International Journal of Geo-Information

MDPI

Article

Modeling Spatiotemporal Associations Between Multiscale Urban Greenness and Metro Ridership: Evidence from Nanjing, China

Xinyu Hu [1], Yong Zhang [2], Zhenjun Zhu [1], Baorui Han [1,*] and Lu Yu [1]

[1] College of Automobile and Traffic Engineering, Nanjing Forestry University, Nanjing 210037, China
[2] Shenzhen Research Institute of Northwestern Polytechnical University, Shenzhen 518057, China
* Correspondence: hbr@njfu.edu.cn

Abstract

Urban green space may be associated with metro ridership, but existing studies commonly rely on a single greenness indicator and provide limited evidence on the differences between area-level vegetation coverage and street-level visible greenery across travel contexts. This study compares the model-based predictive attribution and nonlinear association patterns of the Normalized Difference Vegetation Index (NDVI) and Green View Index (GVI) for metro ridership under weekday, weekend, and holiday scenarios. Using Nanjing Metro as a case study, a Local XGBoost Ensemble was developed by combining local spatial modeling with distance-weighted ensemble prediction. Spatial Ensemble SHAP and partial dependence analysis were used to examine the spatially varying predictive attribution and nonlinear modeled associations of NDVI and GVI. The Local XGBoost Ensemble achieved R^2 values of 0.7593, 0.8048, and 0.7356 for holiday, weekday, and weekend scenarios, respectively, and outperformed Global XGBoost and multiscale geographically weighted regression under the same spatial cross-validation procedure. Within the specified greenness-focused model, NDVI generally showed larger SHAP magnitudes than GVI, particularly during holidays. NDVI contributions displayed relatively continuous spatial patterns, whereas GVI contributions were more fragmented and showed more frequent changes in direction across travel scenarios. Partial dependence curves further revealed nonlinear, scenario-dependent patterns, with greater variation during holidays than during weekdays or weekends. This study provides an interpretable spatial machine-learning framework for comparing complementary greenness indicators in metro-ridership prediction.

Keywords: urban green space; metro ridership; Local XGBoost Ensemble; NDVI; GVI

Check for updates

Academic Editors: Wolfgang Kainz, Agnese Pinto, Filippo Gramegna, Saverio Ieva, Floriano Scioscia and Michele Ruta

Received: 7 August 2026
Revised: 10 October 2026
Accepted: 13 October 2026
Published: 23 November 2026

Copyright: © 2026 by the authors. Published by MDPI on behalf of the International Society for Photogrammetry and Remote Sensing. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

1.1. Research Background and Gaps

With continued urbanization and motorization, large cities are facing growing sustainability challenges related to traffic congestion, energy consumption, and environmental pressures. As a high-capacity and low-carbon mode of public transport that supports compact urban development, metro systems have become an important component of urban transport systems and spatial development. Transit-Oriented Development (TOD) promotes high-density, mixed-use, and pedestrian-friendly environments around metro stations, whereas green urbanism places greater emphasis on environmental quality and residents' living experience. Integrating these two concepts may create synergies among public

transport use, compact development, and environmental improvement [1]. Understanding the relationship between the built environment around metro stations and station-level ridership is therefore important for land-use planning and TOD strategies in station areas. Previous studies have shown that station-level ridership is closely associated with a range of built-environment characteristics, including population and employment, commercial and public service facilities, land-use patterns, transport connectivity, and station location [2].

Compared with conventional built-environment factors such as population, employment, and land use, urban green space (UGS) has received relatively limited attention. Nevertheless, previous studies have shown that UGS is significantly associated with metro ridership. On the one hand, greener street environments can improve the walking experience of passengers accessing or leaving metro stations. On the other hand, green spaces can also serve as attractive destinations, particularly parks, waterfront areas, and large open green spaces that accommodate recreational and leisure activities [3,4]. The Normalized Difference Vegetation Index (NDVI) has traditionally been widely used to characterize vegetation coverage from a remote-sensing, top-down perspective. With the development of street-view imagery and semantic segmentation techniques, the Green View Index (GVI) has increasingly been used to measure greenery that is visually perceived by pedestrians at street level [5]. However, NDVI has limited ability to capture low-lying vegetation, vertical greenery, and fragmented green spaces, whereas GVI is constrained by road accessibility, street-view sampling, and street morphology. Moreover, the relationship between the two indicators exhibits considerable spatial heterogeneity [6].

Station-level ridership relationships may further vary across space and travel contexts. Conventional global regression and global machine-learning models generally impose a common functional relationship across stations. Recent spatial machine-learning studies have sought to accommodate spatially varying and nonlinear predictive relationships, and have reported improved predictive performance relative to some global benchmark models [7]. In addition, different travel scenarios are associated with different urban activity patterns. Pooling ridership observations from different periods may therefore obscure scenario-specific relationships between the built environment and metro ridership [8,9].

Against this background, three research gaps remain. First, few metro-ridership studies have jointly represented station-area greenness through both area-level vegetation coverage and street-level visible greenery. Second, evidence remains limited on whether the model-based predictive attribution and nonlinear association patterns of these two greenness measures differ across weekday, weekend, and holiday contexts. Third, existing global approaches may not adequately represent spatially heterogeneous and nonlinear modeled associations at the station level. Addressing these gaps requires a framework that compares NDVI and GVI under a common station sample, outcome definition, validation structure, and model specification.

1.2. Research Objectives and Contributions

This study uses Nanjing Metro as a case study to compare two complementary representations of station-area greenness: NDVI and GVI. Average daily inbound ridership is modeled separately for weekday, weekend, and holiday scenarios. A Local XGBoost Ensemble is used to model nonlinear predictive relationships while allowing model configurations and predictions to vary spatially through local estimation and distance-weighted ensemble prediction. Spatial Ensemble SHAP and partial dependence analysis are then used to examine relative predictive attribution, local model contributions, and nonlinear modeled response patterns across stations and travel scenarios.

The study makes three contributions. First, it compares overhead-view and eye-level greenness indicators within the same greenness-focused station-level modeling framework.

Second, it distinguishes weekday, weekend, and holiday contexts to examine whether the modeled patterns of NDVI and GVI differ across observed travel scenarios. Third, it combines a Local XGBoost Ensemble with Spatial Ensemble SHAP to describe spatially varying predictive attribution and nonlinear modeled associations.

1.3. Paper Organization

This report consists of five sections. Section 2 presents the literature review. Section 3 presents the study area, data sources, and the Local XGBoost Ensemble methodology. Section 4 reports and discusses the empirical results, including model performance, the relative importance and spatial heterogeneity of NDVI and GVI, and their cross-scenario and nonlinear associations with metro ridership. Section 5 summarizes the main findings, presents planning implications and limitations, and suggests directions for future research.

2. Literature Review

Many studies have examined metro ridership and urban greenness, as well as their spatiotemporal and nonlinear relationships, as summarized in Table 1. With the increasing availability of AFC, smart-card, and multi-source urban data, urban rail transit research has gradually shifted from aggregate passenger-flow description toward more detailed analyses of travel behavior and station-level demand [10,11]. Recent studies have also increasingly employed machine-learning approaches to identify more complex and nonlinear travel patterns [12,13].

Existing research shows that the relationship between the built environment and metro ridership is characterized by nonlinearity, spatial non-stationarity, and temporal variation [14,15]. These relationships can vary across temporal conditions and spatial scales, while fine-scale built-environment characteristics and network-based spatial structures may be associated with different ridership patterns [16,17]. A single global model may therefore fail to capture important differences across stations and time periods.

Ridership patterns also vary across travel conditions. Weekdays, weekends, and holidays correspond to different activity structures and travel demands, and metro-related access behavior and holiday passenger flows exhibit clear spatiotemporal differences [18,19]. Commuting, shopping, leisure activities, and station functions further are associated with differentiated ridership patterns [20,21]. Holiday subway travel also displays distinctive spatial correlations and flow structures [22], while station community structure, temporal scale, and last-mile conditions can further shape ridership outcomes [23,24].

Against this background, green space and recreational resources have gradually been introduced into research on urban travel and metro demand. Nonlinear and spatially heterogeneous relationships between the built environment and urban activity have been identified in studies of urban vitality and ride-hailing demand [25,26], while green space also provides broader environmental and recreational functions [27]. In metro-related studies, parks, tourist resources, and recreational facilities are commonly incorporated as components of the built environment or destination accessibility and have been shown to exhibit spatiotemporally varying relationships with public transport and metro ridership [28,29]. Green space within station catchments and at the origin–destination scale also shows spatially heterogeneous relationships with metro travel [6,30].

Research has increasingly shifted toward human-scale greenery. Green-space exposure around metro stations is associated with mode choice, travel satisfaction, and cycling behavior [31,32], suggesting that perceived street-level greenery may be relevant to station access conditions. More recent work has directly incorporated GVI into metro-ridership analysis and identified nonlinear and spatially heterogeneous relationships [33], but systematic comparison across different greenness perspectives remains limited.

This limitation is closely related to the measurement of urban greenness. NDVI mainly represents vegetation coverage from an overhead perspective, whereas GVI captures visible greenery at street level, and the two measures are not equivalent. Street-view measures can capture human-scale green exposure that conventional remote-sensing indicators may not fully represent [34], while NDVI and GVI, although correlated, still show substantial spatial inconsistency [35]. Their differences are also shaped by urban form [36], and both satellite- and street-view-based measures may be affected by observation perspective, road accessibility, image timing, and vegetation occlusion [37]. NDVI and GVI should therefore be regarded as complementary rather than interchangeable measures of urban greenness.

Overall, existing research has established the importance of temporal heterogeneity, spatial non-stationarity, and nonlinearity in metro-ridership analysis, while green space and street-level greenery have increasingly been introduced into metro-related travel studies. However, greenery is still commonly treated as one explanatory variable within a broader built-environment framework, and comparisons between NDVI and GVI have mainly focused on greenness measurement itself. Systematic evidence remains limited regarding their relative predictive contributions, nonlinear response patterns, and spatial heterogeneity across different travel conditions. The present study therefore focuses on these two greenness representations and compares their differential modeled associations under weekday, weekend, and holiday conditions.

