the expected outcomes due to a mismatch with the corresponding land-use patterns (Tao & Cao, 2024). Second, existing studies often rely on linear assumptions when exploring the impact of the built environment, failing to fully consider potential threshold effects. In particularly, within a specific range, once a built environment attribute reaches a certain threshold, further changes may have little impact (Tao & Cao, 2024). This neglect has resulted in estimation bias in the relationship between the built environment and the combination of M-DBS based on the generalized linear assumption, thus missing out on more refined land planning strategies.

Shanghai provides an ideal setting for investigating these relationships. As one of the world's largest metropolitan areas, Shanghai accommodates approximately 24.76 million residents and operates an extensive metro network of 20 lines, 409 stations, and 831 km, with daily ridership exceeding 10 million (Shanghai Municipal Transportation Commission, 2023). The city's dockless bike-sharing system has expanded in parallel, with over one million shared bicycles and more than 13 million registered users. According to the 2022 Shared Bicycle and Urban Development White Paper, over 85% of dockless shared bicycle trips in Shanghai involve a transfer to or from the metro, underscoring the centrality of M-DBS integration in the city's daily mobility. Moreover, the land-use context around Shanghai's metro stations is highly heterogeneous: decades of rapid urbanization have produced a rich mosaic of land-use patterns, ranging from dense CBDs and planned satellite towns to suburban communities and multi-modal transport hubs (Li et al., 2019; Liu, Zhou, et al., 2022). This diversity makes Shanghai an exceptionally suitable case for examining how built environment factors differentially influence M-DBS combined usage across varying land-use patterns, and for deriving policy insights applicable to other megacities.

To fill these gaps, this study tackles three research questions: 1) What are the land-use patterns around subway stations, and do the effects of built environment on M-DBS combined usage vary across these patterns? 2) Are there nonlinear relationships between built environment and M-DBS combined usage, and how do these relationships differ across land-use patterns? 3) What built environment interventions can more effectively promote M-DBS combined use under different land-use patterns?

To answer these questions, we applied machine learning algorithms, including K-means for clustering and XGBoost for regression, to shared bike travel data around subway stations in Shanghai. After identifying land-use patterns, we explored the relative importance and nonlinear relationships of built environment attributes on M-DBS inflow and outflow within each pattern and derived targeted intervention measures. This study provides theoretical and practical insights for optimizing land resource allocation around subway stations and enhancing overall transport system efficiency.

The contributions of this paper are as follows:

By identifying land-use patterns around subway stations and modeling M-DBS combined usage under each pattern, this study moves beyond single-pattern analytical frameworks, revealing the heterogeneous effects of built environment on M-DBS integration. This enables practitioners to design land-use policies tailored to specific local contexts, such as low-density suburbs versus commercial cores, avoiding one-size-fits-all interventions that may be ineffective or counterproductive.

By employing machine learning for nonlinear modeling, this study transcends the generalized linear assumptions in existing research, uncovering threshold effects and reversed correlations between built environment attributes and M-DBS combined usage across different land-use patterns. These nonlinear insights allow planners to allocate resources more efficiently: recognizing when additional investment yields diminishing or even negative returns prevents waste, while exploiting threshold effects enables substantial improvements in M-DBS usage through minor, low-cost built environment adjustments near critical thresholds, supporting more targeted and cost-effective interventions.

The rest of this article is organized as follows. Section 2 provides a literature review of related research on bike-sharing. Section 3 describes the scope of the study and data sources, presents descriptive statistics of the variables and introduces the methods. The results are presented in Section 4, and the discussion is provided in Section 5. The policy implications, conclusion, limitations, and future work of this study are included in Section 6 and 7.

2 Literature review

2.1 Built environment's impact on the combined usage of M-DBS in diverse land use patterns

In recent years, urban planners have been dedicated to enhancing the built environment to promote the combined utilization of M-DBS (Chen & Ye, 2021; Ji et al., 2017; Liu, Zhou, et al., 2022; Zhang, Hu, et al., 2024; Zhang & Hu, 2024; Zhou et al., 2023; Zhu & Diao, 2020). The development of Transit-Oriented Development (TOD) mixed-use models (Kumar et al., 2020), networked street design (Gao, Hu, et al., 2023), integration of public transportation systems (Zhou et al., 2023), and population densification (Cheng et al., 2022; Lin et al., 2018; Liu, Fan, et al., 2022) have all been implemented and validated, supporting the correlation between the built environment and the combined utilization of M-DBS. For instance, Zhou et al. (2023), using shared bicycle travel data near subway stations in Beijing, explored the impact of the built environment on the combined use of M-DBS. The study revealed that population density, the number of workplaces, and the quantity of shopping and dining establishments play crucial roles in the integrated system of M-DBS. Guo et al. (2021), using the 2017 trip data from a dockless bike-sharing operator in Shenzhen, proposed a people-subway-bicycle-route-urban space framework, which carefully examined the influence of the built environment on the combined utilization of M-DBS. The study indicated that the degree of land mix can enhance the combined utilization of M-DBS, and furthermore, the length of bicycle lanes correlates positively with combined utilization.

However, these interventions related to the built environment may not yield consistent effects across different land-use patterns, such as suburbs, commercial CBD, and transportation hubs (Kim et al., 2023; Kumar et al., 2020; Li et al., 2019). The following sub-sections review these differential effects by location type.

2.1.1 Suburban and low-density areas

In suburban areas, strategies promoting population densification may more easily enhance M-DBS usage, whereas their effectiveness in urban centers may not meet expectations. This is because in suburban areas, where population density has not reached its peak, and with DBS facilities exceeding demand, population densification can more readily stimulate DBS usage and thereby increase the combined utilization of M-DBS (Cheng et al., 2022). Hu et al. (2022) found that in suburban areas or university campuses in Shanghai, an increase in leisure facilities such as shopping may enhance the popularity of combined use of subway systems and shared bicycles. The relationship between buses and shared bicycles in suburban areas also differs from that in central areas. Especially in non-CBD areas, where transportation facilities are less densely distributed and there is a certain distance between subway and bus stations, bike-sharing trips, as an important mode of travel connecting subway and bus trips, not only do not compete with buses but actually promote cooperation between buses, subways, and shared bikes (Ma et al., 2019), thereby mitigating the substitution effect of bike-sharing on buses.

2.1.2 Urban core and CBD areas

In contrast, in saturated urban areas, population densification strategies may appear less effective and could even hinder M-DBS usage due to traffic congestion and safety concerns (Cheng et al., 2022). Cheng et al. (2022) employed a quantile regression approach to analyze the integration of free-floating bike-sharing and the urban rail transit system in Nanjing. The study found that policies encouraging increased residential density in the central business district (CBD) may have very limited effectiveness. The authors attributed this to the combined effect of the CBD's highly compact spatial form and localized cycling restrictions (the CBD has a high density of facilities, and cycling restriction strategies exist in some areas, such as the prohibition of non-motorized vehicles on certain road sections in the core area of Nanjing's Xinjiekou). Furthermore, considering that rail transit stations are extremely dense within the CBD, resulting in a significant overlap of walking catchment areas, walking is generally regarded as a substitute for cycling as the primary commuting mode to metro stations in such high-density environments. This, from another perspective, also weakens the effect of increased residential density on boosting bike-sharing feeder demand (Lin et al., 2018).

Zhao and Li (2017) argue that shopping centers around urban central subway stations in Beijing may not necessarily encourage more shared bicycle usage because of limited dense land supply in urban centers and planners' inclination towards designing pedestrian-friendly streets rather than promoting shared bicycle usage. Additionally, Guo and He (2020) found that an increase in the number of subway stations does not necessarily continue to drive the combined utilization of M-DBS in Shenzhen. Evidence from several first-tier cities in China suggests that a higher density of subway stations may reduce the potential for combined utilization because in densely populated areas, shared bicycles may be replaced by walking as a feeder mode to the subway (Lin et al., 2018).

2.1.3 Mixed-use and transit hub areas

Some variables about land-use facilities show context-dependent effects across different area types. Guo et al. (2020) found that an increase in the density of residential facilities catering to commuting needs significantly promotes the combined utilization of M-DBS during peak hours. Furthermore, empirical findings regarding the impact of public transportation facilities are mixed. Generally, more bus routes or bus stops near subway stations increase the likelihood of taking buses, which is not conducive to the combination of M-DBS (Luo et al., 2023; Martin & Shaheen, 2014; Zhao & Li, 2017). However, the relationship between buses and shared bicycles is not always competitive. This tension between substitution and complementarity of bicycles and public transportation has been recognized as a key theme in the literature (Kilani et al., 2024; Ma et al., 2019; Zhang & Hu, 2024; Zhao & Li, 2017). Ma et al. (2019) identify cooperative patterns between subway systems, buses, and shared bicycles in Chengdu, particularly in areas where transit stations are spaced farther apart.

The aforementioned studies highlight the importance of considering the heterogeneity of land-use patterns when exploring the complex relationship between the built environment and the combined use of M-DBS. However, existing research often covers entire areas, with the results typically reflecting the overall impact of the built environment on M-DBS combined use. Although some studies have explored the mechanisms of the built environment and bike-sharing travel behavior in specific areas, such as suburban and urban districts (Cheng et al., 2022; Hu et al., 2022; Kilani et al., 2024; Tao & Cao, 2024; Zhao & Li, 2017), these studies have insufficiently examined the land-use patterns around subway stations, which hinders the development of refined, cutting-edge intervention measures (Kumar et al., 2020; Li et al., 2019). Therefore, it is essential to systematically uncover the impact of built environment variables on M-DBS usage under different land-use patterns.

2.2 Nonlinear impact of the built environment on the combined use of M-DBS

The above content introduces the heterogeneity of land-use patterns, which is an important source of the inconsistent effects of the built environment on the combined use of M-DBS. Recently, the agglomeration effects and diminishing returns in urban economics have demonstrated the potential nonlinear effects of the built environment (Galster, 2018). The nonlinear and threshold effects of the built environment are considered another important source of these inconsistencies (Ding et al., 2019; Tao & Cao, 2024; Zhang, Chen, et al., 2024). This phenomenon refers to situations where variables of the built environment have significant effects only when they reach a certain level. However, previous studies have often relied on linear assumptions or generalized statistical models to establish the relationship between the built environment and the combined use of M-DBS (Gao, Hu, et al., 2023; Liu et al., 2024; Yan et al., 2024; Zhou et al., 2023). Examples include geographically weighted regression models (Zhou et al., 2023) and their derivatives, such as Multiscale Geographically Weighted Regression (MGWR) (Yan et al., 2024) and spatio-temporal geographically weighted regression (Liu et al., 2024), as well as models following specific probability distributions, such as Gaussian and Poisson distributions (e.g., poisson regression) (Gao, Yang, et al., 2023). Relationships derived from these models provide clues and inspiration for planning suggestions, such as general directions for intervention strategies, but they fail to reflect more complex real-world conditions. They also do not support the development of more precise intervention measures for planners. Recently, a series of studies using machine learning methods have revealed the nonlinear effects of the built environment (Ding et al., 2019; Sun et al., 2024; Zeng et al., 2024; Zhang, Hu, et al., 2024). For instance, Zhang, Hu, et al. (2024) using the XGBoost (eXtreme Gradient Boosting) algorithm and the SHAP interpretability method, found that in Shanghai's CBD, the growth trend of dockless bike-sharing usage significantly slows down when the road length reaches approximately 5.6 km. Additionally, compared to the least squares algorithm, the XGBoost algorithm demonstrates higher predictive accuracy. Yan et al. (2024) also confirmed this, finding that the XGBoost algorithm outperforms the MGWR algorithm on the test set. Gao, Yang, et al. (2023), using the Gradient Boosting Decision Tree (GBDT) algorithm, discovered that when bike lane density reaches 2 km/km², dockless bike-sharing usage increases significantly. Chen, Wang, et al. (2020), using the random forest algorithm, found that the number of bus stops and transit routes has a great impact on bike-sharing usage when the number of bus stops reaches 11 and the number of transit routes reaches 60.

Although the method of using machine learning models to reveal the nonlinear effects of the built environment on the DBS has become widely adopted, the combined use of M-DBS differs from either bike-sharing or subway usage alone, intervention measures suitable for enhancing dockless bike-sharing may not necessarily be effective in promoting the combined use of M-DBS, while existing literature rarely explores the nonlinear relationship between the built environment and the combined use of M-DBS. Therefore, it is necessary to investigate the nonlinear relationship between the built environment and the combined use of M-DBS to obtain more targeted and effective intervention measures.

In summary, we have reviewed two sources of inconsistent effects of the built environment on the combined utilization of M-DBS: land-use patterns and nonlinear relationships. We have emphasized the importance of considering land-use patterns and nonlinear effects in modelling the relationship between the built environment and M-DBS combined utilization. However, there has been no systematic exploration of the nonlinear relationships between built environment attributes and M-DBS combined utilization under different land-use patterns. Neglecting the influence of different land-use patterns around subway stations, interventions based on the nonlinear relationship between built environment attributes and M-DBS combined utilization may still exhibit bias. This results in the loss of more targeted and effective land planning strategies. Therefore, it is crucial to uncover the nonlinear relationships between the built environment and M-DBS combined utilization across different land-use patterns.

2.3 Land-use development patterns in Shanghai

As this study uses Shanghai as the case city, a brief review of its land-use development patterns is warranted to contextualize the empirical analysis. Shanghai's urban spatial structure has evolved through several distinct phases since the 1990s. The central city, encompassing the historic core and the Lujiazui financial district, is characterized by high-density commercial and mixed-use development, with land-use dominated by retail, office, and high-rise residential functions (Li et al., 2019). Beyond the central core, a ring of satellite towns (e.g., Jiading, Songjiang, Qingpu) was developed under the 1999–2020 Master Plan, with the objective of decentralizing population and employment from the congested city center. These satellite towns exhibit a more balanced mix of residential, industrial, and service land-uses, often centered around their own metro stations that connect to the core network (Liu, Zhou, et al., 2022). In the urban periphery, low-density residential communities, industrial parks, and agricultural land coexist, with metro stations serving as regional transit anchors that provide essential connectivity to the rest of the network. Figure 1 illustrates the land-use planning for Shanghai's main urban area, based on the city's Master Urban Plan (2017–2035) (Shanghai Municipal Bureau of Planning and Natural Resources, n.d.).