Table 1. Comparison of representative studies on urban green space and travel demand.

Topic	Study	Area	Methods	Key Findings and Gaps
Spatiotemporal heterogeneity of metro ridership	Gan et al., 2020 [14]	Nanjing	Gradient Boosting Regression Trees (GBRTs)	Revealed nonlinear built-environment effects on metro ridership; greenness was not the focus.
	Chen et al., 2019 [15]	Nanjing	Modified-Distance Geographically Weighted Regression (MD-GWR)	Identified spatiotemporal heterogeneity in built-environment effects; did not examine multi-perspective greenness.
	Li et al., 2020 [38]	Guangzhou	Big-data analysis of fine-scale built environment and ridership patterns	Showed temporal variation in ridership–environment relationships; NDVI and GVI were not distinguished.
	Zhou et al., 2025 [22]	Hangzhou	Mobile-payment data and spatial correlation analysis	Revealed distinctive holiday subway-flow patterns; urban greenness was not considered.
	Gao et al., 2022 [17]	/	Network-distance-based geographically weighted regression	Captured spatially varying built-environment effects; nonlinear NDVI–GVI differences were not examined.

Table 1. *Cont.*

Topic	Study	Area	Methods	Key Findings and Gaps
Urban greenness and metro-related travel behavior	An et al., 2022 [28]	Wuhan	Multiscale Geographically Weighted Regression (MGWR)	Identified multiscale built-environment effects on public transport use; greenness was not the focal factor.
	Du et al., 2022 [29]	/	Nonlinear spatiotemporal accessibility analysis	Found nonlinear effects of recreational accessibility; greenery was measured indirectly.
	Xi et al., 2024 [30]	Xian	Spatiotemporal nonlinear analysis	Identified heterogeneous green-space effects on ridership; NDVI and GVI were not separately compared.
	Wu et al., 2023 [31]	Beijing	Empirical travel-behavior analysis	Linked station-area greenery with mode choice and satisfaction; did not directly model metro ridership.
	Wang et al., 2020 [32]	Shenzhen	Poisson regression	Linked eye-level greenness with cycling frequency; focused on cycling and a single greenness perspective.
	Dong et al., 2026 [33]	Shanghai	Global and local interpretable nonlinear modeling	Revealed nonlinear and spatially varying GVI–ridership relationships; no systematic NDVI–GVI comparison.
Multi-perspective greenness measurement	Larkin & Hystad, 2019 [34]	/	Comparative evaluation of street-view greenness measures	Highlighted the value of street-view greenery; did not examine transport demand.
	Li et al., 2021 [35]	Beijing	Spatial correlation and geographically weighted regression	Found spatial inconsistency between GVI and NDVI; did not link them to metro ridership.
	Hua et al., 2022 [36]	Hongkong	Deep-learning-based street-view extraction and urban-form comparison	Showed that NDVI–GVI differences vary with urban form; travel demand was not examined.
	Huang et al., 2025 [37]	10 cities worldwide	Cross-city comparative measurement-bias analysis and machine learning	Identified measurement biases between satellite and street-view greenery; did not examine metro ridership.

3. Data and Methods

3.1. Study Area

Nanjing is the provincial capital of Jiangsu Province and one of the important central cities in eastern China [39]. By 2025, it had 12 rail transit lines and 234 stations, with a total operating mileage of 484 km. Stations along suburban sections generally exhibit lower ridership than central-area stations, reflecting differences in population density and the spatial concentration of urban activities between central and suburban areas.

A catchment area was defined for each metro station to delimit the spatial extent within which the surrounding built environment could reasonably influence station ridership. The selection of the catchment radius should balance behavioral relevance, spatial coverage, and the risk of including areas that are less directly connected to station access. Comparative research has shown that there is no universally optimal catchment radius and that the appropriate scale may vary across cities and station contexts [40]. Zhu et al. used AFC and POI data within an 800 m radius to investigate the relationship between the built environment and metro station ridership in Nanjing [39]. Therefore, we selected an 800 m buffer zone as the influencing area for the rail transit station. In view of the close proximity of rail transit stations in the city center, the Thiessen polygon method was employed in order to address the issue of overlapping catchment areas.

Figure 1 illustrates the study area at three spatial scales. The upper-left panel locates Nanjing within China, while the lower-left panel further shows the position of Nanjing within Jiangsu Province. The red rectangular frame in the lower-left panel identifies the spatial extent enlarged in the right-hand panel. In the right-hand panel, the red shaded areas delineate the station catchment areas defined using the 800 m buffers and Thiessen polygons. The panel also presents metro stations, metro lines, the road network, water bodies, urban green spaces, and the administrative district names within the study area.

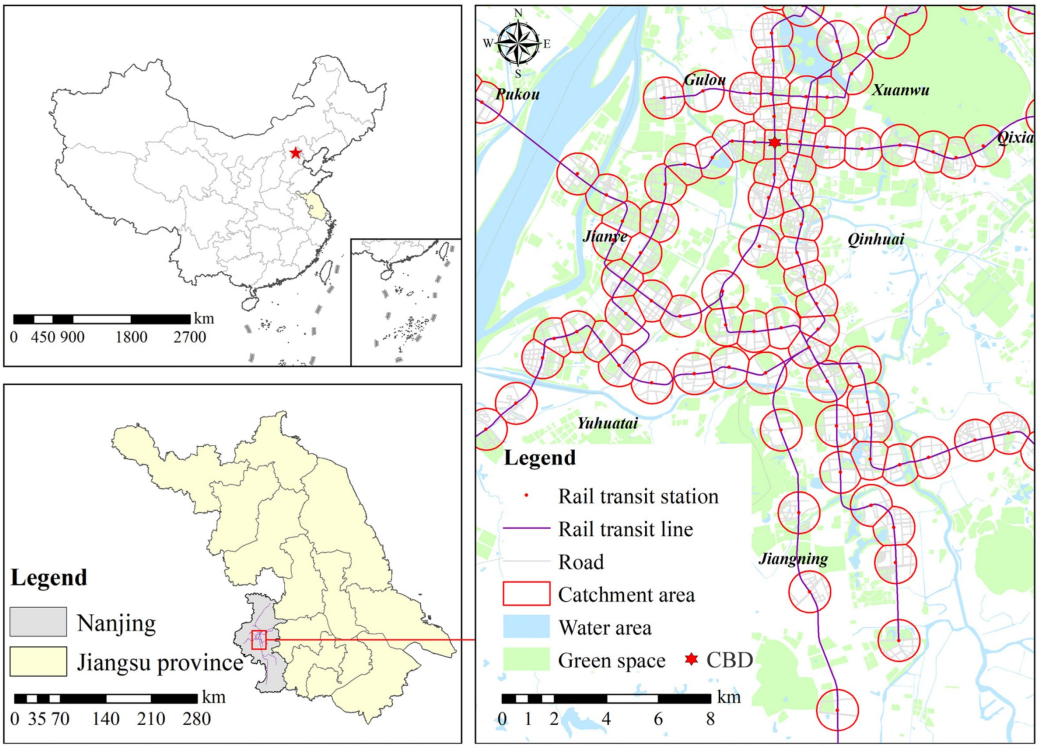

Figure 1. Map of the study area in Nanjing.

3.2. Data

3.2.1. Ridership Data

The ridership data for rail transit stations originated from the automatic fare collection (AFC) data provided by Nanjing Metro Group Co., Ltd. (Nanjing, China). The source tables report inbound and outbound passenger counts for each station by observation date and time interval. During data preparation, station names were standardized across observation dates, line-specific records referring to the same physical interchange station were consolidated, and the time-interval counts were aggregated to obtain daily station-level ridership. After excluding inactive station entries without valid passenger-flow observations, the AFC dataset contained 190 operating stations and 5700 station-day observations.

This data covers weekdays (April 26–28, May 4–6, 8–10, September 22 and 25–28), weekends (May 7 and September 23–24), May Day Holiday (April 29–30 and May 1–3), and National Day Holiday (September 29–30 and October 1–6). For each station, average daily inbound ridership was calculated separately for weekday, weekend, and holiday scenarios as the arithmetic mean of daily inbound ridership across all observed dates in the relevant category. The observations do not cover a full annual cycle, and the holiday data include only the relatively long May Day and National Day holidays. Therefore, the findings may not represent annual ridership variations or shorter public holidays. Table 2 presents the descriptive statistics of the average daily inbound passenger flow at selected stations.

Table 2. Daily average inbound passenger flow statistics.

Station Name	Longitude	Latitude	Average Daily Passenger Flow		
			Holiday	Weekday	Weekend
Xinjiekou	118.778935	32.04386	146,371	82,397	96,179
Jiangwang Temple	118.826032	32.079676	8126	6371	4626
Nanjing Station	118.791322	32.091403	94,102	47,021	40,460
Huashen Temple	118.780455	31.979836	3878	7016	2880
Andemen	118.756919	31.992886	8626	11,938	7289
Zhongsheng	118.727958	31.989938	5288	19,550	5129
Jiming Temple	118.792347	32.059734	37,427	13,731	13,001
Linshan	118.525967	31.920925	1159	915	838
Yuantong	118.716248	31.997509	14,518	20,309	10,335
Gulou	118.778497	32.060724	27,587	35,660	22,022
…	…	…	…	…	…

To assess the stability of the sampled observations, we further examine the within-category variation in daily ridership and the similarity of station-level ridership profiles across observation days. Table 3 reports the within-category variability in systemwide daily inbound ridership and the temporal stability of station-level ridership profiles. The coefficients of variation in systemwide daily inbound ridership are 14.94%, 10.01%, and 6.84% for holidays, weekdays, and weekends, respectively. Meanwhile, the corresponding median pairwise Spearman correlations between the station ridership profiles are 0.9770, 0.9941, and 0.9916, indicating that the relative spatial distribution of ridership across stations remains stable within the observed travel contexts despite day-to-day changes in overall ridership volume.

Table 3. Within-category variability and spatial-profile stability of daily inbound ridership.

Travel Category	CV (%)	Median Pairwise Spearman's ρ	Range
Holiday	14.94%	0.9770	0.9392–0.9969
Weekday	10.01%	0.9941	0.9602–0.9989
Weekend	6.84%	0.9916	0.9872–0.9932
May Day	20.22%	0.9664	0.9392–0.9938
National Day	6.36%	0.9862	0.9580–0.9969

Table 4 evaluates the sensitivity of this choice by comparing inbound ridership with outbound and total ridership. Inbound and outbound ridership exhibit consistently high station-level correlations across the three scenarios, with Pearson correlations ranging from 0.9938 to 0.9982 and Spearman correlations ranging from 0.9958 to 0.9972. Their median normalized differences range from 4.33% to 5.54%, indicating that the typical numerical difference between the two directions is also limited. Accordingly, average daily inbound ridership is retained as the dependent variable.

Table 4. Sensitivity analysis of inbound, outbound, and total station ridership.

Scenario	I–O Pearson's r	I–O Spearman's ρ	I–T Pearson's r	I–T Spearman's ρ	Median Normalized I–O Difference (%)	IQR (%)	Median O/I Ratio
Holiday	0.9938	0.9958	0.9985	0.9989	5.54	2.54–11.24	0.966
Weekday	0.9982	0.9964	0.9996	0.9991	5.26	2.67–9.25	0.964
Weekend	0.9971	0.9972	0.9993	0.9992	4.33	2.43–9.15	0.970

Note: I, O, and T denote average daily inbound, outbound, and total ridership.

3.2.2. Multi-Perspective Urban Green Space Indicators

(1) NDVI

The Normalized Difference Vegetation Index (NDVI) is a remote-sensing index derived from the spectral reflectance characteristics of vegetation; it can characterize the extent of surface vegetation cover and its growth status from a bird's-eye view. Using the Google Earth Engine (GEE) cloud computing platform, Level-2A surface reflectance data from the Sentinel-2 Multispectral Instrument (MSI) were acquired and processed; the dataset used was COPERNICUS/S2_SR_HARMONIZED. Using the administrative boundaries of Nanjing as the image selection and cropping area, Sentinel-2 imagery from May 2023 was acquired to characterize the surface vegetation conditions in Nanjing.

Scene-level filtering was performed using the CLOUDY_PIXEL_PERCENTAGE attribute of Sentinel-2 imagery to retain images with cloud cover not exceeding 80%. The Sentinel-2 Level-2A surface reflectance imagery was spatially and temporally aligned with the Google Cloud Score+ dataset, and the cs_cdf band was selected for pixel-by-pixel quality assessment. Pixels with a cs_cdf value of 0.60 or higher were classified as relatively clear pixels, and only those meeting this threshold were retained for subsequent NDVI calculations. To remove invalid values and anomalous pixels that may exist at the edges of Sentinel-2 image tiles, edge quality control was further performed using valid data masks from the B8A and B9 bands.

Calculate the NDVI for each scene using Sentinel-2's near-infrared band B8 and red band B4:

$$NDVI = \frac{\rho_{B8} - \rho_{B4}}{\rho_{B8} + \rho_{B4}} \quad (1)$$

In the equation, ρ_{B8} represents the surface albedo in the Sentinel-2 near-infrared band B8, and ρ_{B4} represents the surface albedo in the red band B4.

To further mitigate the impact of residual clouds, cloud shadows, atmospheric disturbances, and pixels with anomalous high or low values on the monthly NDVI results, this study performed a pixel-by-pixel median merging of all quality-controlled, scene-by-scene NDVI images within the same month. For a pixel at spatial location s, its NDVI median for month m can be expressed as:

$$NDVI_m(s) = \text{median}\{NDVI_1(s), NDVI_2(s), \ldots, NDVI_{T_m}(s)\} \quad (2)$$

In the equation, $NDVI_m(s)$ represents the composite NDVI value at spatial location s in month m; $NDVI_t(s)$ represents the NDVI value at location s in the t valid image for that month; T_m represents the number of valid observations available for that pixel within month m.

In this study, the mean NDVI within each station catchment represents area-level vegetation coverage and condition from a top-down perspective. It captures the broader ecological context surrounding a station, including continuous vegetation, parks, and other green open spaces that may be associated with the spatial attractiveness and land-use characteristics of the station area. In the subsequent model interpretation, the SHAP values and partial dependence patterns of NDVI are used to examine its relative predictive attribution, spatial heterogeneity, and nonlinear association with ridership across travel scenarios.

(2) GVI

The Green View Index (GVI) is a people-centered green space indicator based on street-view imagery, used to characterize the proportion of vegetation that pedestrians can directly observe in street spaces. Street-view image data provided by the Baidu Maps Open Platform was used to measure the GVI around Nanjing Metro stations. The road network was extracted within the station area, and street-view sampling points were set at fixed intervals of 50 m along the roads. Street-view images were captured from four different directions at each sampling point to maximize coverage of the horizontal field of view surrounding the sampling location.

The original street-view images were cropped and normalized to satisfy the input requirements of the semantic segmentation model. Pixel-level classification of the environmental elements was performed using DeepLabV3+ with publicly available pretrained weights from the Cityscapes dataset. DeepLabV3+ with the Xception backbone has been reported to achieve a mean intersection over union (mIoU) exceeding 82% on the Cityscapes benchmark dataset [41]. Among the 19 semantic classes produced by the model, only pixels classified as "vegetation" were categorized as greenery. The number of vegetation pixels was then divided by the total number of valid pixels to calculate the directional GVI of each street-view image. The Baidu street-view images were retrieved in 2023; however, because image-level capture dates were not consistently available in the retained metadata, the retrieval year does not necessarily indicate that all images were originally captured in 2023. For a street-view image at sampling point p in direction k, its directional GVI is expressed as follows:

$$GVI_{p,k} = \frac{S_{p,k}^{green}}{S_{p,k}^{total}} \quad (3)$$

In the equation, $GVI_{p,k}$ represents the green coverage ratio of the sampling point p in the street-view image along the direction k; $S_{p,k}^{green}$ represents the area or number of pixels identified as vegetation in that image; $S_{p,k}^{total}$ represents the total area or number of valid pixels in that image included in the calculation.

The GVI values from the four directions at the same sampling point are averaged to obtain the comprehensive green view ratio for that sampling point:

$$GVI_p = \frac{1}{4}\sum_{k=1}^{4} GVI_{p,k} \tag{4}$$

Calculate the average of the GVI values for all valid sampling points within the buffer zone of the subway station i to obtain the station-scale GVI:

$$GVI_{i,r} = \frac{1}{M_{i,r}}\sum_{p=1}^{M_{i,r}} GVI_p \tag{5}$$

In the equation, $M_{i,r}$ represents the number of valid street-view sample points within the buffer zone of the subway station i. The resulting metric is used to characterize the average visible green area in the road environment surrounding subway stations at different spatial scales.

The station-scale GVI represents street-level visible greenery along the road network within each station catchment. Unlike NDVI, which captures area-level vegetation coverage from an overhead perspective, GVI reflects vegetation that passengers may directly perceive along station access and egress routes and is therefore more closely related to the visual quality of the pedestrian environment. Therefore, this study interprets GVI as a street-level visual greening index, which, together with NDVI, forms a multiscale urban green space index to comprehensively characterize the green space environment around subway stations from different observational perspectives.

3.3. Method

XGBoost is adopted as the base learner because it is well suited to structured station-level data and can capture nonlinear associations and interactions between greenness indicators without requiring a predefined functional form. Its shrinkage, row subsampling, and L1/L2 regularization mechanisms provide explicit control over model complexity, which is particularly important when models are repeatedly trained using relatively small local neighborhoods. XGBoost also offers sufficient computational efficiency for repeated local-model estimation and hyperparameter optimization and is compatible with TreeSHAP, allowing the prediction and interpretation stages to be implemented within a consistent framework.

To simultaneously capture nonlinear relationships and spatial heterogeneity, we constructed Local XGBoost Ensemble. We trained a set of local XGBoost models centered on subway stations, and during the prediction phase, we performed weighted ensemble analysis of the prediction results from multiple neighboring local models based on the spatial distance between the target station and the center of each local model. Compared to using a single local model for prediction, the spatially weighted fusion of neighboring models reduces the instability of local models with small sample sizes and mitigates the spatial discontinuities that may result from switching between local models.

To ensure a consistent comparison, NDVI and GVI were calculated for the same metro stations using the same 800 m station catchment areas and were matched with the same scenario-specific ridership outcomes. Within each travel scenario, both indicators were included in the same Local XGBoost Ensemble and evaluated using identical spatial cross-validation partitions, neighborhood construction, ensemble-weighting, and parameter-optimization procedures.

NDVI and GVI were the explanatory predictors entered into each XGBoost learner. Station coordinates were used to construct local training neighborhoods and to select and weight neighboring local models during prediction, making location a structural input to the complete framework. Other built-environment characteristics, including population density,

employment, land use, station function, network connectivity, and service characteristics, were not included as covariates. The reported SHAP values therefore describe the predictive attributions of NDVI and GVI within this spatially structured framework and do not estimate their independent causal effects.

To evaluate spatial generalization, stations were partitioned into five spatial groups using K-means clustering based on standardized projected station coordinates. Five-fold GroupKFold validation was then applied, with one spatial group held out in each fold. A random seed of 2026 was used for spatial partitioning, Optuna sampling, and XGBoost estimation. XGBoost hyperparameters were optimized separately for each travel scenario using the TPE sampler in Optuna, with up to 50 trials. The search covered n_estimators (100–900), max_depth (2–6), learning_rate (0.005–0.12), subsample (0.55–1.00), colsample_bytree (0.55–1.00), min_child_weight (1–15), gamma (0–1), reg_alpha (1×10^{-4}–10), and reg_lambda (0.5–25). The Global XGBoost benchmark used the same optimization framework, with the reg:squarederror objective, histogram tree method, and RMSE evaluation metric.