The land-use heterogeneity around Shanghai's metro stations is further shaped by the city's Transit-Oriented Development (TOD) strategy, which promotes differentiated development intensities based on station location and function (Li et al., 2019). Stations in the urban core tend to be surrounded by commercial complexes and high-density housing, while suburban stations are more likely to be adjacent to lower-density residential neighborhoods interspersed with service facilities. Transit hub stations, where multiple metro lines intersect, exhibit exceptionally high commercial agglomeration and pedestrian flow, creating a distinct land-use profile. This spatial differentiation of land-use around metro stations provides a natural setting to examine whether and how the influence of the built environment on M-DBS combined usage varies across land-use patterns and makes Shanghai an ideal case for deriving policy insights that may inform other megacities with comparable urban structures (Li et al., 2019; Liu, Zhou, et al., 2022). Topographically, Shanghai is situated on the Yangtze River Delta, and the entire metropolitan area features flat terrain with minimal natural slopes, meaning that cycling requires overcoming no significant climbing resistance and thus enjoys inherently favorable topographic conditions (Yin et al., 2013). Climatologically, Shanghai has a subtropical monsoon climate with four distinct seasons, and the influence of weather conditions on cycling behavior has become a key focus of recent empirical research. Li et al. (2021), using Mobike trip data from Shanghai, systematically analyzed the effects of multiple meteorological variables, including temperature, rainfall, humidity, and air quality, on the usage of free-floating bike-sharing. They found that high temperatures and rainfall significantly suppress cycling demand. Notably, however, even under adverse weather conditions, the demand for cycling trips connecting to metro stations exhibits strong rigidity. Compared with bus stops, cyclists in poor weather are more inclined to choose metro stations as their trip destinations, reflecting the buffering effect of rail feeder functions against weather impacts.

Figure 1. Land-use planning for Shanghai's main urban area
Source: The Master Plan of Shanghai 2017–2035 (Shanghai Municipal Bureau of Planning and Natural Resources, n.d.)

3 Methodology

The research framework is first introduced, as shown in Figure 2. Leveraging population data, Point of Interest (POI) data, OpenStreetMap (OSM) data, and riding data from Mobike, we obtain the dependent variables, the inflow and outflow of M-DBS combined usage. Additionally, we identify built environment attributes comprising population density, land use facility proportion, road network density, and accessibility indicators. Prior to incorporating the M-DBS combined utilization frequency and the built environment variables into the nonlinear model computation, a cluster analysis is performed on the characteristics of the built environment to discern the land-use patterns near subway stations. On this foundation, we delve deeper into analyzing the nonlinear relationship between built environmental factors and the combined utilization of M-DBS under varying land-use patterns and put forward the corresponding policy recommendations.

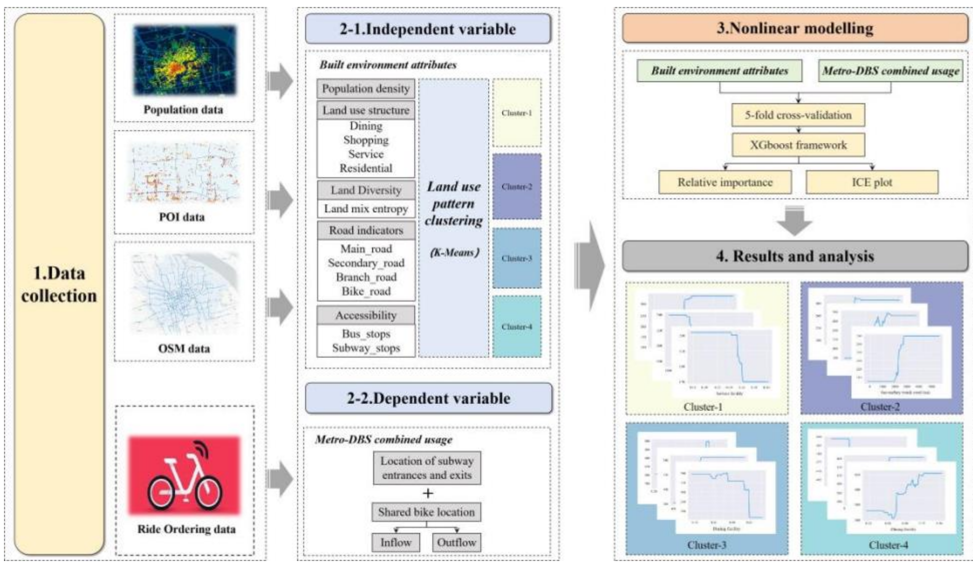

Figure 2. Research framework

3.1 Research scope

Shanghai is a globally renowned international city with a resident population of approximately 24.76 million and a total area of about 6,341 km^2. The city's urban rail transit has undergone rapid development, and by the end of 2023, Shanghai had opened 20 metro lines, comprising 409 stations, with a total operational distance of 831 km. The daily ridership of Shanghai's urban rail transit has reached 10.28 million, largely owing to the diverse development of land within the service scope of rail transit stations by the government and businesses. Such rich and diverse development models have attracted significant travel demand, driving the rapid expansion of Shanghai's rail transit system. The combined usage of metro and other modes of transportation, particularly the combination of metro and shared bicycles, has gained popularity. Currently, the number of shared bicycles in operation in Shanghai has surpassed one million, with over 13 million registered users. According to the 2022 shared bicycle and urban development white paper, over 85% of dockless shared bicycles in the urban area of Shanghai transfer to the subway (China Academy of Urban Planning & Design, 2022), so this study chooses Shanghai for case analysis, the research scope is illustrated in Figure 3.

Figure 3. Research scope

3.2 Data source

This study draws on four types of data: dockless shared bike order data, road network data, population data, and Point of Interest (POI) data. The following subsections describe each data source, its provenance, and the pre-processing steps applied.

Dockless shared bike order data. It is sourced from Mobike, one of the top three bike-sharing operators in China. We obtained DBS order data spanning from August 19th to August 23rd, 2019, and specifically selected data with no precipitation and temperatures between 20°C and 30°C to minimize the impact of weather on DBS order data. This temperature range is consistent with the range identified in the cycling literature as optimal for bike-sharing usage (Bean et al., 2021; El-Assi et al., 2017). Additionally, we selected only daytime order data to focus on commuting trips, for which M-DBS integration is the most prevalent and functionally significant. According to the 2021 Report on Shared Bike/E-Bike Riding in Major Chinese Cities released by the China Academy of Urban Planning & Design (CAUPD), the proportion of shared bike trips around subway stations in Shanghai reached as high as 39%, and existing travel surveys consistently identify M-DBS integration as a major commuting mode (Zhou et al., 2023; Guo et al., 2020). Moreover, daytime commuting trips are less sensitive to non-built-environment confounders such as recreational travel variability and night-time safety concerns, thus providing a cleaner setting for examining the effects of built environment variables (Sun et al., 2024; Zhang, Hu, et al., 2024).

It should be noted that by focusing on daytime data, this study primarily captures the effects of built environment variables on M-DBS usage during commuting periods. As reviewed in Section 2, the presence of shopping and dining facilities influences M-DBS usage; however, shopping- and dining-related trips are not confined to commuting hours. Consequently, the estimated effects of shopping and dining facility variables reported in this study reflect their influence specifically within the commuting time window, and their effects during non-commuting periods (e.g., evening leisure or weekend shopping trips) may differ. This limitation should be kept in mind when interpreting the policy implications related to these land-use variables.

After filtering, our raw data ultimately included a total of 4.76 million order records. The dataset provides the following information: bicycle ID, start time, end time, and the longitude and latitude of each bicycle at the beginning and end of the trip. The sample data is shown in Table 1. Using the Geopandas library in Python, we removed trips outside the study area by setting longitude and latitude boundaries. Following the recommendation of Shen et al. (2018) and Zhou et al. (2023), we excluded trips with durations less than 30 seconds or exceeding 1 hour, as well as those covering distances less than 100 m or more than 4 km, thereby eliminating all invalid data, such as duplicate records, missing crucial information, and location errors.

Notably, the DBS data used in this study were collected before the COVID-19 pandemic (Liu et al., 2026), since 2020, bike-sharing and public transit ridership patterns in Shanghai have undergone significant changes. During the pandemic, metro ridership dropped substantially while bike-sharing usage increased as travelers shifted away from crowded public transport; by 2021, metro ridership had recovered to approximately 86% of 2019 levels, while the share of non-motorized travel (including bike-sharing) in Shanghai's modal split rose from 25.5% (2019) to 28.6% (2021). Shi et al. (2023), analyzing bike-sharing data across pre-, during-, and post-pandemic periods, found that while absolute usage volumes fluctuated with pandemic waves, the core spatial structure of bike-sharing demand, including its relationship with land use and transit accessibility, remained largely stable. This suggests that the pandemic primarily affected the magnitude of travel demand rather than the fundamental relationships between the built environment and M-DBS usage. We therefore expect the directional findings and threshold patterns identified in this study to remain informative, while acknowledging that the absolute usage volumes reported here may not directly reflect the post-pandemic baseline. Future validation with post-pandemic data is warranted.

Table 1. Sample order data

Item	Sample
Bike ID	BA8C09291
Start time	2019/8/19 14:01
Start location	121.4134, 31.33624
End time	2019/8/19 14:08
End location	121.5139, 31.13796

Road network data is sourced from OpenStreetMap (OSM), a global, user-created, open-source geo-spatial data platform from which we crawled spatial and vector data for main roads, secondary trunk roads, branch roads, bike lanes, as well as public transportation facilities and metro stations in Shanghai (Zhang et al., 2015).

Population data is obtained from the 7th National Population Census of China. This dataset provides information on the resident population and population density in Shanghai. The dataset offers 2020 population data with a resolution of 100 m grid. Compared to other population grid datasets, this population data exhibits higher accuracy and resolution in China (Zhou et al., 2023).

Point of Interest (POI) data is sourced from Amap, which is referred to as AutoNavi Map. It is one of the most popular map navigation applications in China, providing a wealth of map information and an integrated API. We crawled the POI data for the urban area of Shanghai from this platform. The latitude and longitude of each POI are included in the data. The POI dataset consists of a total of 17 categories of data points as shown in Table 2. Notably, to avoid the over-fitting risk that accompanies high-dimensional features in a limited sample (Zhang, Hu, et al., 2024), we first ran a correlation screen and retained only the four POI categories that were both highly correlated with shared-bike usage and consistently emphasized in the residential facilities, shopping facilities, dining facilities, and service facilities. Their strong associations with cycling demand have been detailed in the literature review (Cheng et al., 2022; Guo et al., 2020; Zhou et al., 2023) and are therefore not repeated here.

Table 2. The POI categories

Code	POI Category
1	Car Service
2	Car Sales
3	Car Repair
4	Motorcycle Service
5	Dining
6	Shopping
7	Daily Life service
8	Sports
9	Hospital Related
10	Hotel Related
11	Tourist Attraction
12	Residence
13	Governmental Organization
14	Education Related
15	Transport Related
16	Finance Service
17	Enterprises

3.3 Selection and calculation of the variables

3.3.1 Calculation of M-DBS combined usage

The combined use of M-DBS includes the following two scenarios: one in which a user rides a shared bike to another destination after exiting the subway station, and the other in which a user travels to a subway station for the next segment of their journey by shared bike from another starting point (Guo et al., 2020; Zhao & Li, 2017). Since these two scenarios typically occur near subway stations, most existing studies quantify M-DBS combined usage by calculating the usage of dockless bikes around subway stations (Guo et al., 2020; Yan et al., 2024; Zhou et al., 2023).

Following Yan et al. (2024), we define a 100 m buffer zone around each subway station entrance and exit to identify M-DBS combined usage. While prior studies commonly use a 300 m buffer centered on the station's geometric center (Guo et al., 2020; Zhou et al., 2023), the 100 m entrance-level buffer is preferred because shared bicycles near subway stations are not uniformly distributed, they cluster near entrances and exits rather than around the station centroid (Guo et al., 2021). Yan et al. (2024) demonstrated that the 100 m entrance-level buffer provides a more accurate identification of M-DBS integration than the conventional station-centered 300 m buffer. We therefore calculate the sum of lock and unlock order frequencies within each subway station's entrance-and-exit buffer zones as the outflow and inflow of M-DBS combined usage, respectively (Zhou et al., 2023). Equations (1) and (2) give the daily-level formulas.

It is important to emphasize a key assumption when calculating the M-DBS combined usage flow in this paper: It is assumed that all shared bicycle trips that start or end near subway stations are related to the combined use of M-DBS. Although existing studies and surveys show that the use of dockless shared bicycles around subway stations can largely be considered as a combined use of subways and dockless shared bicycles (Yan et al., 2024; Zhou et al., 2023), there is still a small portion of users who start or stop using shared bicycles near subway stations without transferring to the subway. In particularly, trips transferring from shared bikes to buses are likely to confound the estimated effect of bus-accessibility variables on the M-DBS combined usage identified in this study. Because bike-order data contain neither metro entry/exit records nor individual continuous trajectories, this false-positive error cannot be fully eliminated (Yin et al., 2024).

$$Outflow_{m,t} = \frac{\sum_{e=1}^{d} E_{e,m,t}^{unlock}}{d} \tag{1}$$

$$Inflow_{m,t} = \frac{\sum_{e=1}^{d} E_{e,m,t}^{lock}}{d} \tag{2}$$

Where, $Outflow_{m,t}$ represents the M-DBS outflow usage of the *m* th subway station on day *t*, where $E_{e,m,t}^{unlock}$ is the number of unlock orders within the 100 m buffer zone of the *e* th entrance of the *m* th subway station on day *t*, and *d* is the total number of entrances of the *m* th subway station on day *t*. The M-DBS inflow usage of the *m* th subway station is similar, as shown in Equation (2), where $E_{e,m,t}^{lock}$ is the number of lock orders within the 100 m buffer zone of the *e* th entrance of the *m* th subway station on day *t*. Based on the above equations, we obtained the average daily combined use of M-DBS within the surveyed time range. The visualization of the identified calculation results and a simplified flowchart of the M-DBS calculation are shown in Figure 4.