For the MGWR benchmark, predictor variables were standardized using z-score normalization. Separate MGWR models were estimated for the three travel scenarios using covariate-specific adaptive bandwidths (multi = True) and an adaptive bisquare kernel. Bandwidths were selected by minimizing AICc through golden-section search. No manual parameter grid search or stochastic optimization was applied to MGWR.

Step 1: Data Preparation

Using Nanjing Metro stations as the basic spatial analysis units, we established a station-scale database containing passenger flow, urban green spaces, and station spatial coordinates. For the station i, its vector of explanatory variables can be expressed as:

$$x_i = (x_{i1}, x_{i2}, \ldots, x_{ip}) \tag{6}$$

where x_i denotes the vector of explanatory predictors for station i.

The dependent variables are passenger ridership at stations on weekdays, weekends, and holidays, respectively. All types of spatial data are aggregated into a unified station analysis unit and matched with station IDs and spatial coordinates.

$$y_i^{(s)}, s \in \{\text{weekday}, \text{weekend}, \text{holiday}\} \tag{7}$$

In the equation, s represents the passenger flow scenario.

Step 2: Build a local XGBoost model centered on the station

For a local learner centered at station i, only stations available in the training set of the current cross-validation fold are considered as candidates. Let N_{train} denote the number of eligible training stations in that fold. K_{local} denotes the proportion of these stations included in the local training neighborhood, and the corresponding number of selected stations is defined as $n_{\text{local}} = [K_{\text{local}} \times N_{\text{train}}]$. The n_{local} nearest eligible training stations to center i are then used to train the local learner. Validation stations are excluded from this candidate pool.

$$N_i^{\text{local}} = \text{KNN}(i, n_{\text{local}}) \tag{8}$$

The corresponding local training dataset is:

$$S_i = \left\{(x_j, y_j) \mid j \in N_i^{\text{local}}\right\} \tag{9}$$

For the central station i, train a separate XGBoost regression model $f_i()$ on a local dataset S. XGBoost gradually adds regression trees to fit the residuals from the previous round's model; its prediction function can be expressed as:

$$\hat{y}_j^{(i)} = f_i(x_j) = \sum_{t=1}^{T} g_t^{(i)}(x_j) \tag{10}$$

In the equation, $\hat{y}_j^{(i)}$ represents the predicted value for station j by the local model centered on station i; T denotes the number of trees in the model; $g_t^{(i)}$ represents the t regression tree in the local model i.

The training objective for a local model consists of a loss function and a model complexity penalty term:

$$J_i = \sum_{j \in N_i} l\left(y_j, \hat{y}_j^{(i)}\right) + \sum_{t=1}^{T_i} \Omega\left(g_{i,t}\right) \tag{11}$$

$l()$ represents the prediction loss function; in continuous passenger flow forecasting, the mean squared error is typically used; $\Omega()$ represents the tree complexity regularization term, which can be written as:

$$\Omega(g) = \gamma Q + \frac{\lambda}{2} \sum_{q=1}^{Q} w_q^2 + \alpha \sum_{q=1}^{Q} |w_q| \tag{12}$$

Q denotes the number of leaf nodes in the regression tree g, w denotes the weight of the q leaf node, γ denotes the leaf-node-complexity penalty coefficient, and λ and α denote the L2 and L1 regularization coefficients, respectively.

This training process is repeated for all stations, resulting in a model ensemble consisting of N local XGBoost models:

$$M = \{f_1, f_2, \ldots, f_N\} \tag{13}$$

Step 3: Distance-Weighted Ensemble Prediction Using Nearby Local Models

For a target station i, first select the K_{ens} local models closest to the target station from among all local model centers; their set is denoted as:

$$E_i = \mathrm{KNN}(i, K_{\mathrm{ens}}) \tag{14}$$

The framework involves two levels of locality. K_{local} controls the training scale of each local learner; training neighborhoods centered at nearby stations may overlap, which improves model stability under the limited station sample. K_{ens} controls prediction-scale locality. Only nearby local models are selected for each target station and combined using normalized distance-based weights. Thus, the final prediction remains location-specific through station-specific model selection and weighting.

For any local model m in the set, the explanatory variable x_i for the target station is input into model $f_m()$ to obtain that model's prediction for the target station:

$$\hat{y}_i^{(m)} = f_m(x_i) \tag{15}$$

Because spatially adjacent stations may exhibit more similar observed ridership patterns and station-area characteristics, an inverse-square distance function was used to determine the weights assigned to neighboring local models. After normalizing the weights, the following is obtained:

$$w_{im} = \frac{1/(d_{im}^2 + \varepsilon)}{\sum_{l \in E_i} 1/(d_{il}^2 + \varepsilon)} \tag{16}$$

d_{im} represents the spatial distance between the two; ε is a very small positive number used to prevent the denominator from becoming zero when the target station coincides with the center of the local model.

The final prediction for the target station i is obtained by weighting the predictions from all selected local models:

$$\hat{y}_i = \sum_{m \in E_i} w_{im} \hat{y}_i^{(m)} \tag{17}$$

To ensure consistency between ensemble prediction and model interpretation, SHAP values were first calculated separately for the neighboring local models selected for each target station. The resulting contributions were then combined using the same normalized distance-based weights employed for ensemble prediction:

$$\Phi_{if} = \sum_{m \in E_i} w_{im} \phi_{if}^{(m)} \tag{18}$$

Φ_{if} denotes the Spatial Ensemble SHAP value of feature f for target station i, $\phi_{if}^{(m)}$ denotes the SHAP contribution generated by local model m, $\mathcal{N}i$ is the set of neighboring local models selected for station i, and $w\{im\}$ is the corresponding normalized distance-based weight. The baseline values of the local models were aggregated using the same weights, thereby preserving the additive relationship between the ensemble prediction, the weighted baseline, and the Spatial Ensemble SHAP contributions.

For each station-centered local learner, feature importance was calculated as the mean absolute SHAP value across the samples in its training neighborhood, separately for NDVI and GVI. Taking absolute values prevents positive and negative contributions from canceling, while averaging summarizes the contribution magnitude within each neighborhood.

For each grid value z of feature f, the focal feature was replaced by z for all stations, while the remaining predictor values and station coordinates were retained. Predictions were generated using the same neighboring local models and distance-based ensemble weights, and the partial dependence value was calculated as:

$$PD_f(z) = \frac{1}{n} \sum_{i=1}^{n} \hat{y}_i\left(z, x_{i,-f}\right) \tag{19}$$

where $x_{i,-f}$ denotes the remaining predictor values of station i, and $\hat{y}i(z, x_i, -f)$ denotes the prediction generated by the complete Local XGBoost Ensemble after replacing feature f with z. The values obtained across the feature grid form the scenario-specific PDP. Spatial Ensemble SHAP values and PDPs were interpreted as predictive attribution and modeled response patterns within the fitted model.

Step 4: Hyperparameter Optimization for Model Groups Using Optuna

Parameters to be optimized include:

$$\Theta = \{T, \eta, d_{\max}, s_{\text{row}}, \alpha, \lambda, K_{\text{local}}, K_{\text{ens}}\} \tag{20}$$

T represents the number of trees, η represents the learning rate, $d_{\max}$ represents the maximum tree depth, s_{row} represents the sample sampling ratio, α and λ represent the L1 and L2 regularization parameters, respectively, K_{local} represents the proportion of stations included in the local training neighborhood, and K_{ens} represents the number of local models participating in the ensemble during the prediction phase.

The optimization objective for the model is defined as:

$$\Theta^* = \arg\min_{\Theta} O(\Theta) \tag{21}$$

$O(\Theta)$ represents the model's prediction error in spatial cross-validation. The root mean square error (RMSE) is used as the optimization objective, and its formula is as follows:

$$O(\Theta) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i(\Theta))^2} \tag{22}$$

n represents the number of samples in the validation set; y_i and $\hat{y}_i$ denote actual ridership and predicted ridership, respectively.

Step 5: Spatial Batch Cross-Validation and Model Evaluation

Subway station data exhibits distinct spatial dependence. Spatially adjacent stations typically share similar built environments, land-use patterns, and passenger flow characteristics. If standard random K-fold cross-validation is used, adjacent stations may appear simultaneously in both the training and validation sets, leading to spatial information leakage. Therefore, spatial block cross-validation is employed. First, K-means spatial clustering is performed based on the stations' projected coordinates to group stations with similar geographic locations into the same spatial cluster:

$$G_i = \mathrm{KMeans}(u_i, v_i) \tag{23}$$

G_i denotes the spatial group to which station i belongs. The spatial group is used as the grouping variable for GroupKFold. In each round of cross-validation, one or more complete spatial groups are selected as the validation set, and the remaining spatial groups serve as the training set.

At each cross-validation split, the local model is built using the stations in that training set. Accordingly, the spatial cross-validation procedure evaluates geographically separated generalization rather than random station-level interpolation. Holding out an entire spatial group may remove some stations that are geographically close to the validation stations from the training pool. The local neighborhoods used during validation are therefore restricted to the stations available outside the held-out group and may differ from those available when the model is calibrated using the complete station network.

$$N_{i,f}^{\mathrm{local}} = \mathrm{KNN}\left(i, n_{\mathrm{local}}; S_{\mathrm{train}}^{(f)}\right) \tag{24}$$

f denotes the current cross-validation fold. The model is evaluated using the coefficient of determination, mean absolute error, and root mean square error:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2} \tag{25}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{26}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{27}$$

As illustrated in Figure 2, the research framework mainly consists of two components: multi-source data collection and processing, and modeling methods and visualization of results. AFC records were aggregated at the station level to calculate average daily inbound ridership under weekday, weekend, and holiday scenarios. Sentinel-2 remote-sensing imagery and Baidu street-view imagery were processed to derive NDVI and GVI, representing area-level vegetation coverage and pedestrian-level visible greenery, respectively. These multi-source data were subsequently matched according to station identifiers and spatial

coordinates to establish a unified station-scale database. Separate Local XGBoost Ensemble models were developed for the three travel scenarios. Spatial block cross-validation and Optuna-based optimization were employed to evaluate model performance and determine the optimal model and spatial-ensemble parameters. Finally, Spatial Ensemble SHAP and partial dependence analysis were used to visualize and compare the relative predictive attribution, spatial patterns, and nonlinear model responses associated with NDVI and GVI across stations and travel scenarios.