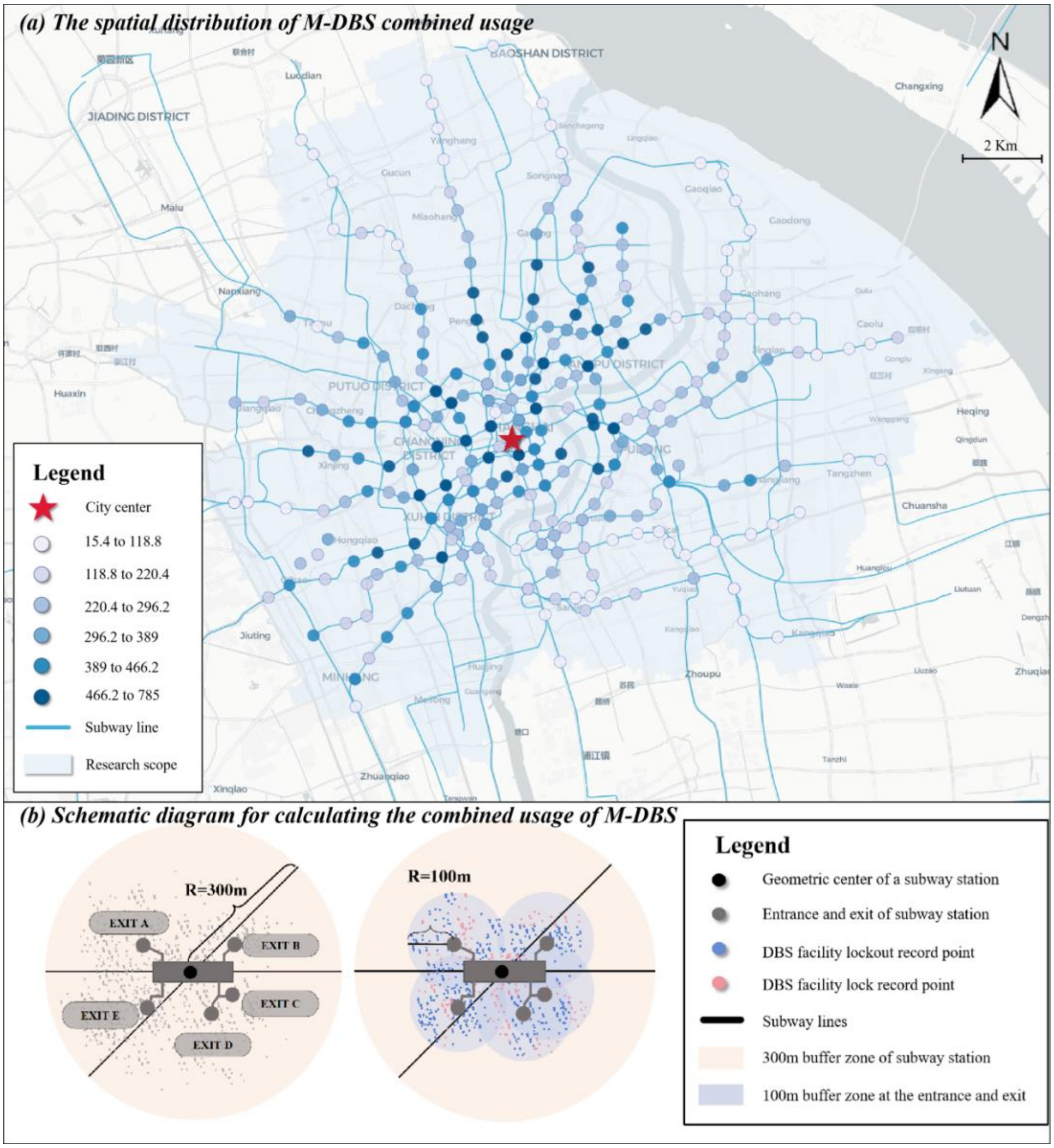

Figure 4. Calculation of the inflow and outflow of M-DBS combined usage

3.3.2 Calculation of variables affecting the M-DBS combined usage

Based on the established studies (Cheng et al., 2022; Guo et al., 2020; Zhou et al., 2023), we obtain the key variables affecting the M-DBS combined usage within the service area of the subway station.

The service area of metro stations has been widely discussed in previous studies, with the most commonly accepted definition being a static buffer based on walking accessible distance (Hu et al., 2022; Lin et al., 2019; Yan et al., 2024; Zhou et al., 2023). However, in studies on metro and dockless bike-sharing (M-DBS) integration, the involvement of bike-sharing significantly expands the actual catchment area of stations. Relevant research indicates that the cycling distance for most bike-sharing trips feeding into metro stations is approximately 1 km (Yang et al., 2018); therefore, 1 km is regarded as a more reasonable spatial unit for analyzing the mixed usage pattern of M-DBS (Hu et al., 2022;

Zhou et al., 2023). Accordingly, this study sets the static buffer radius of metro stations as 1 km.

To address the potential overlap of built environment characteristics within the buffer zones of multiple stations, Yan et al. (2024) and Zhou et al. (2023), in their studies covering the entire urban areas of Shenzhen and Beijing (both encompassing central urban and suburban areas), indicate that a significant overlap of built environment characteristics across multiple station buffers only occurs when the buffer radius exceeds 2000 meters. Therefore, the 1000-meter radius adopted in this paper is sufficient to effectively avoid such overlaps, ensuring the independence of the study units.

The built environment is typically characterized by demographic statistics, land use, street design, accessibility et al. (Chen, Wang, et al., 2020; Ding et al., 2018; Li et al., 2021). In this study, we compute the built environment variables within the service scope of rail transit stations and consider them as key variables influencing M-DBS combined usage. Table 3 presents the descriptive statistics of these variables. Specifically, concerning demographic statistics, we calculate the population density within the service scope of metro stations based on census results (Chen, Van Lierop, et al., 2020; Zhang et al., 2017). Regarding land use, we utilize Point of Interest (POI) data to calculate the proportion of various land-use facilities within the service scope of metro stations, such as service facilities, residential facilities, shopping facilities, and dining facilities. These land-use types have been considered closely related to DBS usage (Dong et al., 2022; Gao et al., 2021). In addition, we also calculated the land-use diversity using the land-use mixture entropy, the higher this value indicates a more diverse types of functions within the area. The calculation formula is shown in Equation (3), where p_q represents the proportion of the q th type of land-use facility, and n is the total number of land-use patterns. From the perspective of street design, we calculated the lengths of main roads, secondary trunk roads, branch roads, and bike lanes within the service scope of metro stations using road network data obtained from OSM. Since the service area has the same area, the influence of road length can be directly regarded as the influence of road density. Regarding accessibility, we computed the number of bus stops and metro stations within the service scope of metro stations to characterize both bus and metro accessibility (Hu et al., 2022; Ma et al., 2024).

It is noteworthy that the diverse cycling cultures and spatial configurations of cities lead to significant variations in the combined utilization of M-DBS and the attributes of the built environment across different urban settings (Tao & Cao, 2024). The present study conducts a case analysis focused on the urban area of Shanghai, where Table 3 solely presents the variable characteristics within the research scope of Shanghai's urban districts. This does not imply that the combined utilization of M-DBS and the attributes of the built environment in other cities exhibit similar characteristics.

$$Q_{\text{entropy}}=-(\sum_{q=1}^{n} p_q{*}\ln(p_q))/\ln(n) \tag{3}$$

Table 3. Descriptive statistics of variables

Variables	Description	Mean	Std
Dependent Variable			
Outflow of M-DBS Combined usage	The sum of the unlocking frequencies of DBS facilities within a 100 m buffer zone around each entrance and exit of the subway station.	302.37	214.41
Inflow of M-DBS Combined usage	The sum of the locking frequencies of DBS facilities within a 100 m buffer zone around each entrance and exit of the subway station.	307.02	169.10
Independent Variables			
Social-demographic attribute			
Population density	Population per square kilometer(person/km^2)	15589.91	13261.85
Land-use Facility			
Dining facility	Ratio of the number of POIs for Dining facilities to the number of all POIs, from 0 to 1.	0.26	0.13
Shopping facility	Ratio of the number of POIs for Shopping facilities to the number of all POIs, from 0 to 1.	0.33	0.09
Service facility	Ratio of the number of POIs for Service facilities to the number of all POIs, from 0 to 1.	0.27	0.07
Residential facility	Ratio of the number of POIs for Residential facilities to the number of all POIs, from 0 to 1.	0.20	0.11
Land Diversity			
Mix_index	Land Mixing Entropy Index, the higher the entropy value, the higher the land diversity, and the lower the entropy value, from 0 to 1.	0.76	0.23
Street design			
Main_road	The length of the main road within a 1 km radius centered on the subway station (m)	7378.43	2713.48
Secondary_road	The length of the secondary trunk road within a 1 km radius centered on the subway station (m)	2791.44	726.97
Branch_road	The length of the branch road within a 1 km radius centered on the subway station (m)	4337.84	1637.34
Bike_road	The length of the bike road within a 1 km radius centered on the subway station (m)	1896.36	2544.18
Accessibility			
Bus stops	Number of bus stops within a 1 km radius centered around the subway station	10.98	4.50
Subway stops	Number of subway stations within a 1 km radius centered around the subway station	1.67	1.24

Before incorporating these variables into the calculation, we need to perform a multicollinearity analysis of the variables (Liu, Zhou, et al., 2022). When there is collinearity between independent variables, the parameters of the model become extremely unstable and the predictive ability of the model decreases. We used correlation analysis to test the multicollinearity problem between variables. Generally, two or more Pearson correlation coefficients greater than 0.7 are considered to have multicollinearity (Hong et al., 2020; Zhang, Hu, et al., 2024). As shown in Figure 5, all values are less than 0.7, and we did not find multicollinearity problems in the selected variables (Liu, Zhou, et al., 2022). Therefore, all variables can be used for subsequent clustering analysis.

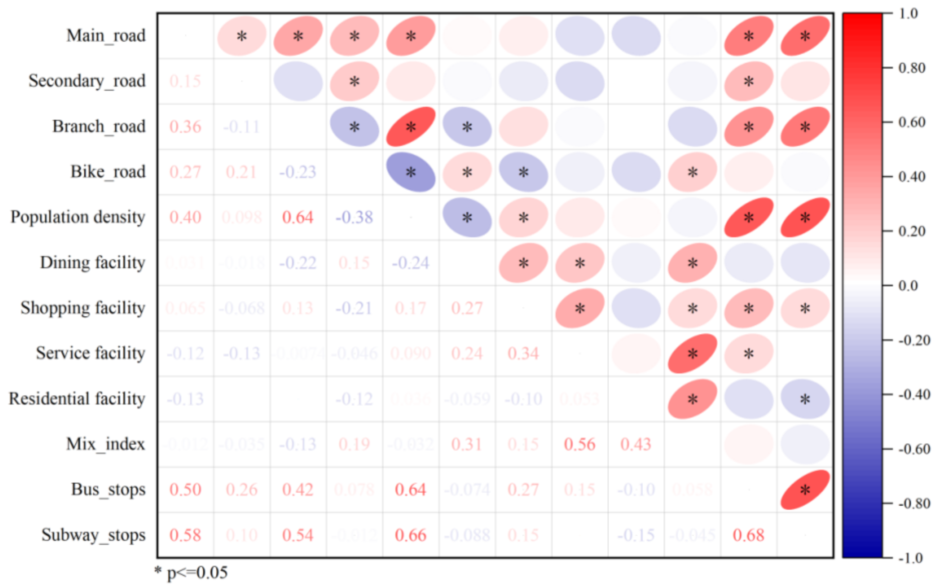

Figure 5. Correlation analysis results of built environment variables

3.4 Methods

3.4.1 K-means clustering algorithm

First, we used unsupervised clustering methods to identify land-use patterns within the service area of metro stations based on built environment variables. The rationale for selecting the K-means algorithm is detailed in Appendix A.

We selected the optimal number of clusters based on two criteria:

1. Using SSE (Sum of Squared Errors) to identify the optimal K (Wang et al., 2019), commonly referred to as the elbow method. SSE measures the sum of squared distances between each observation and its assigned cluster centroid. The optimal K corresponds to the point beyond which additional clusters yield diminishing reductions in SSE.

2. Using the silhouette coefficient to evaluate clustering quality (Chang et al., 2020). This metric combines cluster cohesion and separation, ranging from −1 to 1; higher values indicate that observations are better matched to their assigned cluster than to adjacent clusters. Both criteria converge on K = 4 (Figure 6), suggesting well-separated clusters for which hard and soft clustering methods typically yield convergent assignments (Steinley & Brusco, 2011). We acknowledge that stations located at spatial boundaries between patterns may possess transitional characteristics that discrete classification cannot fully represent; future research with expanded samples could apply fuzzy clustering approaches to model gradual spatial transitions (Bezdek et al., 1984).

3.4.2 XGBoost algorithm

After identifying the land-use patterns, we analyze the key built environment variables affecting M-DBS combined usage within each pattern and model their nonlinear relationships. The selection of an appropriate modeling method is constrained by four requirements imposed by our research design: (a) the model must capture nonlinear and threshold effects without pre-specifying functional forms, (b) it must produce interpretable feature importance measures that are comparable across land-use patterns,

(c) it must accommodate small and unequal sample sizes across clusters (ranging from n = 30 to n = 132), and (d) it must handle inherent multicollinearity among built environment variables. We evaluate candidate methods against these criteria in turn. In Appendix B, we provide a detailed introduction and discussion of the characteristics of various modeling methods as well as the rationale for selecting the XGBoost algorithm.

In summary, XGBoost is selected because it satisfies all four design constraints simultaneously: it captures nonlinearity and interactions without parametric pre-specification (Chen & Guestrin, 2016; Friedman, 2001), provides gain-based feature importance for cross-pattern comparison (Lundberg & Lee, 2017), incorporates regularization critical for small-sample estimation (Hastie et al., 2009), and demonstrates superior predictive accuracy in our empirical comparison (Table 4). These properties make XGBoost uniquely suited to the analytical task of identifying pattern-specific, threshold-sensitive relationships between the built environment and M-DBS usage.

The processed dataset includes both independent variables x_i and dependent variables y_i, Here, x_i represents the independent variable in the i-th sample, and y_i represents the dependent variable in the i-th sample. With 12 characteristics per sample, including 4 indicators related to the proportion of land-use facilities (service facilities, residential facilities, shopping facilities, dining facilities), 4 indicators related to road length (main roads, secondary trunk roads, branch roads, bicycle lanes), 2 accessibility indicators (number of subway stations, number of bus stations), as well as population density and land mix entropy. The tree ensemble model uses K additive functions to estimate the target value, as shown in Equation (4).

$$\hat{y}=\varphi(x_i)=\sum_{k=1}^{K} f_k(x_i), f_k\in F \tag{4}$$

Where f_k is the independent tree structure and F is the tree space.

$$\mathrm{L}(\varphi)=\sum_{i} l(y_i,\hat{y}_i) + \sum_{k} \Omega(f_k) \tag{5}$$

The objective function that the XGBoost algorithm aims to minimize, as shown in Equation(5), is expressed as the sum of the loss function $\sum_i l(y_i,\hat{y}_i)$ and the regularization term $\sum_k \Omega(f_k)$. $\sum_i l(y_i,\hat{y}_i)$ represents the sum of the loss values for i samples, $\sum_k \Omega(f_k)$ represents the sum of the complexity for k trees. For the complexity Ω, we give a more detailed formula as shown in Equation (6), where T is the number of leaves and w_j is the weight of leaf j. w_j^* is the optimal value of w_j as shown in Equation (7), $\partial_{\hat{y}^{t-1}} l(y_i,\hat{y}^{t-1})$ and $\partial^2_{\hat{y}^{t-1}} l(y_i,\hat{y}^{t-1})+\lambda$ are used to denote the first-order and second-order gradient statistics of the i th sample at the t-1 th iteration in terms of the y_i and the $\hat{y}^{t-1}$. I_j represents the index set of all training samples x_i assigned to the j-th leaf node by the tree model. Notably, in XGBoost's regression problem, specifically the algorithmic challenge addressed in this study, the second-order gradient is consistently defined as the second-order partial derivative of the loss function with respect to the current model's output. For regression tasks, this naturally corresponds to the full Hessian matrix, eliminating any cross-term components. Given spatial constraints, readers can refer to Chen and Guestrin (2016) to learn more about XGBoost's technical details.

$$\Omega(f_k)=\gamma T+\frac{1}{2}\lambda w_j^2 \tag{6}$$

$$w_j^*=-\frac{\sum_{i\in I_j}\partial_{\hat{y}^{t-1}}l(y_i,\hat{y}^{t-1})}{\sum_{i\in I_j}\partial^2_{\hat{y}^{t-1}}l(y_i,\hat{y}^{t-1})+\lambda} \tag{7}$$

4 Results

4.1 Land-use patterns around urban rail transit stations

As shown in Figure 6, we use the elbow plot and silhouette coefficient plot to find the optimal K value for the K-means clustering algorithm. The results show that when the number of clusters is four, the downward trend in the elbow plot is significantly reduced, and the silhouette coefficient is maximized, indicating that four clusters are the most appropriate.

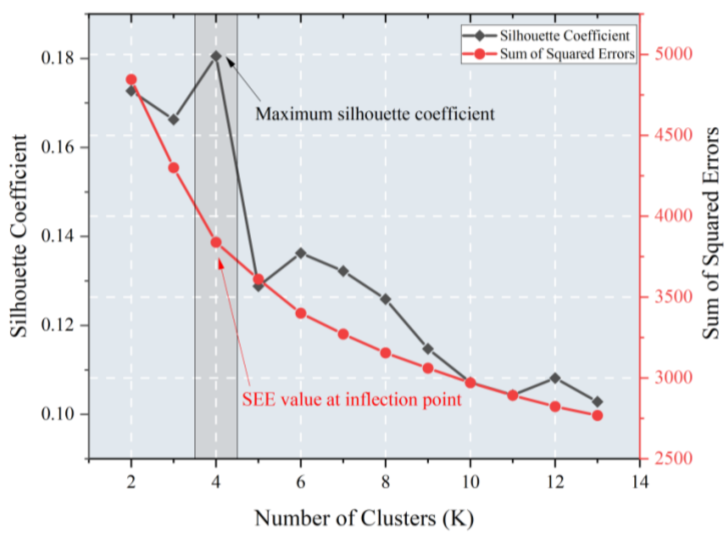

Figure 6. Elbow plot and silhouette coefficient plot

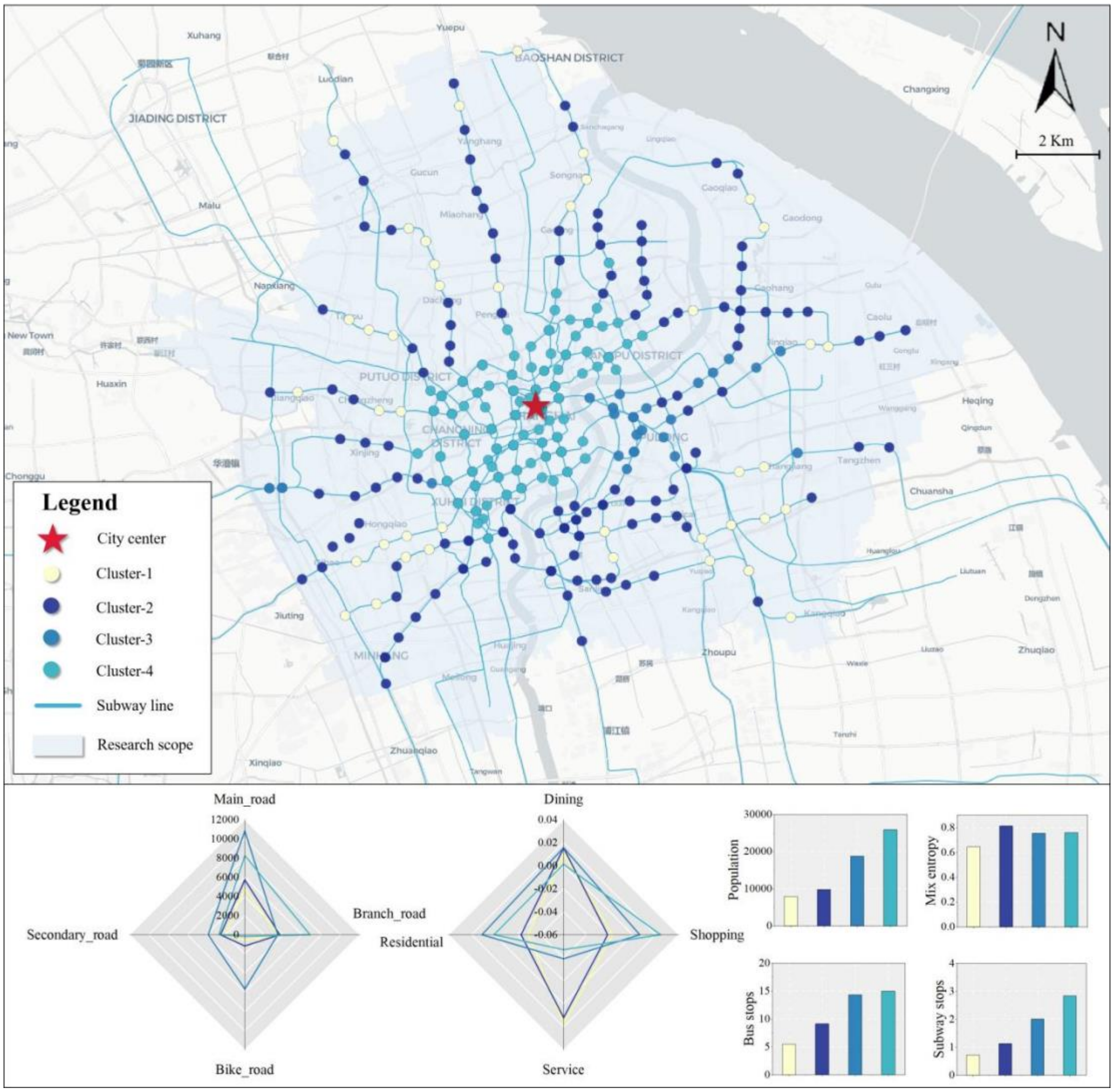

Figure 7. Four land-use patterns around urban rail transit stations and distribution

Figure 7 depicts the land-use patterns surrounding urban subway stations classified into four types. We visually represent the spatial distribution of these patterns and describe their characteristics based on built environment attributes, including road indicators, land use, land diversity, public transit accessibility, and population density et al. To better observe the differences in land-use facilities among these patterns, we calculate and visualize the differences between the indicator values and the mean values of all patterns. For example, for the proportion of shopping facilities, Cluster 4 is approximately 2 percentage points higher than Cluster 3. We draw insights from the community types oriented towards potential public transit development identified by Kumar et al. (2020), as well as the Transit-Oriented Development (TOD) patterns around urban rail transit stations in Shanghai, China, as delineated by Li et al. (2019), to describe the four land-use patterns obtained in this study. These four clustering patterns are: low density pattern, mixed use pattern, transit hub pattern, and commercial core pattern. It is worth noting that, limited by the approach of case analysis and considering the significant differences in urban structure and spatial layout, the four types of land-use patterns obtained in this article cannot cover all potential land-use patterns (Tao & Cao, 2024).

Furthermore, we have also conducted a statistical analysis of the combined usage of M-DBS corresponding to the land-use patterns around subway stations, as illustrated in Figure 8. Specifically, Figure 8 utilizes grouped box plots to illustrate the distribution of inflow and outflow volumes associated with the combined use of four land-use patterns and presents the mean values accordingly.

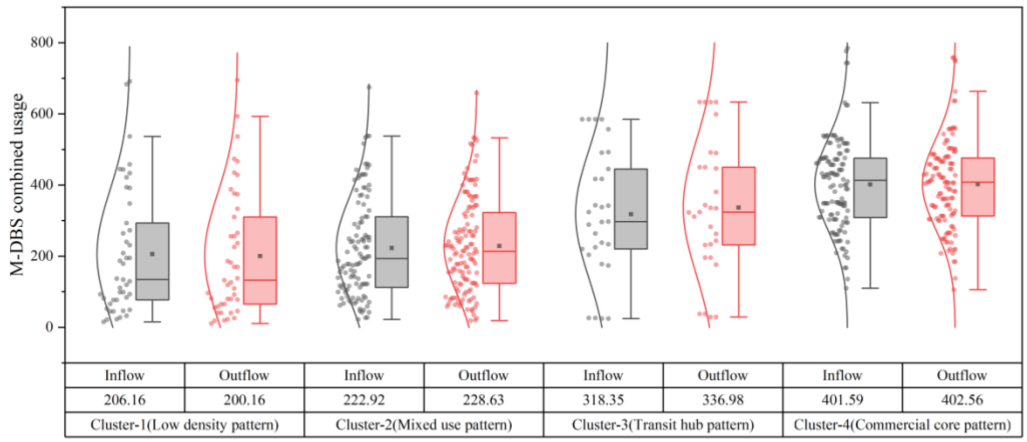

Figure 8. The combined usage of M-DBS corresponding to four land-use patterns

4.1.1 Low-density pattern (Cluster-1)

This pattern is primarily situated on the urban outskirts and encompasses 45 stations, the second smallest group. All road length indicators are the lowest among the four patterns, consistent with the limited transport infrastructure investment characteristic of peripheral urban areas. The land structure is dominated by dining (27.8%) and service (29.6%) facilities, which together account for nearly 60% of all POIs, reflecting a land-use profile oriented toward meeting the daily needs of local residents rather than attracting external visitors. Population density is the lowest among all patterns, at approximately 7,911 persons/km², and public transit accessibility is correspondingly limited, with relatively few bus and subway stations within the service area. These attributes collectively define a low-intensity, locally oriented urban fringe environment where travel demand is constrained by sparse population and limited destination diversity. As expected, M-DBS usage is the lowest among the four patterns (inflow: 206/day; outflow: 200/day). The near-parity between inflow and outflow suggests balanced bidirectional commuting patterns, though at low absolute volumes.

4.1.2 Mixed-use pattern (Cluster-2)

This pattern is the largest group, comprising 132 stations. Spatially, it is situated on the urban periphery but extends further toward the city center than the low-density pattern, forming a transitional belt between the urban fringe and the core. The land structure shares a similar emphasis on dining (28.3%) and service (29.1%) facilities with the low-density pattern, yet the defining feature of this pattern is its exceptionally high land-use diversity: the land mix entropy index is the highest among all four patterns. This elevated diversity likely reflects Shanghai's satellite city planning initiatives, which promote a balanced mix of residential, commercial, and service functions on the urban periphery to decentralize population and employment from the congested city center (Li et al., 2019; Liu, Zhou, et al., 2022). Compared to the low-density pattern, all road lengths are slightly longer, population density is moderately higher at approximately 9,757 persons/km², and public transit accessibility is more developed, with a greater number of bus and subway stations. M-DBS usage is moderately higher than in the low-density pattern (inflow: 223/day; outflow: 228/day). The slight asymmetry favoring outflow over inflow may indicate that these mixed-use peripheral areas function more as residential origins than as employment destinations during commuting hours, a pattern consistent with their satellite-city role of housing workers who commute inward.

4.1.3 Transit hub pattern (Cluster-3)

This pattern comprises only 30 stations, the smallest group among the four patterns, and is spatially concentrated at major transfer nodes in the city's core area. Its defining characteristic is the primacy of transportation infrastructure. The road network profile is distinctive: the lengths of main roads, secondary trunk roads, and notably, bicycle lanes all exceed those of the commercial core pattern, while the branch road length falls below it. This configuration reflects the hub's functional logic: the dense grid of arterial and collector roads facilitates high-volume passenger throughput and regional accessibility, and the greater bicycle lane provision supports shared-bike connections to surrounding areas; the relatively sparse branch road network, by contrast, indicates limited fine-grained, pedestrian-scale urban fabric, consistent with a node designed for transfer rather than for dwelling or street-level retail. In essence, the transit hub's road network is optimized for channeling flows in and out, whereas the commercial core's network, with its higher branch road density, supports the capillary circulation characteristic of dense urban neighborhoods. While the land structure shares a general emphasis on shopping (34.1%) and residential (13.7%) facilities with the commercial core pattern, the underlying functional logic differs fundamentally. In transit hub areas, commercial activity is derivative of passenger flow, retail and service establishments cluster around transfer stations to serve the transient population of commuters changing between lines. Population density in this pattern is substantially lower than in the commercial core, consistent with the hub's primary orientation toward through-movement rather than residential or employment agglomeration. Public bus accessibility is high relative to peripheral areas, reflecting the integration of bus routes with major transfer stations, but falls below that of the commercial core, consistent with the hub's role as a regional transit interchange rather than a dense employment center. M-DBS usage is moderately high (inflow: 318/day; outflow: 336/day), matching the pattern's function as a transfer node where shared bicycles serve as an intermediate feeder mode between bus, metro, and surrounding destinations.

4.1.4 Commercial core pattern (Cluster-4)

This pattern encompasses 128 stations and represents the densest urban form in Shanghai. In contrast to the transit hub pattern, the commercial core is defined not by transportation infrastructure but by high-density commercial and residential agglomeration. Population density reaches approximately 25,861 persons/km², substantially higher than in the transit hub, and public transit accessibility is the highest among all four patterns, with up to 15 bus stops and nearly 3 subway stations within the 1 km buffer zone of an average station. Critically, although the aggregate proportions of shopping (35.8%) and residential (14.7%) facilities appear similar to those of the transit hub pattern, the composition and spatial logic differ. In the commercial core, these facilities form a dense, mixed-use urban fabric where commercial establishments primarily serve the resident and worker populations rather than transient commuters; the high residential density generates substantial origin-destination demand, while the concentration of employment attracts inbound trips throughout the day. The road network, while maintaining relatively high levels across all indicators except bicycle lane length, is less developed than that of the transit hub in terms of main roads, secondary trunk roads, and bicycle lanes, a consequence of historical urban development where arterial expansion is constrained by existing high-density building stock. Its higher branch road density, however, supports the fine-grained pedestrian and cyclist circulation typical of dense urban neighborhoods. This distinction is substantive: the transit hub's road network is oriented toward channeling external flows, whereas the commercial core's road network, with more extensive branch roads but fewer high-capacity arterials, reflects its role as a destination rather than a conduit. M-DBS usage is the highest among the four patterns (inflow: 401/day; outflow: 402/day), driven jointly by high population density, employment concentration, and extensive transit accessibility.

4.2 Impact of built environment on M-DBS combined usage under various land-use patterns

We use the XGBoost algorithm to reveal the impact of built environment factors on the combined usage of M-DBS under different land-use models. We divide the dataset into a 60% training set, a 20% validation set, and a 20% test set, and test different combinations of hyperparameters using a 5-fold cross-validation method and grid search to mitigate overfitting and obtain the best parameter settings. We report the key hyper-parameters of the XGBoost, GBDT, and random forest model (maximum depth, learning rate, etc.) and evaluate its performance using the R^2 on the test set. As shown in Table 4, we find that the XGBoost model performs better than others. It is worth noting that the essence of machine learning regression-deriving complex relationships between variables from large datasets-leads to stronger interpretability for specific datasets, but at the same time, it somewhat limits the transferability of model parameters and regression results to other contexts. This is an inevitable drawback of data modeling (Tao & Cao, 2024).