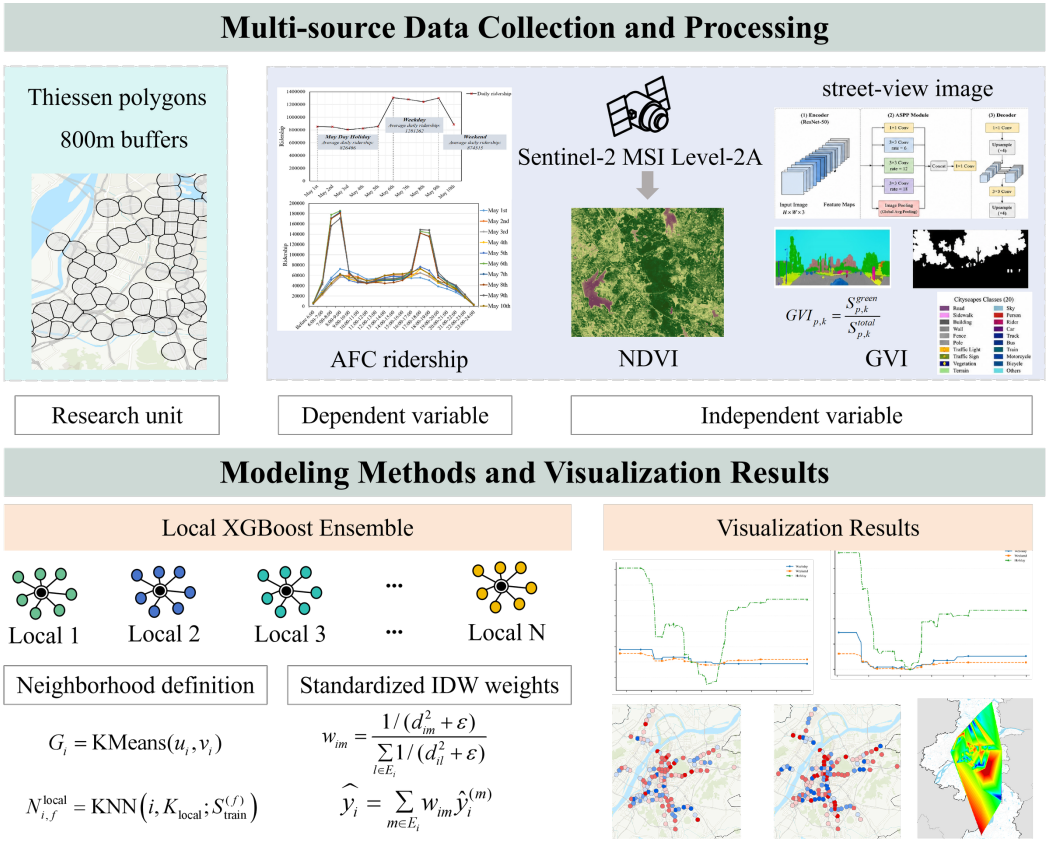

Figure 2. Research framework.

4. Results and Discussion

4.1. Model Tuning

To obtain the optimal model configuration for different travel scenarios, this study employs an Optuna-based Bayesian optimization method to jointly optimize the spatial structure parameters of the Local XGBoost Ensemble and the XGBoost hyperparameters. The optimal parameter combinations for each travel scenario and the corresponding model prediction performance are shown in Table 5. As shown in Table 5, the optimal parameter combinations differed across the three travel scenarios. The weekday model used a local neighborhood ratio of 0.6 and five local models for ensemble prediction and achieved the highest R^2 among the three scenario-specific models. The holiday and weekend models selected local neighborhood ratios of 0.5 and 0.8, respectively, with three and five local models included in their ensemble predictions. These results demonstrate that scenario-specific parameter optimization produced distinct model configurations for holidays, weekdays, and weekends.

Table 5. Optimal model parameters and prediction performance under different travel scenarios.

Scenario	Learning Rate	Max Depth	n_estimators	Subsample	K_{local}	K_{ens}	R^2	MAE	RMSE
Holiday	0.0484	2	700	0.55	0.5	3	0.7593	2102.77	4112.95
Weekday	0.0551	4	450	0.70	0.6	5	0.8048	1540.43	4411.83
Weekend	0.0366	6	850	0.65	0.8	5	0.7356	2330.57	3461.98

Table 6 provides additional prediction-error diagnostics for the three travel scenarios. The WAPE values were 9.18%, 11.94%, and 13.36%, respectively, indicating that the prediction errors remained relatively limited after accounting for differences in ridership magnitude. Taken together with the R^2 results, these findings show that the models adequately captured the overall cross-station variation in ridership under all three travel scenarios. Therefore, the overall model performance provides a reasonable basis for the subsequent comparative analysis of the spatial patterns and nonlinear associations of NDVI and GVI with station ridership.

Table 6. Prediction-error diagnostics under different travel scenarios.

Scenario	MAE 95% CI	Median AE	WAPE	WAPE 95% CI
Holiday	[1774.76, 2646.53]	1683	9.18%	[6.23%,11.45%]
Weekday	[1302.17, 1908.85]	1213	11.94%	[10.31%,13.76%]
Weekend	[2080.88, 2547.30]	1594	13.36%	[11.71%,15.16%]

Figure 3 presents prediction errors stratified by observed-ridership quartile for each travel scenario, with Q1 representing the lowest-ridership stations and Q4 the highest. Absolute errors were greatest in Q4, whereas WAPE was highest in Q1 and decreased substantially across the higher-ridership quartiles. Thus, although low-ridership stations had smaller absolute errors than high-ridership stations, their errors were larger relative to observed ridership. These results complement the overall performance metrics by showing that predictive accuracy varies with station ridership scale.

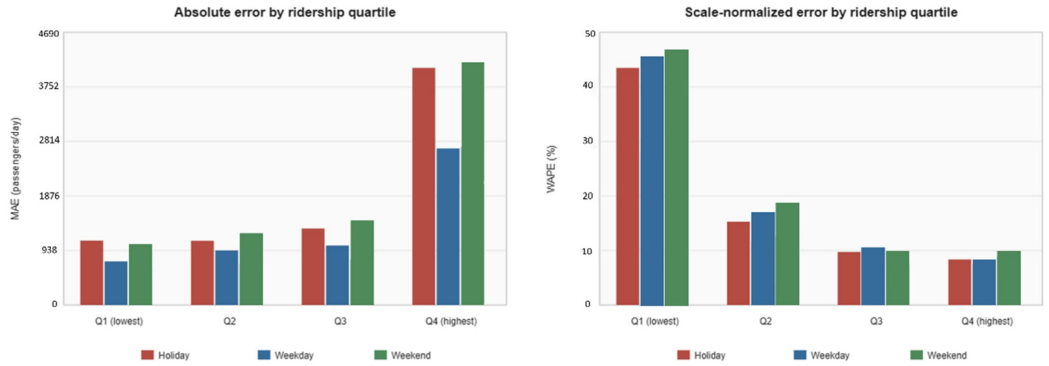

Figure 3. Predictive performance across observed-ridership quartiles.

As shown in Table 7, the Local XGBoost Ensemble consistently outperformed both Global XGBoost and MGWR across all three travel scenarios under the same spatial cross-validation procedure. The Local XGBoost Ensemble also achieved higher R^2 and lower absolute prediction errors than MGWR in all three scenarios.

Table 7. Comparison of model performance under different travel scenarios.

Scenario	Model	R^2	MAE	RMSE
Holiday	MGWR	0.6926	7513.13	9413.16
	Global XGBoost	0.7358	4135.79	5822.53
	Local XGBoost	0.7593	2102.77	4112.95
Weekday	MGWR	0.7439	5466.56	8027.69
	Global XGBoost	0.7672	4861.06	6570.97
	Local XGBoost	0.8048	1540.43	4411.83
Weekend	MGWR	0.6644	8495.62	10,671.73
	Global XGBoost	0.7175	4336.44	5092.45
	Local XGBoost	0.7356	2330.57	3461.98

4.2. Relative Importance of Multi-Perspective Urban Green Space

Figure 4 shows the distribution of local feature importance for NDVI and GVI in the Local XGBoost Ensemble model under different travel scenarios. Compared with weekdays, the absolute SHAP values of both greenness indicators are generally higher during holidays and weekends, indicating greater model-based predictive attribution within the specified model under these observed travel scenarios.

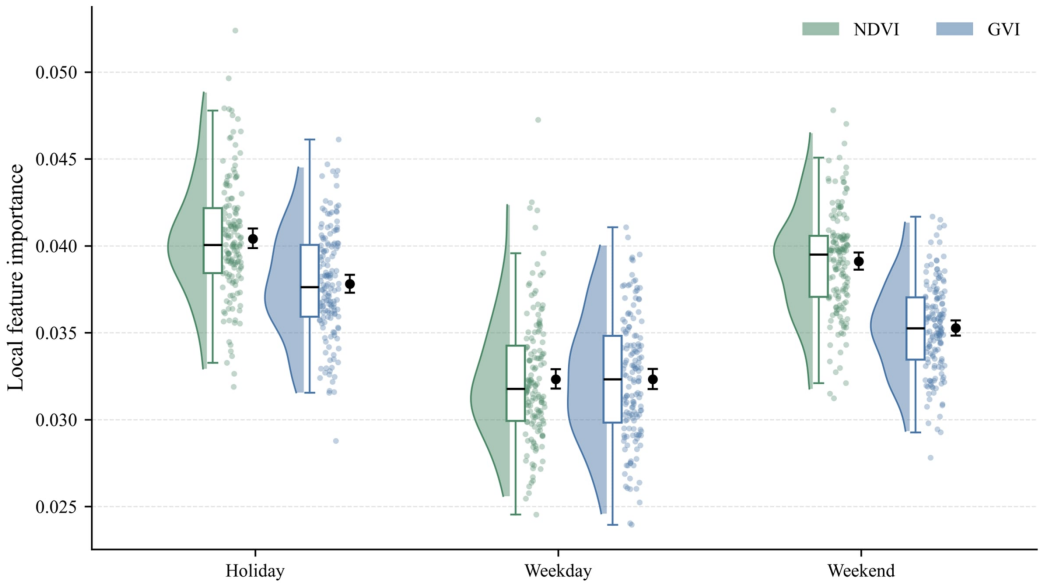

Figure 4. Local feature importance for NDVI and GVI under different travel scenarios.

NDVI showed greater local feature importance than GVI during holidays and weekends, whereas their importance was broadly comparable on weekdays. NDVI importance was highest on holidays, with a median of approximately 0.040, while on weekdays it declined to about 0.032, a reduction of approximately 20% relative to holidays. The holiday and weekend distributions shifted toward higher values and showed longer upper tails, with local importance exceeding 0.05 at some stations. By contrast, weekday values were concentrated mainly between 0.030 and 0.034, with less spatial variation. These patterns indicate that NDVI received greater model-based predictive attribution in the fitted holiday and weekend models than in the weekday model.

For both indicators, local importance was highest on holidays, followed by weekends, and lowest on weekdays. The difference between NDVI and GVI was most pronounced

during holidays and narrowed on weekdays. Across the three observed travel scenarios, NDVI generally exhibited larger local SHAP magnitudes than GVI, particularly during holidays. Within the specified feature set, this pattern indicates that overhead-view vegetation coverage contributed more strongly to model predictions of station ridership than street-level visible greenery.

Figure 5 further illustrates the station-level distribution of local feature importance for NDVI and GVI. Both indicators exhibit substantial spatial heterogeneity, with considerable variation among stations. The spatial patterns of local importance also differ across travel scenarios, indicating that the fitted models relied on the two greenness indicators differently across station contexts and observed travel scenarios.

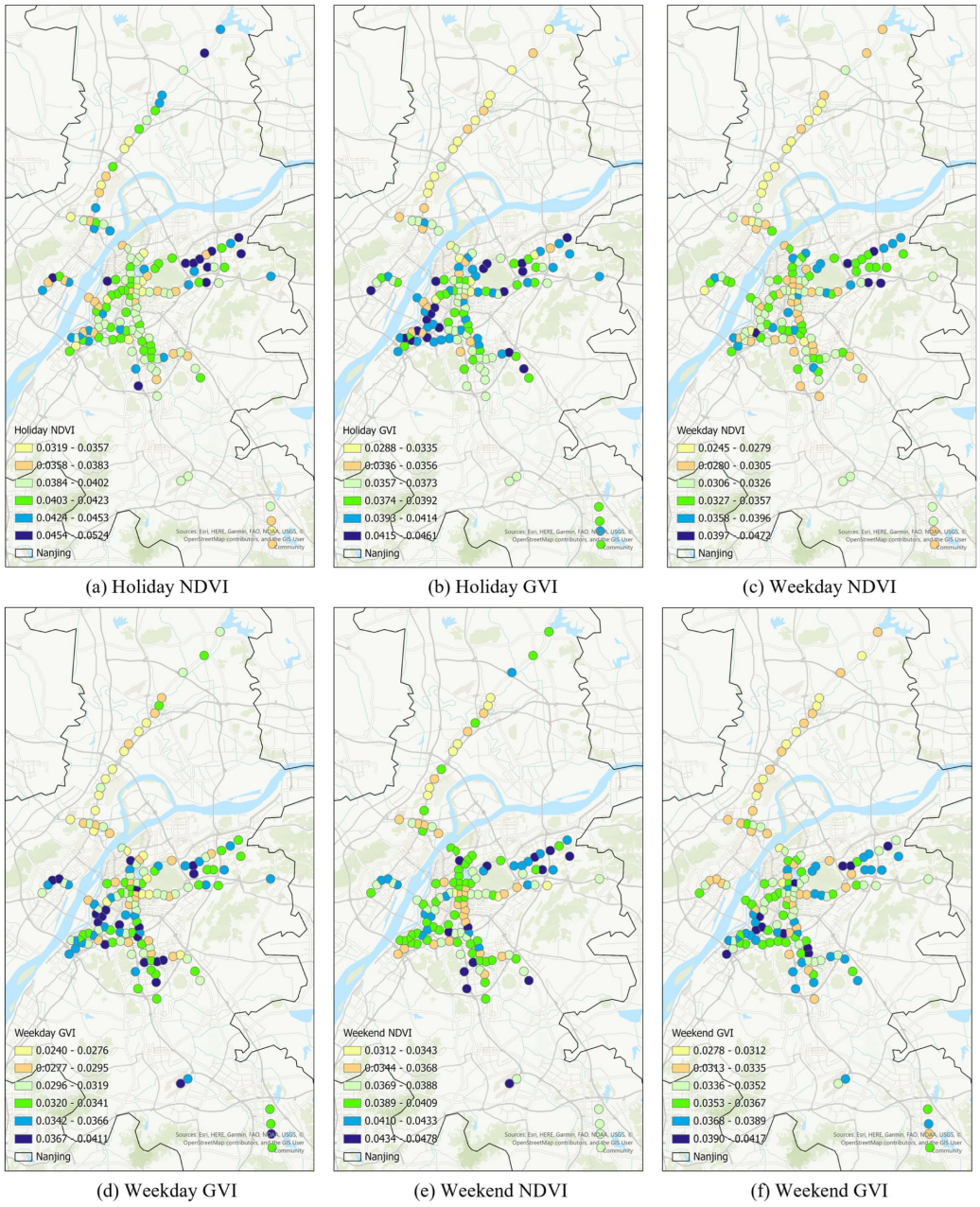

Figure 5. Spatial distribution of local feature importance for NDVI and GVI under different travel scenarios.

Stations with high NDVI importance are primarily distributed in new urban districts on the city's outskirts, university clusters, and areas with high green-space coverage, while most stations in the central urban area show relatively low importance. Under weekday scenarios, high-value stations are mainly scattered across the eastern and northern outskirts. During weekends, high-importance stations become more common and extend toward Xianlin University Town and the Jiangbei area. During holidays, high-importance stations form a broader cluster in areas with ecological and landscape resources or recreational functions. Only a few representative stations, such as Xuzhuang Suning Headquarters and Qilinmen, reach the highest importance levels across the network.

The spatial distribution of GVI importance is more dispersed. High values occur more frequently in commercial centers, transportation hubs, and high-density built-up areas, rather than being concentrated only in areas with high vegetation cover. On weekdays, high-importance stations are primarily distributed in the central urban area and large mixed-use districts. During weekends and holidays, high-value stations extend toward tourism and leisure zones, large public-event areas, and urban subcenters, producing a more dispersed pattern than that of NDVI. This pattern suggests that the modeled importance of GVI varies with street settings and station contexts.

Comparing the two indicators, high NDVI importance shows stronger peripheral clustering around areas with ecological resources and open spaces, whereas high GVI importance is more dispersed and occurs more frequently in built-up areas with active urban functions. A small number of stations have high importance for both indicators. These contrasting spatial patterns suggest that NDVI and GVI encode complementary greenness information at regional and pedestrian-observation scales.

4.3. Spatial Patterns of Local SHAP Contributions

Figure 6 illustrates the spatial patterns of station-level NDVI and GVI SHAP contributions across the three travel scenarios. To facilitate comparison across the six panels, the values were visualized using Delaunay triangulation and linear interpolation within the data boundary, without extrapolation or additional smoothing. The resulting surfaces are interpolated visualizations of station-level contributions and do not represent directly observed SHAP values between metro stations.

The mapped NDVI SHAP contributions show relatively strong regional continuity. Positive SHAP contributions form a relatively continuous band along the southern urban development axis of Nanjing, while negative contributions are more prevalent in the north and northeast. Under the weekday scenario, positive SHAP contributions are primarily distributed along the southern urban development axis of Nanjing and the Jiangning cluster, while the northern part of the central urban area, Xianlin satellite city, and parts of Jiangbei are dominated by negative contributions. As the travel scenario shifts from weekdays to weekends and holidays, the positive southern zone expands toward the southern edge of the central urban area, whereas Xianlin and parts of the northeast remain predominantly negative. These maps describe the spatial configuration of station-level SHAP contributions across the observed travel scenarios.

GVI exhibits more pronounced local heterogeneity. Positive and negative GVI SHAP contributions form discrete patches across the main urban area, Hexi New Town, Jiangning, and outlying clusters, and adjacent stations often show different contribution directions. Under weekday scenarios, SHAP values at most stations cluster around zero, with only a few stations showing stronger positive or negative contributions. During weekends and holidays, the magnitude of both positive and negative SHAP values increases, and localized high- and low-value areas become more evident, while the overall distribution

remains fragmented. This pattern suggests that the model-based association between visible greenery and predicted ridership varies across local station contexts.