Table 4. Statistics of model fitting indicators

Model	Cluster-1	Cluster-2	Cluster-3	Cluster-4
R^2_XGBoost	0.83	0.81	0.86	0.84
R^2_GBDT	0.63	0.72	0.65	0.73
R^2_RF	0.60	0.63	0.66	0.71
XGBoost_Max_Depth	19	16	6	8
XGBoost_Learning_Rate	0.20	0.25	0.15	0.05

4.2.1 Relative importance

In the realm of machine learning, the relative importance of a feature serves as a pivotal metric for interpreting model predictions, revealing the predictive prowess of a given variable. This measure quantifies the significance and contribution of a feature attribute in the construction of decision trees within an ensemble model. In this study, a high relative importance ascribed to a particular feature indicates its substantial contribution to the predictive power of the combined M-DBS usage. Additionally, given that previous studies have demonstrated varying effects of the built environment on the two feeder modes of dockless bike sharing systems in proximity to subway stations, this study delves into the relative importance of the built environment in shaping the combined utilization of M-DBS. Specifically, we analyze this from two distinct angles:

the inflow volume, which signifies the transition from DBS to subway, and the outflow volume, representing the reverse transition from subway to DBS.

Tables 5 and 6 present the relative importance of key built environmental factors that influence the inflow and outflow of combined usage of M-DBS under different land-use patterns around urban rail transit stations. Due to the constraints of space, our focus primarily lies on the top three variables ranked by their relative significance.

Table 5. The relative importance of built environment attributes on inflow

	Relative importance			
Variable	Cluster-1	Cluster-2	Cluster-3	Cluster-4
Socio-demographic attribute				
Population density	8.57%	32.41%	13.38%	4.88%
Street design				
Main_road	5.49%	9.79%	2.59%	8.53%
Secondary_road	8.00%	13.36%	5.84%	15.15%
Branch_road	6.90%	3.92%	5.38%	5.47%
Bike_road	3.10%	5.63%	2.01%	0.97%
Land use				
Shopping facility	7.38%	0.00%	25.49%	11.64%
Service facility	11.95%	11.45%	10.71%	4.36%
Residential facility	8.50%	8.39%	3.03%	8.66%
Dining facility	4.27%	0.00%	4.05%	10.84%
Land Diversity				
Mix_index	0.00%	0.00%	8.33%	3.83%
Accessibility				
Bus_stops	35.83%	8.57%	16.63%	3.59%
Subway_stops	0.00%	6.48%	2.57%	22.06%

In the low-density mode, both the number of bus stops and the proportion of life service facilities exhibit high relative importance in influencing the inflow and outflow of M-DBS combined usage. The relative importance of the former in affecting inflow and outflow is 35.83% and 33.22%, respectively, while the latter's relative importance is 11.95% and 12.17%, respectively. Road network density, especially the density of branch roads, has a more pronounced impact on outflow, with a relative importance of 12.43%. Meanwhile, for inflow, population density takes precedence, with a relative importance of 8.57%.

In the mixed-use mode, population density emerges as the most significant factor influencing the combined usage of M-DBS, accounting for 32.41% and 28.31% in relative importance for the inflow and outflow, respectively. Notably, under this mode, road network indicators and the proportion of land-use facilities exert a greater impact on the inflow. Specifically, the relative importance of secondary trunk roads and service facilities is pronounced, accounting for 13.36% and 11.45%, respectively. However, for the outflow of M-DBS combined usage, the influence of secondary trunk roads and the number of bus stops is more apparent, with relative importance scores of 13.25% and 16.90%, respectively.

In the transportation hub mode, population density remains a crucial variable influencing the usage of M-DBS, accounting for 13.38% and 18.74% of the relative importance for inflow and outflow, respectively. The proportion of shopping facilities also exhibits a strong influence on the inflow and outflow of M-DBS combined usage, with relative importance scores of 25.49% and 20.48%, respectively. However, within this mode, dining facilities only significantly impact the outflow of M-DBS combined usage, contributing a relative importance of 17.64%, while their impact on inflow is minimal, accounting for just 4.05%.

In the commercial center mode, population density plays a less critical role, giving way to road network indicators, land-use facilities, and accessibility. Among these, the density of secondary trunk roads and the number of subway stops emerge as decisive factors influencing the inflow and outflow of M-DBS combined usage. Specifically, for the inflow, these two indicators contribute a relative importance of 15.15% and 22.06% in this mode. For the outflow, their relative importance stands at 12.81% and 21.06%, respectively.

Table 6. The relative importance of built environment attributes on outflow

	Relative importance			
Variable	Cluster-1	Cluster-2	Cluster-3	Cluster-4
Socio-demographic attribute				
Population density	11.61%	28.31%	18.74%	2.96%
Street design				
Main_road	4.44%	10.98%	4.87%	10.15%
Secondary_road	7.96%	13.25%	1.27%	12.81%
Branch_road	12.43%	0.93%	3.79%	6.78%
Bike_road	0.73%	2.23%	9.09%	3.65%
Land use				
Shopping facility	6.82%	4.69%	20.48%	11.50%
Service facility	12.17%	6.40%	0.93%	2.79%
Residential facility	1.25%	0.93%	3.20%	5.40%
Dining facility	0.73%	1.25%	17.64%	12.16%
Land Diversity				
Mix_index	7.71%	13.22%	6.08%	5.03%
Accessibility				
Bus_stops	33.22%	16.90%	3.80%	5.70%
Subway_stops	0.93%	0.90%	10.11%	21.06%

4.2.2 Nonlinear relationship

Utilizing the results of the XGBoost model, we have employed Individual Conditional Expectation (ICE) plots to elucidate the nonlinear relationships between built environment attributes and M-DBS utilization across different land-use patterns. Due to space constraints, this section highlights the nonlinear associations of the top three variables that significantly impact M-DBS usage. Figures 9 to 12 illustrate the nonlinear relationships between the built environment of low-density, mixed-use, transportation hub, and commercial core modes, respectively, and the inflow and outflow of M-DBS. To assess the availability of these nonlinear relationships, we have incorporated a rug plot on the axis, providing insights into the distribution of data samples. Sparse whisker regions, for instance, may indicate insufficient data, rendering the nonlinear relationship less meaningful in those segments. Notably, we refrain from utilizing scatter plots from the SHAP library to visualize these nonlinear relationships, as they are inadequate in capturing the intuitive trend between built environment variables and M-DBS utilization.

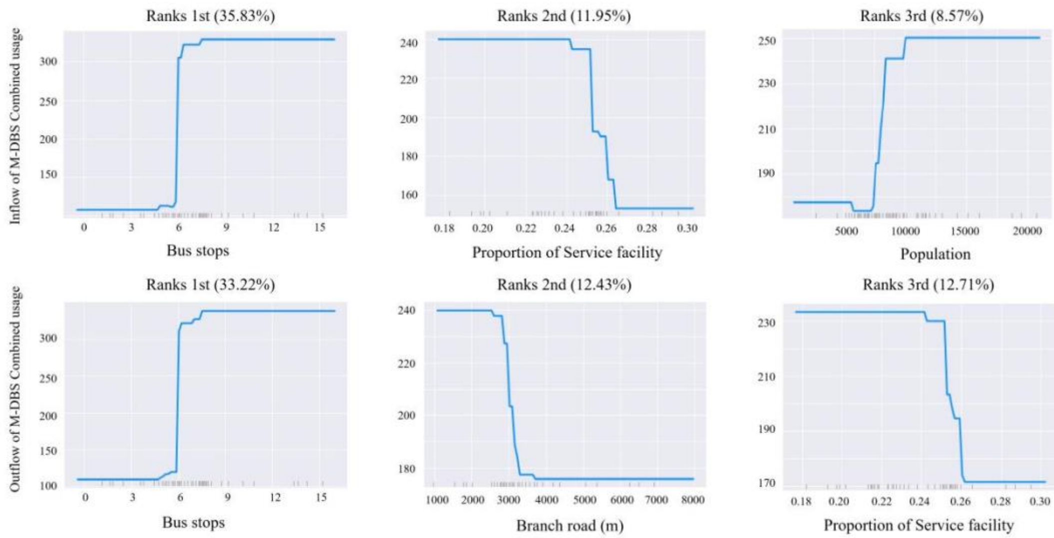

Figure 9. Nonlinear relationship under low-density pattern

In the low-density pattern, we observe a positive correlation between the number of bus stops and both the inflow and outflow of M-DBS combined usage. A threshold effect becomes apparent when the number of bus stops reaches approximately six. Within this mode, there is a negative correlation between life service facilities and the inflow and outflow of M-DBS usage. Specifically, when the proportion of service facilities hits 25%, there is a notable decrease in the inflow and outflow of M-DBS usage. Additionally, population density exhibits a significant positive correlation with the inflow of M-DBS usage in the low-density mode. When the population density reaches approximately 7500 person/km^2, there is a substantial increase in the inflow of M-DBS usage in the low-density mode. Furthermore, there is a pronounced negative correlation between the length of branch roads and the outflow of M-DBS usage in the low-density mode. Once the branch road length attains 3 km, there is a sharp decline in M-DBS outflow.

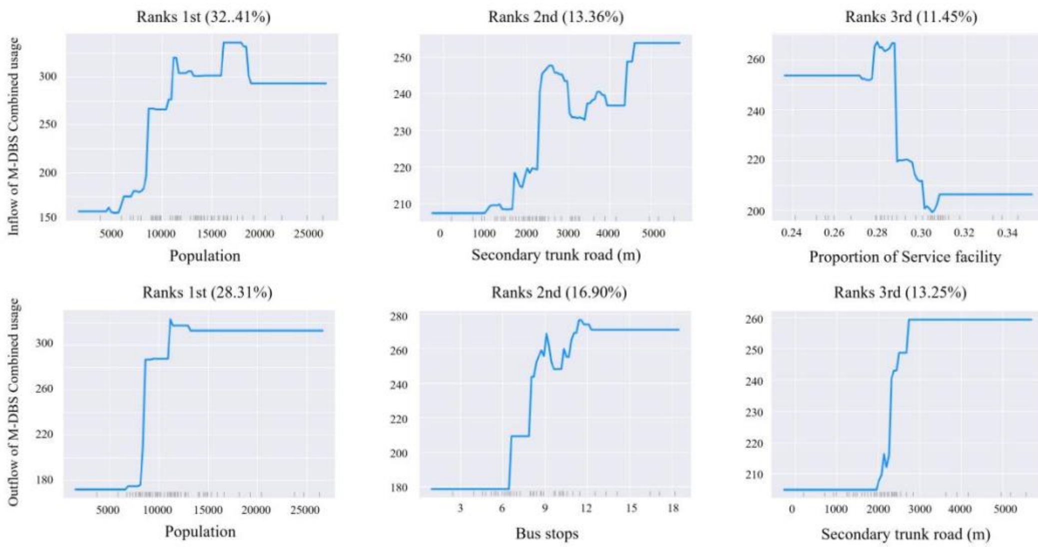

Figure 10. Nonlinear relationship under mixed-use pattern

In the mixed-use pattern, we observe a positive correlation between population density as well as the length of secondary trunk roads and the combined usage of M-DBS. The threshold for population density exerting the role is approximately 8000 person/km^2. However, the threshold for the secondary trunk road lengths exerting the role differs for the inflow and outflow of M-DBS combined usage. Specifically, there are two significant increases for the inflow of M-DBS combined usage when the length of secondary trunk roads reaches 2200 m and 4200 m, respectively, whereas the outflow only experiences a notable increase when the length reaches 2200 m. Additionally, there is a clear negative correlation between the proportion of life service facilities and the inflow of M-DBS combined usage. When the proportion of life service facilities is around 29%, the inflow experiences a significant decline. It is worth mentioning that in the low-density mode, the proportion of life service facilities also exhibits a negative correlation with the inflow, but the threshold shifts from 25% to 29%, indicating the heterogeneous impact of different land-use patterns. This heterogeneity is also evident in the impact of bus stop numbers. In the mixed-use mode, the number of bus stops positively correlates with the outflow of M-DBS combined usage, with a significant increase in outflow when the number reaches 6, plateauing at 12 bus stops. However, in the low-density mode, the outflow plateaus at around 7 bus stops.

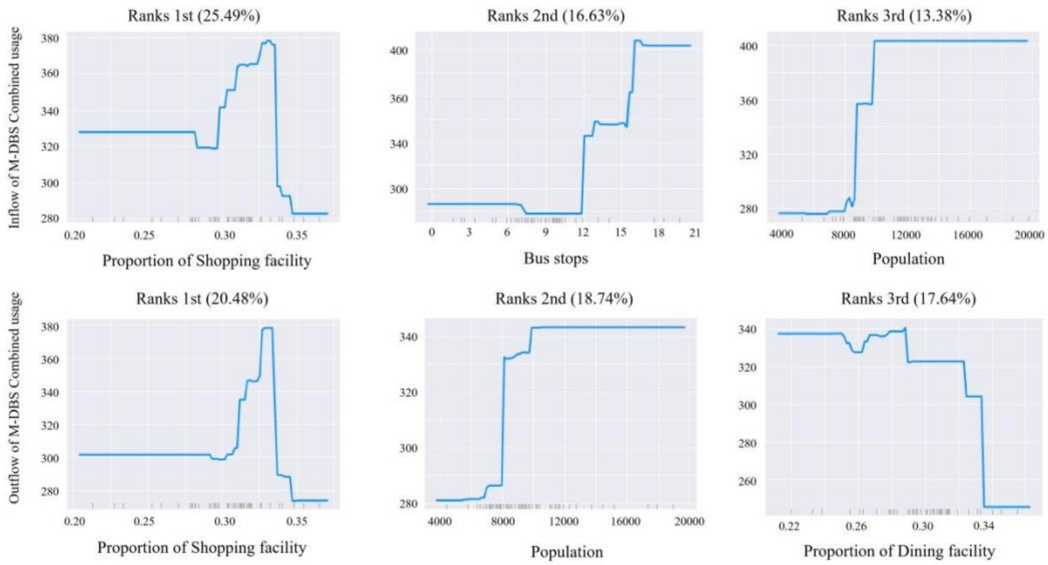

Figure 11. Nonlinear relationship under transportation hub pattern

In the transportation hub pattern, the proportion of shopping facilities and population density are pivotal variables that concurrently influence the M-DBS combined usage. Specifically, the nonlinear relationship between the share of shopping facilities and the inflow as well as outflow exhibits a consistent trend resembling an “inverted U” shape, suggesting an optimal range of shopping facility proportions that corresponds to higher levels of M-DBS combined usage. When the proportion of shopping facilities is approximately 32.5%, the inflow and outflow of M-DBS combined usage can be sustained at a relatively high level. Notably, the samples with a shopping facility proportion below 27.5% are sparse, rendering the regression relationship within this range of little meaningful (Tao & Cao, 2024). Population density exhibits a positive correlation with the inflow and outflow of M-DBS combined usage. When the population density reaches approximately 8000 person/km^2, the inflow and outflow of M-DBS combined usage experience a significant increase. Furthermore, the proportion of dining facilities is negatively correlated with the outflow of M-DBS combined usage. Specifically, when the proportion of dining facilities reaches 29% or 33%, the outflow experiences a marked decline. Moreover, the number of bus stops demonstrates a clear positive correlation with the inflow, yet it differs from the threshold exhibited by the number of bus stops in the low-density model. In the transportation hub model, a threshold effect on the inflow is observed when the number of bus stops reaches 12, whereas in the low-density model, this threshold stands at 6. This disparity is likely attributed to the higher-than-average number of bus stops in the transportation hub land-use pattern, further emphasizing the heterogeneous impacts of different land-use patterns.

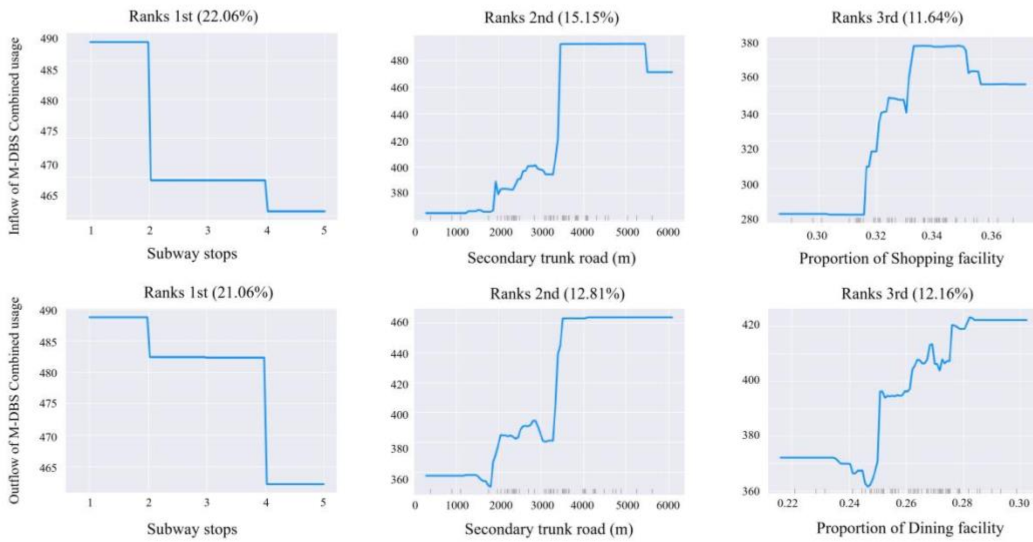

Figure 12. Nonlinear relationship under commercial core pattern

In the commercial core model, there exists a distinct nonlinear relationship between the number of subway stations, the length of secondary trunk roads, and the M-DBS combined usage. The findings indicate that the number of subway stations is negatively correlated with the usage of M-DBS. Specifically, when the number exceeds 2 and 4, it significantly impacts the inflow and outflow of M-DBS combined usage. The length of secondary trunk roads positively correlates with the inflow and outflow of M-DBS combined usage, with a marked increase observed when the length reaches approximately 3.2 km. Notably, in mixed-use pattern, the length of secondary trunk roads also positively influences the inflow and outflow of M-DBS combined usage, however, threshold value is different compared to the commercial core model. Furthermore, the proportion of dining facilities is positively correlated with the outflow of M-DBS combined usage. When this proportion reaches approximately 25% and 27.5%, the outflow of M-DBS combined usage maintains a relatively high level. It is worth mentioning that in the transportation hub model, the relationship between the proportion of dining facilities and the outflow is negative, and an obvious threshold effect is only observed when the proportion reaches 33%. These results demonstrate not only the inconsistent thresholds but also the changing correlation between the proportion of dining facilities in the transportation hub and commercial core models, further emphasizing the heterogeneous impacts of different land-use patterns.

5 Discussion

5.1 Relative importance of built environment on M-DBS combined use under different land-use patterns

We identified four land-use patterns around Shanghai metro stations: low-density, mixed-use, transportation hubs, and commercial cores. As expected, we found significant differences in the relative importance of built environment factors across these land-use patterns, with varying performance observed in the two feeder modes.

Specifically, under the low-density pattern, the proportion of service facility and bus accessibility significantly affect the inflow and outflow of M-DBS combined usage. This could be attributed to the higher demand for life service-related facilities in low-density areas, a finding also mentioned in studies by Zhou et al. (2023) and Liu, Zhou, et al. (2022). Moreover, compared to other patterns, bus accessibility has a more significant impact on M-DBS combined use in low-density areas. Generally, bike-sharing and bus are seen as competing modes, as noted by Zhou et al. (2023) and Guo et al. (2020), with the substitution effect between bus and bike-sharing being one of the key reasons for the higher impact of bus accessibility on M-DBS combined use (Bocker et al., 2020; Liu, Zhou, et al., 2022; Tong et al., 2023). However, in the low-density pattern, attention must also be given to the cooperative relationship between metro, bus, and bike-sharing. In the low-density pattern, metro stations need to extend public transport service coverage over a larger area. Given that bike-sharing is human-powered and that Shanghai's flat terrain imposes minimal physical burden on cycling, buses play a more critical role in extending the range of public transport services (Nawaro et al., 2021; Tilahun et al., 2007). It should be noted, however, that this complementarity between bus and bike-sharing as feeder modes may manifest differently in cities with steeper topography, where cycling's physical demands could shift the balance toward bus dependence. In this context, the cooperative model of connecting subway stations and bus stations through dockless shared bicycles, thereby allowing the bus network to further extend public transport service coverage, becomes more prominent. This is another key reason why bus accessibility has a higher influence on M-DBS combined use in low-density areas, a point that is rarely addressed in existing studies. We also found that under the low-density pattern, the impact of built environment attributes in different feeder modes was inconsistent. Compared to its influence on outflows, population density has a more significant effect on the inflows of M-DBS combined use. This is easy to understand: low-density areas are located on the urban periphery, and considering daytime commuting activities, residents have a higher demand for travel from the urban edge to the city center for work. Population density is the key factor driving this travel demand, which is why its effect on M-DBS combined use inflows is more significant (Chen et al., 2021; Luo et al., 2023; Ma et al., 2019). Additionally, this directional asymmetry may partly reflect weather sensitivity: during Shanghai's rainy summer mornings or cold winter days, some outbound commuters may switch from cycling to bus or other sheltered modes, thereby amplifying the observed inflow–outflow disparity in low-density peripheral areas.

Compared to the low-density pattern, the built environment variables that significantly influence M-DBS combined usage in the mixed-use pattern differ. We found that, in this pattern, road network indicators and land structure indicators play a more important role in influencing M-DBS inflows, particularly the density of secondary trunk roads and the proportion of service facility. Zhou et al. (2023) and Liu, Zhou, et al. (2022) obtained similar results. For outflows of M-DBS combined usage, land-use mix has a greater impact. This is not difficult to understand, as higher levels of land use combination make these areas more likely to become commuting destinations (Zhou et al., 2023). Notably, while the land-use mix is also high in the commercial core pattern, its impact on M-DBS combined use is relatively small. We attribute this difference to policy variations caused by spatial location factors. Specifically, metro stations in the commercial core pattern are mostly located in high-density areas such as the urban center, where commuters typically walk rather than cycle (Zhang, Hu, et al., 2024), a preference that is likely reinforced during Shanghai's frequent summer rainfall and winter cold spells, which disproportionately suppress cycling relative to short-distance walking. Additionally, planners restrict the use of dockless bike-sharing facilities in certain areas of the CBD to reduce collision risks in traffic-dense zones (Orvin et al., 2021; Zhang, Hu, et al., 2024). In contrast, as far as we know, mixed-use areas are mostly located in the satellite cities of Shanghai's urban area, where traffic planning does not explicitly restrict cycling. This regional policy difference can be attributed to spatial location factors, as noted in the research of Yan et al. (2024).

In the transportation hub pattern, in addition to population density, the proportion of shopping facilities significantly affects both the inflow and outflow of M-DBS combined usage. This finding reinforces the notion that a balance between transport and non-transport functions (such as leisure, entertainment, and public services) is crucial in land-use planning, a point emphasized by Liu, Fan, et al. (2022) regarding its importance in transportation planning. In the commercial core pattern, we found that subway accessibility and secondary trunk road density have particularly significant impacts on both the inflow and outflow of M-DBS combined usage. This is easy to understand: higher secondary trunk road density improves the connectivity of the traffic network within the commercial core, facilitating the movement of cyclists between starting points and metro stations. It is noteworthy that, regarding metro accessibility, our results differ from those of Zhou et al. (2023), who found that higher metro accessibility is associated with increased use of shared bikes around metro stations. However, this relationship is not absolute. We found that although higher metro station density can improve subway accessibility within a certain range, an excessively high metro station density in the commercial core may dilute the inherent travel demand. As pointed out by Guo et al. (2020, 2021), ignoring the heterogeneity of land-use patterns near metro stations can hinder metro accessibility from fully exerting its positive effects. Notably, this conclusion, derived from Shanghai's flat topography, warrants caution when extended to hilly cities, where slope gradients may moderate or even dominate the relationship between metro accessibility and cycling usage.

5.2 Nonlinear impacts of built environment on M-DBS combined use under different land-use patterns

We not only identified the heterogeneous impacts of built environment on M-DBS combined usage across different land-use patterns but also uncovered the nonlinear relationships. These nonlinear relationships vary across different land-use patterns, meaning that built environment attributes only exert significant effects on M-DBS combined use once they reach certain threshold levels. Moreover, these threshold levels and the shape of the nonlinear relationships change depending on the land-use pattern.

Specifically, we found that in both low-density and mixed-use patterns, bus accessibility shows a clear positive correlation with M-DBS outflows. However, compared with the low-density mode, the threshold at which bus accessibility takes effect in the mixed-use mode has undergone a lateral shift. Similarly, the threshold at which the proportion of service facilities influences M-DBS combined usage differs between the low-density and mixed-use patterns. This highlights the heterogeneous impact of land-use patterns on the nonlinear relationship between built environment attributes and M-DBS combined usage, a topic that has been scarcely addressed in prior research. It should be emphasized that these threshold values are estimated under Shanghai's flat terrain conditions; in cities with pronounced topographic variation, slope effects may interact with or partially override the land-use-driven thresholds identified here, and the specific threshold levels may not be directly transferable. Although Yan et al. (2024) separately revealed the heterogeneity and nonlinear effects of the built environment, they did not delve deeply into how the nonlinear relationship affected by heterogeneity changes. Furthermore, we discovered a non-monotonic "inverted U-shaped" nonlinear relationship between the proportion of shopping facilities and M-DBS combined use in the transportation hub pattern. This suggests that, for high levels of M-DBS combined use in transportation hubs, there exists an optimal proportion of shopping facilities. Previous studies, such as those by Chen et al. (2021) and Gao, Yang, et al. (2023), have also highlighted this phenomenon. However, this optimal proportion may exhibit seasonal variation: during Shanghai's hot, rainy summer months, reduced cycling tolerance may shift the peak of the inverted-U curve leftward relative to annual-average conditions, a possibility that warrants further investigation with finer temporal resolution data. Our findings deepen these insights by reaffirming the conditions under which these results hold, specifically emphasizing that in transportation hub areas, for high levels of M-DBS combined use, there is likely an optimal proportion of shopping facilities. In the commercial core pattern, based on our nonlinear results, we found that when there are more than two metro stations within a 1 km service radius, both the inflow and outflow of M-DBS usage at each individual station significantly decrease. This negative correlation becomes even more pronounced when the number of metro stations exceeds four. Guo et al. (2021) also observed this phenomenon, suggesting that too many metro stations serving overlapping areas can dilute the total M-DBS usage, reducing the usage at each individual station. This dilution effect may be further intensified during adverse weather, as commuters gravitate toward a subset of better-sheltered or more accessible stations, concentrating usage and exacerbating the decline at less-favored stations. Planners should be aware of this; simply investing transportation resources in the core area does not always solve the problem (Lin et al., 2018; Zhao & Li, 2017).

Additionally, it is important to note that the variation in nonlinear relationships between built environment attributes and M-DBS combined use across different land-use patterns is not only reflected in shifts in threshold levels but also in changes in the direction of correlation. The correlation between the proportion of dining facilities and outflows of M-DBS combined usage is opposite in the transportation hub pattern and the commercial core pattern. These results further confirm the heterogeneous impact of land-use patterns on the nonlinear relationship between built environment attributes and M-DBS combined use. They also underscore the importance for planners to consider the differences in land-use patterns and avoid one-size-fits-all policies, as such measures may not always yield consistent results and may even have counterproductive effects. Therefore, there is an urgent need for personalized and targeted built environment interventions that can flexibly adapt to different metro station land-use patterns as well as to local climatic conditions and topographical constraints, promoting the integration of metro and dockless bikes (Zhou et al., 2023).

A further observation concerns the interpretation of threshold values across patterns. The service facility proportion at which M-DBS inflow declines is 25% in the low-density pattern and 29% in the mixed-use pattern. Notably, although these two values are numerically close, this does not mean that the specific implementation measures are also similar; the two cannot be directly compared when designing actual intervention measures. The nonlinear relationships and their thresholds are conditional on the land-use pattern within which they are estimated: a given percentage of service facilities corresponds to substantially different absolute numbers of establishments in the two contexts, because the total POI stock in a mixed-use buffer is far larger and more diverse than in a low-density buffer. Consequently, a four-percentage-point adjustment, reducing service facility proportion from a level above the threshold to a level below it, would imply a far greater absolute change in the number of service facilities in the mixed-use pattern than in the low-density pattern. In low-density areas, where the baseline POI count is modest, shifting the service share by a few percentage points requires altering only a handful of establishments; in mixed-use satellite towns, where the baseline POI count runs into the hundreds, the same percentage shift entails a substantially larger physical reconfiguration of the station-area land-use mix. This asymmetry has direct practical implications. In low-density contexts, planners can achieve meaningful changes in service concentration through targeted, small-scale interventions, for instance, limiting a few additional service-oriented land parcels around specific metro stations. In mixed-use contexts, where comparable percentage adjustments require far more extensive and often infeasible land use restructuring, planners should consider alternatives that enhance M-DBS integration without altering land-use composition.

6 Policy implications

Our findings highlight the importance for urban policymakers and planners to recognize the differential impacts of built environment interventions under specific land-use patterns when promoting the integrated use of metro and bike-sharing systems. Because the influential variables, their threshold levels, and even the direction of their effects vary across the four land-use patterns, a given built environment intervention, such as expanding bus routes, adjusting land-use mix, or adding metro stations, is likely to produce different outcomes depending on the pattern context of the target station area. Recognizing this heterogeneity enables planners to allocate resources to the interventions most effective for each pattern, rather than applying a single strategy across all station areas and bearing the cost of uneven results. Therefore, current planning should adopt more personalized and targeted built environment interventions. Drawing on the Shanghai case, we outline the following pattern specific recommendations.