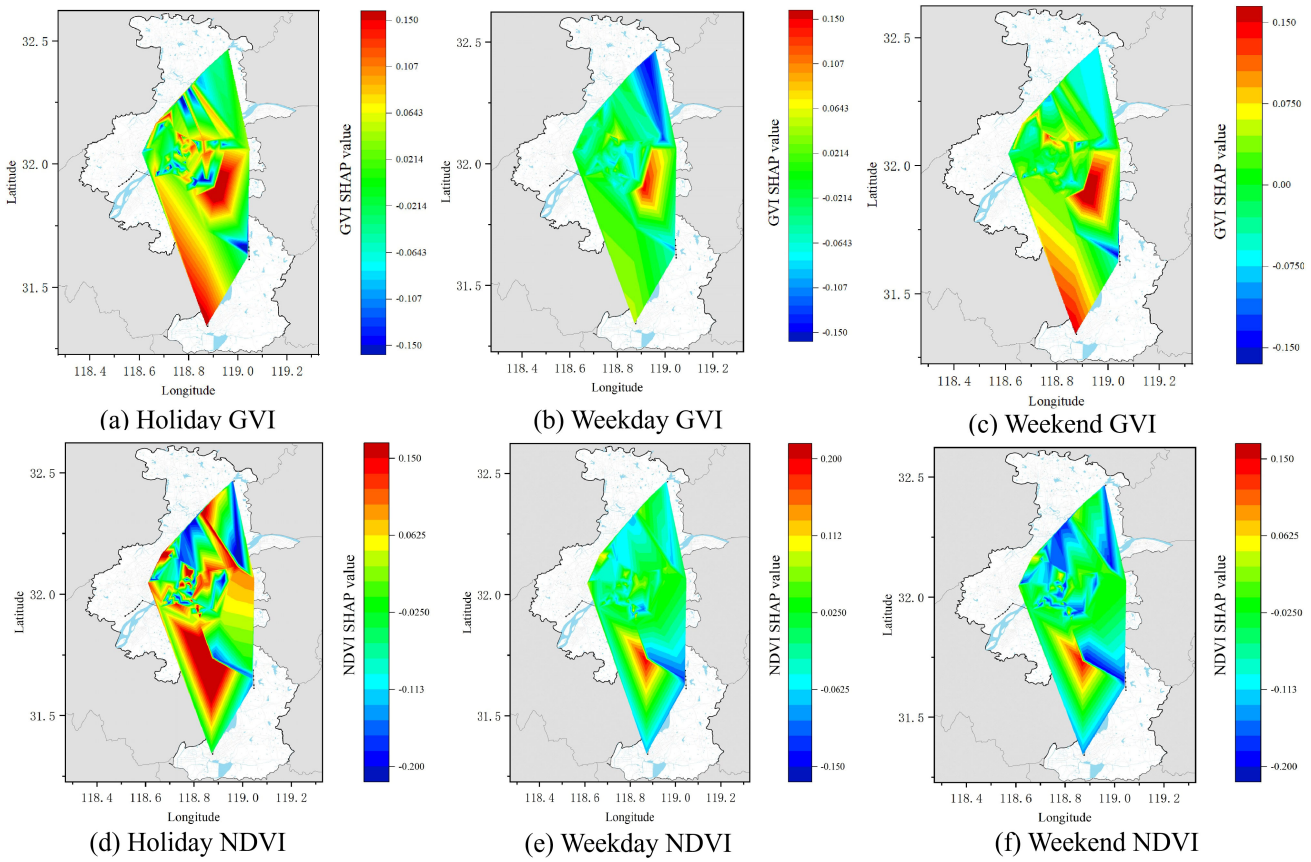

Figure 6. Spatial distribution of local SHAP contributions of NDVI and GVI under different travel scenarios.

Figure 7 shows station-level SHAP contributions across the three travel scenarios. Integrated transportation hubs and core urban functional zones exhibit substantial scenario-dependent variation. For example, Xinjiekou, Nanjing Station, Nanjing South Station, and Wudingmen show marked changes in NDVI SHAP contributions across scenarios. While weekday and weekend values differ relatively little, their holiday SHAP values are more negative. In contrast, outlying clusters, university districts, and recreational stations primarily exhibit changes in the direction of GVI SHAP contributions. For example, Hongshan Zoo, Maqun, and Xuezhe Road stations shift from negative contributions on weekdays to positive contributions during holidays. These stations are located in ecological, recreational, university, or peripheral urban areas, suggesting that the model-based association between visible greenery and ridership varies with travel context and co-occurring station functions.

4.4. Cross-Scenario SHAP Patterns and Partial Dependence Results

Figure 8 shows the distribution of three-dimensional SHAP values for NDVI and GVI across weekdays, weekends, and holidays. Each data point represents a subway station, and the three axes correspond to the local SHAP values of the variable under the holiday, weekday, and weekend scenarios, respectively. Different colors and symbols indicate combinations of positive and negative local SHAP values across the three scenarios. A "+" indicates a positive contribution to the model prediction relative to its baseline, whereas a "−" indicates a negative contribution.

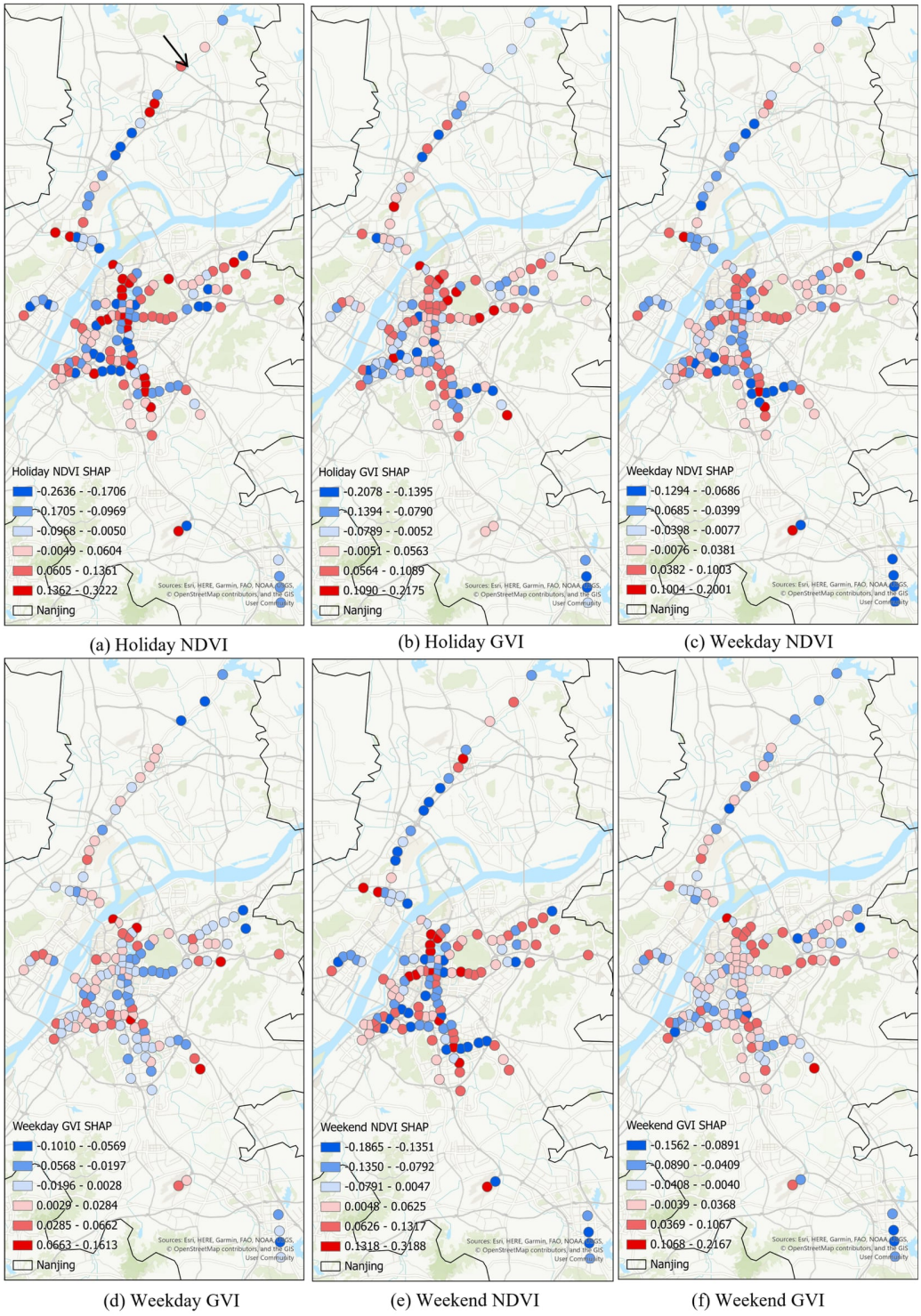

Figure 7. Station-level variation in local SHAP contributions of NDVI and GVI across different travel scenarios.

For most stations, NDVI values are concentrated in the "H+ W+ WE+" and "H− W− WE−" combinations, with only a small number of stations distributed across other quadrants. This indicates that NDVI generally maintains a consistent SHAP-contribution direction across travel scenarios, while the magnitude of its local contribution varies among stations. Sign changes in local NDVI SHAP contributions are relatively uncommon across travel scenarios.

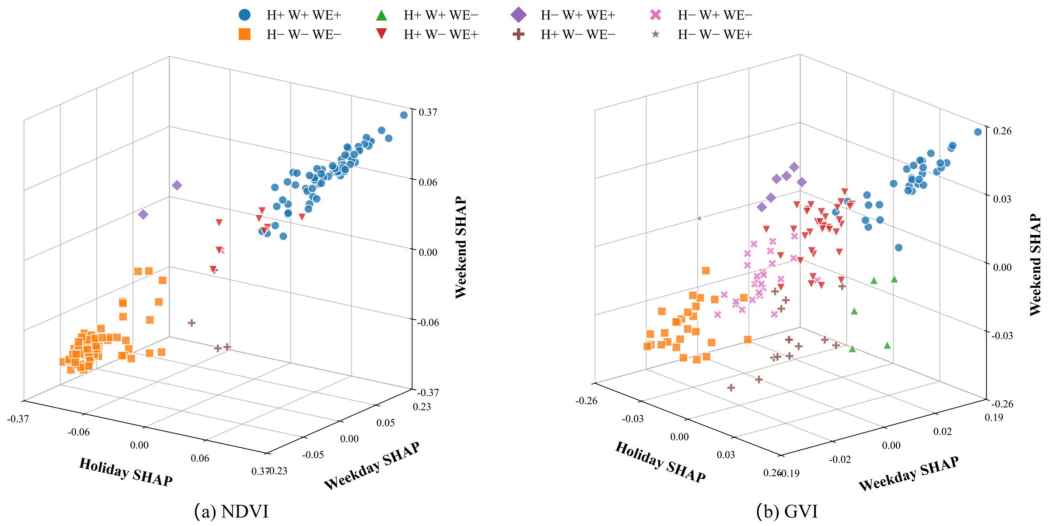

Figure 8. Comparison of local SHAP contribution directions for NDVI and GVI.