For the low-density pattern, we recommend increasing the number of bus stops in suburban areas to improve bus accessibility within the service area of metro stations. This can help effectively integrate bike-sharing, bus, and metro systems, fostering stronger M-DBS combined use. This not only expands the service range of metro stations but is also crucial for the development of integrated transportation systems in non-core urban areas. Our nonlinear results show that the positive effect of bus stop count on M-DBS inflow and outflow becomes apparent at approximately six stops and plateaus at around seven stops. Therefore, for low-density stations along the outer segments of Lines 5, 9, and 11, where current bus stop counts typically fall below six, bus route expansion should be prioritized; stations already at seven or more stops are unlikely to benefit further from additional bus connections.

Furthermore, to enhance the integration of metro and bike-sharing in low-density areas, planners should carefully control the proportion of service facilities around metro stations. Our results show that when the share of service facilities reaches 25%, M-DBS combined usage drops sharply. The mechanism behind this phenomenon may be that an excessively high density of service facilities creates a highly convenient walkable living circle around the station, inducing some travelers who would otherwise use cycling for medium to long distance metro access to instead rely solely on walking for short local trips, thereby generating a substitution effect of walking over cycling (Guo et al., 2021). It should be clearly noted that this does not mean we wish to reduce walking or weaken walkability around stations; walking remains a core component of a sustainable transport system, and the walkability supported by existing service facilities should be fully protected. The real concern of this recommendation is that, in the specific spatial context of low-density suburbs, the core value of bike-sharing is to expand the catchment area of a metro station from the walkable radius of approximately 800 meters to 2 to 3 kilometers. Once the feeder function of cycling is significantly substituted by excessive agglomeration of service facilities, the effective service range of the metro station contracts sharply, forcing potential travelers beyond comfortable walking distance to shift to private motor vehicles, thereby undermining overall transport accessibility at the system level. Therefore, the policy focus is not to reduce existing service facilities or restrict walking, but rather to avoid adding new commercial facilities around stations that would push the service facility share beyond the approximate 25% threshold, preventing suburban metro stations from shifting from regional transit hubs toward purely localized walkable community centers. For stations already at or near this threshold, priority should be given to investing in bicycle parking facilities and cycling connectivity, improving cycling feeder conditions while fully preserving existing walkability, thereby achieving synergistic benefits between walking and cycling in the metro feeder system. This is particularly critical for solving the last-mile problem in low-density patterns and for suburbs that rely on bike-sharing to extend the service range of metro stations.

Moreover, it must be acknowledged that this study did not incorporate safety related variables, such as protected bicycle lanes, street lighting, and social safety concerns. These factors may disproportionately affect female riders (Wang et al., 2019), especially in low-density areas. In contexts where cycling infrastructure is deficient or perceived safety is low, walking or bus may remain the only accepted feeder modes, implying that the built environment effects we observed may be partly moderated by unmeasured safety conditions. Future research should incorporate safety variables to more comprehensively evaluate M-DBS promotion strategies in low-density areas.

In areas with a high degree of land-use mix, planners should pay attention to population density. The nonlinear effects of population density suggest that once a threshold is reached, the marginal benefits for M-DBS combined use diminish, implying that population densification strategies should be implemented within a certain range of population density. Our results identify this threshold at approximately 8,000 persons per square kilometer. In mixed-use satellite towns such as Songjiang and Jiading, densification efforts should target station areas currently below this level. Additionally, secondary trunk road length shows differentiated thresholds for inflow (2,200 meters and 4,200 meters) and outflow (2,200 meters), indicating that extending the secondary road network up to approximately 2,200 meters yields the broadest benefit for both feeder directions, while further extension beyond that length requires careful trade-offs.

Planners who develop bike-friendly or bike-priority measures near transportation hubs need to carefully consider the proportion of land related to commercial activities such as shopping facilities. They must ensure that the land share of such commercial facilities attracts enough residents and travelers while also preventing the substitution effect of walking over cycling caused by excessively high shopping facility density, thereby maintaining a reasonable balance between walking and cycling. Our results show that this relationship follows an inverted U shape, with M-DBS usage peaking at a shopping facility proportion of approximately 32.5%. For major transfer stations such as People's Square, Xujiahui, and Century Avenue, where shopping facilities are already highly agglomerated, planners should treat 32.5% as a reference ceiling. For stations approaching this threshold, priority should be given to investing in bicycle parking facilities and cycling access rather than adding further commercial floor space.

For the city's core area, often located in the CBD, planners should recognize that simply increasing metro station density may not attract more M-DBS usage as expected. From a planning perspective, we recommend that, while ensuring that existing metro stations can accommodate current travel demands in the CBD, the future expansion of metro station density should be carefully controlled, thereby saving urban transportation resources while maintaining high levels of M-DBS combined use in core urban areas. Specifically, our results indicate that when more than two metro stations fall within a one-kilometer radius, both M-DBS inflow and outflow begin to decline, with the decline accelerating beyond four stations. For Shanghai's central business district, particularly the dense station cluster bounded by Lines 1, 2, and 8, future metro expansion should avoid adding new stations within one kilometer of existing ones unless required for capacity relief. Resources saved could be redirected to improving bicycle lane connectivity, which our results show is positively associated with M-DBS usage in this pattern.

Beyond these pattern specific thresholds, a cross-cutting consideration is that threshold values cannot be interpreted in isolation from the baseline land-use intensity of each pattern. As discussed in Section 5.2, a given percentage point adjustment in a land-use variable corresponds to substantially different absolute changes in the number of facilities across patterns, because the total POI stock differs by an order of magnitude between low-density and mixed-use contexts. Planners should therefore evaluate not only whether a station's built environment exceeds a given threshold, but also the practical feasibility of bringing it below that threshold. When land use-side adjustments are too costly, infrastructure or operational interventions, such as bicycle lane expansion or reconfiguring parking zones, offer alternative pathways to enhance M-DBS integration.

Several limitations should be clarified when applying the above policy implications. First, the policy implications derived from this study are based on the Shanghai case and are to some extent context specific. When transferring them to other cities, similarities with Shanghai in terms of city size, public transport systems, and land-use patterns need to be considered, and the validity of the recommendations requires further testing. However, the more important contribution of this study is to provide a general analytical framework that identifies fine-grained built environment interventions based on nonlinear relationships. This framework can be transferred and adapted to other cities. Second, a fundamental methodological qualification applies to all the above recommendations. This study treats the built environment as given and does not model the feedback effect of improved metro accessibility on land use intensification around stations. In reality, transport improvements can induce land use change over time, which in turn reshapes travel demand. This is a dynamic interaction not captured by the cross-sectional design of this study. The thresholds and variable rankings reported here should therefore be understood as conditional on current land-use conditions. Where a station's surrounding land use is undergoing rapid change, planners should re-evaluate the local built environment profile before applying pattern specific guidance thresholds and continuously monitor and adjust during implementation.

7 Conclusion

Existing studies on the impact of the built environment on travel behavior are often constrained by generalized linear assumptions and rarely account for spatial heterogeneity across different land-use contexts, which limits a deeper understanding of the mechanisms underlying Metro and Dockless Bike Sharing (M-DBS) combined usage. To address this research gap and answer the key question of how the built environment nonlinearly influences M-DBS combined usage under different land-use patterns, thereby providing scientific support for improving public transport system efficiency and solving first-mile/last-mile connectivity problems, this study takes Shanghai as a case and adopts a hybrid framework that integrates clustering algorithms and machine learning. This framework accurately identifies land-use patterns around metro stations and deeply models and reveals the complex nonlinear relationships between the built environment and M-DBS combined usage (both inflow and outflow). The results uncover the complex nonlinear mechanisms and spatial heterogeneity of built environment effects on M-DBS combined usage, providing planners with a data-driven decision-making framework that accounts for both spatial heterogeneity and nonlinear threshold effects.

Empirically, this study identifies four typical land-use patterns around metro stations—low-density, mixed-use, transportation hubs, and commercial cores—and demonstrates that built environment effects on M-DBS feeder trips are fundamentally pattern-dependent. The relative importance of built environment variables, their threshold levels, and even the direction of their effects vary substantially across patterns. For the same variable, nonlinear relationships manifest as threshold shifts and correlation reversals depending on the land-use context. These findings confirm that ignoring land-use heterogeneity leads to biased estimates and ineffective interventions.

Based on these empirical insights, this study offers the following policy-oriented implications. First, pattern-specific threshold identification enables resource allocation efficiency: rather than applying uniform strategies across all station areas, planners should prioritize interventions that are most effective for each land-use pattern. For low-density suburbs, improving bus accessibility within metro station service areas can substantially enhance M-DBS integration, while in mixed-use satellite towns, population density management and secondary road network extension offer greater returns. For transportation hubs, balancing commercial facility proportions within an optimal range is critical, whereas in commercial cores, controlling excessive metro station density and improving bicycle lane connectivity are more relevant. Second, the identified threshold effects provide actionable planning targets: built environment adjustments near critical thresholds can produce substantial M-DBS usage gains at relatively low cost, whereas investments beyond saturation thresholds yield diminishing or even negative returns. This allows planners to avoid overinvestment and focus resources where marginal benefits are highest. Third, the observed correlation reversals across patterns underscore the danger of one-size-fits-all policies. A variable that enhances M-DBS usage in one pattern may suppress it in another, necessitating careful diagnosis of the local land-use context before intervention. Fourth, a cross-cutting principle is that thresholds should be interpreted relative to baseline land-use intensity; where land-side adjustments are too costly, infrastructure improvements such as bicycle lane expansion offer alternative pathways. Finally, these policy implications should be applied with awareness of two limitations: the findings are context-specific to Shanghai and require validation before transfer to other cities, and the cross-sectional design does not capture dynamic feedback between transport improvements and land use change over time. Planners should re-evaluate local conditions periodically, especially in areas undergoing rapid development.

Despite these contributions, this study has several limitations. First, reliance on POI data and a fixed 100-meter transfer threshold may introduce minor biases, and the absence of user IDs limits the validation of individual joint travel chains. More importantly, this study did not incorporate the effects of weather and topography on M-DBS combined usage. The case city Shanghai is flat, and slope is generally not a major obstacle to combined usage; however, its subtropical monsoon climate brings extreme weather such as the plum rain season, summer heat, and typhoons, which significantly affect users' willingness to cycle. Without these environmental variables, the current model struggles to capture the temporal dynamics of combined usage. Moreover, due to differences in spatial layouts, urban structures, and cycling cultures among cities, the findings of this study may be more applicable to megacities similar to Shanghai. Consequently, the identified land-use patterns may not fully represent all possible patterns around metro stations, and the nonlinear results are also somewhat constrained by the case study. This limitation affects the transferability of policy implications. Nevertheless, while the specific implementation of the recommendations in other cities may not achieve the same expected outcomes, the policy insights remain valuable for reference and comparison. Future research should integrate multi-source spatiotemporal data such as Digital Elevation Models (DEM), high-resolution meteorological data, and mobile phone signaling to quantify the effects of topography and weather, and to more accurately identify true joint travel behavior. It should also be acknowledged that dockless bike-sharing systems face operational challenges such as vandalism and bike deterioration, which may affect the availability and reliability of bike-sharing services

(Chen, 2019; Leszczynski et al., 2025). In Shanghai, reports indicate that bike destruction and misuse remain non-negligible issues: Mobike recorded over 270,000 reports of bike damage and misuse in 2019, with one in seven coming from Shanghai residents (Xin & Fan, 2017). Future research could incorporate these operational factors to provide a more comprehensive assessment of M-DBS integration strategies. Furthermore, Furthermore, the cross-sectional design of this study treats land use as exogenously given and does not capture the bidirectional interaction between transport and land use. Subsequently, we will further investigate the interaction between the two and conduct a detailed analysis. Moreover, the analytical framework should be extended to cities with different topographic characteristics (e.g., mountainous or hilly cities) and climatic conditions to test the generalizability of the "threshold shift" and "correlation reversal" mechanisms, thereby providing more robust policy references for integrated transport planning worldwide.

Acknowledgments

This work was supported by the Higher education institutions under the jurisdiction of Heilongjiang Province Basic Research Business Expenses for Research Projects (Grant No. 2025-KYYWF-ZR0443), National Natural Science Foundation of China (Grant No. 52272332), and the Fundamental Research Funds for the Central Universities (Grant No. HIT.OCEF.2022026). Special thanks to Professor Jason (Xinyu) Cao of the University of Minnesota and Professor Ding Chuan of Beihang University for their writing suggestions and detailed comments, which greatly improved the quality of this paper.

Author contribution

The authors confirm their contribution to the paper as follows: Conceptualization, Methodology, Writing - original draft: Yantang Zhang; Data curation, Formal analysis, Writing - review & editing: Xiaowei Hu.

The authors declare that they have no conflict of interest.

Appendix

Appendix available as a supplemental file at https://doi.org/10.5198/jtlu.2026.2859.

References

Bean, R., Pojani, D., & Corcoran, J. (2021). How does weather affect bikeshare use? A comparative analysis of forty cities across climate zones. *Journal of Transport Geography, 95*, 103155. https://doi.org/10.1016/j.jtrangeo.2021.103155

Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy c-means clustering algorithm. *Computers & Geosciences, 10*(2–3), 191–203. https://doi.org/10.1016/0098-3004(84)90020-7

Böcker, L., Anderson, E., Uteng, T. P., & Throndsen, T. (2020). Bike sharing use in conjunction to public transport: Exploring spatiotemporal, age and gender dimensions in Oslo, Norway. *Transportation Research Part A: Policy and Practice, 138*, 389–401. https://doi.org/10.1016/j.tra.2020.06.009

Chang, X., Wu, J., He, Z., Li, D., Sun, H., & Wang, W. (2020). Understanding user's travel behavior and city region functions from station-free shared bike usage data. *Transportation Research Part F: Traffic Psychology and Behaviour, 72*, 81–95. https://doi.org/10.1016/j.trf.2020.03.018

Chen, R. (2019). "Bike litter" and obligations of the platform operators: Lessons from China's dockless sharing bikes. *Computer Law & Security Review, 35*(5), 105317. https://doi.org/10.1016/j.clsr.2019.03.011

Chen, E., & Ye, Z. (2021). Identifying the nonlinear relationship between free-floating bike sharing usage and built environment. *Journal of Cleaner Production, 280*, 124281. https://doi.org/10.1016/j.jclepro.2020.124281

Chen, M., Wang, D., Sun, Y., Waygood, E. O. D., & Yang, W. (2020). A comparison of users' characteristics between station-based bikesharing system and free-floating bikesharing system: Case study in Hangzhou, China. *Transportation, 47*, 689–704. https://doi.org/10.1007/s11116-018-9910-7

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). https://doi.org/10.1145/2939672.2939785

Chen, Z., Van Lierop, D., & Ettema, D. (2020). Dockless bike-sharing systems: What are the implications? *Transport Reviews, 40*(3), 333–353. https://doi.org/10.1080/01441647.2019.1710306

Cheng, L., Wang, K., De Vos, J., Huang, J., & Witlox, F. (2022). Exploring non-linear built environment effects on the integration of free-floating bike-share and urban rail transport: A quantile regression approach. *Transportation Research Part A: Policy and Practice, 162,* 175–187. https://doi.org/10.1016/j.tra.2022.05.022

China Academy of Urban Planning & Design. (2022). *2022 annual report on sharing bikes and electric bikes in major Chinese cities*. CAUPD. https://www.199it.com/archives/1507123.html

Ding, C., Cao, X., & Wang, Y. (2018). Synergistic effects of the built environment and commuting programs on commute mode choice. *Transportation Research Part A: Policy and Practice, 118*, 104–118. https://doi.org/10.1016/j.tra.2018.08.041

Ding, C., Cao, X., Dong, M., Zhang, Y., & Yang, J. (2019). Non-linear relationships between built environment characteristics and electric-bike ownership in Zhongshan, China. *Transportation Research Part D: Transport and Environment, 75*, 286–296. https://doi.org/10.1016/j.trd.2019.09.005

Dong, J., Chen, S., Li, W., Zhou, Y., & Si, H. (2022). Evaluation of dockless bike-sharing transfer services around metro stations considering spatial heterogeneity. *Journal of Advanced Transportation*, *2022,* 7732485. https://doi.org/10.1155/2022/7732485

El-Assi, W., Salah Mahmoud, M., & Nurul Habib, K. (2017). Effects of built environment and weather on bike sharing demand: a station level analysis of commercial bike sharing in Toronto. *Transportation, 44*(3), 589–613. https://doi.org/10.1007/s11116-015-9669-z

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics, 29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Galster, G. C. (2018). Nonlinear and threshold effects related to neighborhood: Implications for planning and policy. *Journal of Planning Literature, 33*(4), 492–508. https://doi.org/10.1177/0885412218793693

Gao, F., Li, S., Tan, Z., Wu, Z., Zhang, X., Huang, G., & Huang, Z. (2021). Understanding the modifiable areal unit problem in dockless bike sharing usage and exploring the interactive effects of built environment factors. *International Journal of Geographical Information Science, 35*(9), 1905–1925. https://doi.org/10.1080/13658816.2020.1863410

Gao, K., Yang, Y., Gil, J., & Qu, X. (2023). Data-driven interpretation on interactive and nonlinear effects of the correlated built environment on shared mobility. *Journal of Transport Geography, 110*, 103604. https://doi.org/10.1016/j.jtrangeo.2023.103604

Gao, W., Hu, X., & Wang, N. (2023). Exploring spatio-temporal pattern heterogeneity of dockless bike-sharing system: Links with cycling environment. *Transportation Research Part D: Transport and Environment, 117*, 103657. https://doi.org/10.1016/j.trd.2023.103657

Guo, Y., & He, S. Y. (2020). Built environment effects on the integration of dockless bike-sharing and the metro. *Transportation Research Part D: Transport and Environment, 83*, 102335. https://doi.org/10.1016/j.trd.2020.102335

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*, 102594. https://doi.org/10.1016/j.scs.2020.102594

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer Nature.

Hong, H., Liu, J., & Zhu, A. X. (2020). Modeling landslide susceptibility using LogitBoost alternating decision trees and forest by penalizing attributes with the bagging ensemble. *Science of the Total Environment*, 718, 137231. https://doi.org/10.1016/j.scitotenv.2020.137231

Hu, S., Chen, M., Jiang, Y., Sun, W., & Xiong, C. (2022). Examining factors associated with bike-and-ride (BnR) activities around metro stations in large-scale dockless bikesharing systems. *Journal of Transport Geography, 98*, 102171. https://doi.org/10.1016/j.jtrangeo.2021.103271

Ji, Y., Fan, Y., Ermagun, A., Cao, X., Wang, W., & Das, K. (2017). Public bicycle as a feeder mode to rail transit in China: The role of gender, age, income, trip purpose, and bicycle theft experience. *International Journal of Sustainable Transportation, 11*(4), 308–317. https://doi.org/10.1080/15568318.2016.1253802

Kilani, M., Bennaya, S., & Zerguini, S. (2024). Complementarity and substitution between public transport and bicycles. *Journal of Transport and Land Use, 17*(1), 781–803. https://doi.org/10.5198/jtlu.2024.2465

Kim, S. H., & Mokhtarian, P. L. (2023). Finite mixture (or latent class) modeling in transportation: Trends, usage, potential, and future directions. *Transportation Research Part B: Methodological, 172*, 134–173. https://doi.org/10.1016/j.trb.2023.03.001

Kumar, P. P., Sekhar, C. R., & Parida, M. (2020). Identification of neighborhood typology for potential transit-oriented development. *Transportation Research Part D: Transport and Environment, 78*, 102186. https://doi.org/10.1016/j.trd.2019.11.015

Leszczynski, A., Cinnamon, J., Asa, S., & Jahiu, L. (2025). Docklessness, aesthetic governance, and the urban 'micromobility mess.' *Urban Studies, 62*(12), 2508–2525. https://doi.org/10.1177/00420980251316773

Li, W., Chen, S., Dong, J., & Wu, J. (2021). Exploring the spatial variations of transfer distances between dockless bike-sharing systems and metros. *Journal of Transport Geography, 92*, 103032. https://doi.org/10.1016/j.jtrangeo.2021.103032

Li, Z., Han, Z., Xin, J., Luo, X., Su, S., & Weng, M. (2019). Transit oriented development among metro station areas in Shanghai, China: Variations, typology, optimization and implications for land-use planning. *Land Use Polic*y*, 82*, 269–282. https://doi.org/10.1016/j.landusepol.2018.12.003

Lin, D., Zhang, Y., Zhu, R., & Meng, L. (2019). The analysis of catchment areas of metro stations using trajectory data generated by dockless shared bikes. *Sustainable Cities and Society, 49*, 101598. https://doi.org/10.1016/j.scs.2019.101598

Lin, J. J., Zhao, P., Takada, K., Li, S., Yai, T., & Chen, C. H. (2018). Built environment and public bike usage for metro access: A comparison of neighborhoods in Beijing, Taipei, and Tokyo. *Transportation Research Part D: Transport and Environment, 63*, 209–221. https://doi.org/10.1016/j.trd.2018.05.007

Liu, H., Zhang, W., Wang, S., Cheng, Z., Wei, L., & Huang, W. (2024). Exploring the effect of built environment on spatiotemporal evolution of traffic congestion using a novel GTWR model: a case study of Hefei, China. *Transportation Letters*, 1–12. https://doi.org/10.1080/19427867.2024.2396773

Liu, S., Zhou, C., Rong, J., Bian, Y., & Wang, Y. (2022). Concordance between regional functions and mobility features using bike-sharing and land-use data near metro stations. *Sustainable Cities and Society, 84*, 104010. https://doi.org/10.1016/j.scs.2022.104010

Liu, X., Fan, J., Li, Y., Shao, X., & Lai, Z. (2022). Analysis of integrated uses of dockless bike sharing and ridesourcing with metros: A case study of Shanghai, China. *Sustainable Cities and Society, 82*, 103918. https://doi.org/10.1016/j.scs.2022.103918

Liu, J., Thill, J.-C., Du, J., & Liu, K. (2026). Spatial heterogeneity in the association between land-use patterns and subway accessibility: Micro-level evidence from Shanghai. http://dx.doi.org/10.2139/ssrn.7082927

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_files/paper/2017

Luo, S., & Nie, Y. M. (2023). Integrated design of a bus-bike system considering realistic route options and bike availability. *Transportation Research Part C: Emerging Technologies, 153*, 104192. https://doi.org/10.1016/j.trc.2023.104192

Ma, J., Zheng, C., Yu, M., Shen, J., Zhang, H., & Wang, Y. (2024). The analysis of spatio-temporal characteristics and determinants of dockless bike-sharing and metro integration. *Transportation Letters, 16*(2), 182–195. https://doi.org/10.1080/19427867.2023.2170493

Ma, X., Zhang, X., Li, X., Wang, X., & Zhao, X. (2019). Impacts of free-floating bikesharing system on public transit ridership. *Transportation Research Part D: Transport and Environment, 76*, 100–110. https://doi.org/10.1016/j.trd.2019.09.014

Martin, E. W., & Shaheen, S. A. (2014). Evaluating public transit modal shift dynamics in response to bikesharing: a tale of two US cities. *Journal of Transport Geography, 41*, 315–324. https://doi.org/10.1016/j.jtrangeo.2014.06.026

Nawaro, Ł. (2021). E-scooters: competition with shared bicycles and relationship to public transport. *International Journal of Urban Sustainable Development, 13*(3), 614–630. https://doi.org/10.1080/19463138.2021.1981336

Orvin, M. M., & Fatmi, M. R. (2021). Why individuals choose dockless bike sharing services? *Travel Behaviour and Society, 22*, 199–206. https://doi.org/10.1016/j.tbs.2020.10.001

Shanghai Municipal Bureau of Planning and Natural Resources. (n.d.). Shanghai master urban plan (2017–2035). https://ghzyj.sh.gov.cn

Shanghai Municipal Transportation Commission. (2023). *Shanghai Municipal Transportation Commission*. https://jtw.sh.gov.cn

Shen, Y., Zhang, X., & Zhao, J. (2018). Understanding the usage of dockless bike sharing in Singapore. *International Journal of Sustainable Transportation, 12*(9), 686–700. https://doi.org/10.1080/15568318.2018.1429696

Shi, X., Zhao, J., He, J., & Xu, H. (2023). Exploring year-to-year spatiotemporal changes in cycling patterns for bike-sharing system in the pre-, during and post-pandemic periods. *Sustainable Cities and Society, 98*, 104814. https://doi.org/10.1016/j.scs.2023.104814

Steinley, D., & Brusco, M. J. (2011). Choosing the number of clusters in K-means clustering. *Psychological Methods, 16*(3), 285. https://doi.org/10.1037/a0023346

Sun, Y., Wang, Y., & Wu, H. (2024). How does the urban built environment affect dockless bikesharing-metro integration cycling? Analysis from a nonlinear comprehensive perspective. *Journal of Cleaner Production, 449*, 141770. https://doi.org/10.1016/j.jclepro.2024.141770

Tao, T., & Cao, J. (2024). Ineffective built environment interventions: How to reduce driving in American suburbs? *Transportation Research Part A: Policy and Practice*, *179*, 103924. https://doi.org/10.1016/j.tra.2023.103924

Tilahun, N. Y., Levinson, D. M., & Krizek, K. J. (2007). Trails, lanes, or traffic: Valuing bicycle facilities with an adaptive stated preference survey. *Transportation Research Part A: Policy and Practice, 41*(4), 287–301. https://doi.org/10.1016/j.tra.2006.09.007

Tong, Z., Zhu, Y., Zhang, Z., An, R., Liu, Y., & Zheng, M. (2023). Unravel the spatio-temporal patterns and their nonlinear relationship with correlates of dockless shared bikes near metro stations. *Geo-Spatial Information Science, 26*(3), 577–598. https://doi.org/10.1080/10095020.2022.2137857

Wang, J., & Lindsey, G. (2019). Neighborhood socio-demographic characteristics and bike share member patterns of use. *Journal of Transport Geography, 79*, 102475. https://doi.org/10.1016/j.jtrangeo.2019.102475

Xin, X., & Fan, W. (2017, March 16). *Has Shanghai halted the deployment of shared bicycles? Mobike claims it has "not received any notification."* Daily Economic News. https://m.cnr.cn/tech/20170316/t20170316_523661083.html

Yan, Y., & Chen, Q. (2024). Spatial heterogeneity and nonlinearity study of bike-sharing to subway connections from the perspective of built environment. *Sustainable Cities and Society, 114*, 105766. https://doi.org/10.1016/j.scs.2024.105766

Yang, X. H., Cheng, Z., Chen, G., Wang, L., Ruan, Z. Y., & Zheng, Y. J. (2018). The impact of a public bicycle-sharing system on urban public transport networks. *Transportation Research Part A: Policy and Practice, 107*, 246–256. https://doi.org/10.1016/j.tra.2017.10.017

Yin, J., Yu, D., Yin, Z., Wang, J., & Xu, S. (2013). Modelling the combined impacts of sea-level rise and land subsidence on storm tides induced flooding of the Huangpu River in Shanghai, China. *Climatic Change, 119*(3), 919–932. https://doi.org/10.1007/s10584-013-0749-9

Yin, Z., Rybarczyk, G., Zheng, A., Su, L., Sun, B., & Yan, X. (2024). Shared micromobility as a first-and last-mile transit solution? Spatiotemporal insights from a novel dataset. *Journal of Transport Geography, 114*, 103778. https://doi.org/10.1016/j.jtrangeo.2023.103778

Zeng, Q., Wu, H., Zhou, L., Huang, G., Li, Y., & Dewancker, B. J. (2024). Toward pedestrian-friendly cities: Nonlinear and interaction effects of building density on pedestrian volume. *Journal of Transport Geography, 119*, 103954. https://doi.org/10.1016/j.jtrangeo.2024.103954

Zhang, Y., & Hu, X. (2024). The nonlinear impact of cycling environment on bicycle distance: A perspective combining objective and perceptual dimensions. *Journal of Transport and Land Use, 17*(1), 241–267. https://doi.org/10.5198/jtlu.2024.2434

Zhang, Y., Chen, X. J., Gao, S., Gong, Y., & Liu, Y. (2024). Integrating smart card records and dockless bike-sharing data to understand the effect of the built environment on cycling as a feeder mode for metro trips. *Journal of Transport Geography, 121*, 103995. https://doi.org/10.1016/j.jtrangeo.2024.103995

Zhang, Y., Hu, X., Wang, H., & An, S. (2024). How does the built environment affect the usage efficiency of dockless-shared bicycle? An exploration of time-varying nonlinear relationships. *Journal of Transport Geography, 118*, 103908. https://doi.org/10.1016/j.jtrangeo.2024.103908

Zhang, Y., Li, X., Wang, A., Bao, T., & Tian, S. (2015). Density and diversity of OpenStreetMap road networks in China. *Journal of Urban Management, 4*(2), 135–146. https://doi.org/10.1016/j.jum.2015.10.001

Zhang, Y., Thomas, T., Brussel, M., & Van Maarseveen, M. (2017). Exploring the impact of built environment factors on the use of public bikes at bike stations: Case study in Zhongshan, China. *Journal of Transport Geography, 58*, 59–70. https://doi.org/10.1016/j.jtrangeo.2016.11.014

Zhao, P., & Li, S. (2017). Bicycle-metro integration in a growing city: The determinants of cycling as a transfer mode in metro station areas in Beijing. *Transportation Research Part A: Policy and Practice, 99*, 46–60. https://doi.org/10.1016/j.tra.2017.03.003

Zhou, X., Dong, Q., Huang, Z., Yin, G., Zhou, G., & Liu, Y. (2023). The spatially varying effects of built environment characteristics on the integrated usage of dockless bike-sharing and public transport. *Sustainable Cities and Society, 89*, 104348. https://doi.org/10.1016/j.scs.2022.104348

Zhu, Y., & Diao, M. (2020). Understanding the spatiotemporal patterns of public bicycle usage: A case study of Hangzhou, China. *International Journal of Sustainable Transportation, 14*(3), 163–176. https://doi.org/10.1080/15568318.2018.1538400