In contrast, the distribution of GVI's three-dimensional SHAP values is more dispersed. In addition to "H+ W+ WE+" and "H− W− WE−," many stations occur in combinations such as "H+ W− WE+," "H− W+ WE−," and "H+ W− WE−." Thus, a station is more likely to show changes in the direction of its local GVI SHAP contribution across travel scenarios. This pattern is consistent with the sensitivity of GVI to building shading, road cross-sections, canopy morphology, vertical vegetation structure, and the continuity of street-facing interfaces. A higher amount of green space does not necessarily correspond to higher visible greenery, nor does a higher GVI value necessarily correspond to a larger positive contribution to predicted ridership.

The results suggest hypotheses for differentiated station-area assessment rather than a uniform greening strategy. For commuter-oriented stations, shade, pedestrian connectivity, and transfer comfort may be relevant dimensions for further evaluation. At commercial, tourism, and leisure-oriented stations, particularly during weekends and holidays, the relationship between street-level greenery, pedestrian experience, and metro use warrants further investigation. At stations with high NDVI but low GVI, planners may evaluate whether regional vegetation is visible and accessible from station entrances and pedestrian routes, rather than focusing solely on increasing total green-space area.

Figure 9 shows the partial dependence curves for NDVI and GVI under weekday, weekend, and holiday scenarios. Both indicators exhibit nonlinear modeled patterns, with substantial differences in the magnitude of predicted-ridership variation across scenarios. The holiday curves show the greatest variation, whereas the weekday and weekend curves remain comparatively flat across much of the observed range. The PDPs reveal nonlinear and scenario-specific associations between greenness and predicted ridership. Under the holiday scenario, predicted ridership is relatively high at the lower end of the NDVI and GVI ranges, declines toward the middle range, and then increases as greenness reaches higher values before becoming relatively stable. This pattern is U-shaped or threshold-like rather than monotonic. The two indicators show similar broad nonlinear patterns but differ in the timing and magnitude of their turning points. The GVI holiday curve shows a sharper rebound after its lowest range, whereas the NDVI curve increases more gradually before reaching a plateau.

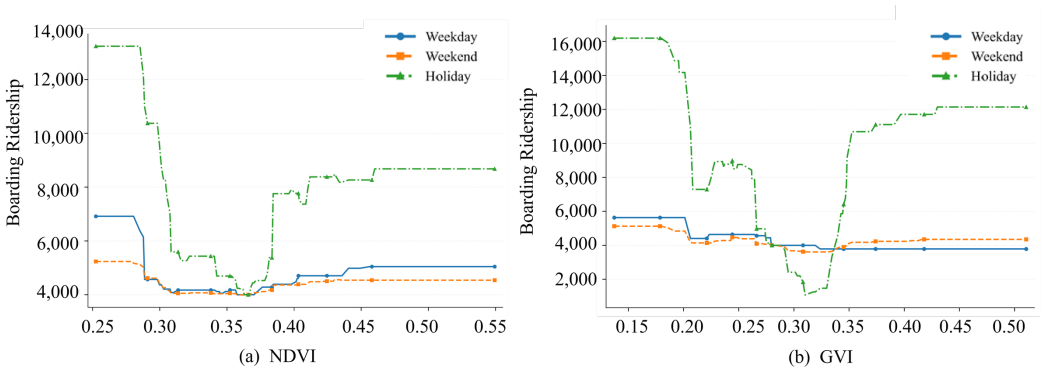

Figure 9. Partial dependence of metro ridership on NDVI and GVI under different travel scenarios.

5. Conclusions

This study investigated Nanjing Metro and characterized urban green space around metro stations from two perspectives: area-level vegetation coverage and street-level visible greenery. Three travel scenarios—weekdays, weekends, and holidays—were further distinguished. A Local XGBoost Ensemble was developed to account for station-level spatial heterogeneity while retaining the ability of machine learning to capture nonlinear relationships through local modeling and distance-based ensemble prediction. The Local XGBoost Ensemble achieved good predictive performance across all three travel scenarios, with the highest accuracy observed for weekdays. NDVI and GVI exhibited distinct patterns of relative importance and spatial distribution in relation to metro ridership. Overall, NDVI was more important than GVI, while the importance of both indicators increased from weekdays to weekends and holidays. High-importance NDVI stations showed a stronger peripheral clustering pattern and were mainly located in areas with abundant ecological resources and large open spaces. In contrast, high-importance GVI stations were more concentrated in built-up areas with active urban functions and exhibited a more dispersed spatial pattern. The local contributions of NDVI showed greater spatial continuity, with relatively continuous positive associations along the southern urban development corridor and in the Jiangning area. These positive areas expanded further from weekdays to weekends and holidays. By comparison, the local SHAP values of GVI displayed a more fragmented spatial pattern, with neighboring stations sometimes showing different directions of association. Both indicators also exhibited nonlinear relationships characterized by thresholds and diminishing marginal responses.

These findings provide a basis for further investigation of green-space conditions around metro stations under the TOD framework. For example, stations with extensive surrounding vegetation but relatively low street-level visible greenery may warrant field assessment of pedestrian access routes, street-facing vegetation, and green-space visibility. Similarly, the differing spatial patterns of NDVI and GVI attribution suggest that station-area greenness should be considered together with station function, pedestrian conditions, land use, and travel demand rather than through a uniform greening strategy. Because the model does not adjust for these broader contextual characteristics, the results should not be interpreted as evidence that a particular greening intervention will independently increase metro ridership.

Several limitations should be acknowledged. First, the analysis focuses on the model-based predictive attribution of NDVI and GVI and does not adjust for a broader set of station-area and built-environment characteristics. The findings should therefore be interpreted as greenness-focused predictive associations rather than independent causal effects. Second, the AFC observations cover selected dates between April and October 2023, with more limited weekend coverage, and the greenness data are not fully temporally synchronized

with every AFC observation. The May–October NDVI comparison supports the relative stability of the station-level NDVI pattern, but the findings remain specific to the observed travel contexts. In addition, GVI was derived from pretrained image-segmentation results, and consistent street-view capture dates were unavailable. The spatial validation design represents a conservative geographically separated prediction setting. Its local training neighborhoods are restricted by the exclusion of complete spatial groups and therefore do not fully reproduce the local information available when the complete station network is observed. Finally, the model evaluation characterizes overall predictive performance but does not provide formal uncertainty intervals for individual station predictions or SHAP values. Future research can extend this work using year-round and temporally matched data, broader station-area variables, repeated spatial resampling, and comparative analyses across cities.

Author Contributions: Conceptualization, Xinyu Hu and Yong Zhang; methodology, Xinyu Hu, Zhenjun Zhu, and Lu Yu; software, Xinyu Hu and Zhenjun Zhu; validation, Xinyu Hu and Baorui Han; formal analysis, Zhenjun Zhu and Lu Yu; investigation, Yong Zhang; resources, Baorui Han; data curation, Xinyu Hu and Yong Zhang; writing—original draft preparation, Xinyu Hu; writing—review and editing, Zhenjun Zhu and Lu Yu; visualization, Baorui Han; supervision, Baorui Han; project administration, Baorui Han. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 52402385.

Data Availability Statement: The Sentinel-2 imagery used in this study is publicly available through the Google Earth Engine platform. The street-view images used for GVI calculation were obtained from the Baidu Maps Open Platform. The AFC data were provided by Nanjing Metro Group Co., Ltd., and are not publicly available due to data-sharing restrictions.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Cervero, R.; Sullivan, C. Green TODs: Marrying Transit-Oriented Development and Green Urbanism. *Int. J. Sustain. Dev. World Ecol.* **2011**, *18*, 210–218. [CrossRef]
2. An, D.; Tong, X.; Liu, K.; Chan, E.H.W. Understanding the Impact of Built Environment on Metro Ridership Using Open Source in Shanghai. *Cities* **2019**, *93*, 177–187. [CrossRef]
3. Lu, Y.; Yang, Y.; Sun, G.; Gou, Z. Associations between Overhead-View and Eye-Level Urban Greenness and Cycling Behaviors. *Cities* **2019**, *88*, 10–18. [CrossRef]
4. Song, J.; Zhou, S.; Kwan, M.-P.; Liang, S.; Lu, J.; Jing, F.; Wang, L. The Effect of Eye-Level Street Greenness Exposure on Walking Satisfaction: The Mediating Role of Noise and PM2.5. *Urban For. Urban Green.* **2022**, *77*, 127752. [CrossRef]
5. Liu, Q.; Dai, S.; Yang, G.; Wu, C. Understanding Multi-Perspective Urban Green Space Patterns under Urban Expansion: Evidence from Guangzhou. *Sustain. Cities Soc.* **2026**, *147*, 107534. [CrossRef]
6. Li, Z.; Zhang, D. Relative Attractiveness of Green Spaces On Subway Ridership: An OD-GWR Analysis of the Residence-Green Mismatch. *Netw. Spat. Econ.* **2026**, *26*, 493–524. [CrossRef]
7. Luo, Y.; Li, B.; Zhang, H.; Kang, M.; Su, S. Unlocking the Nonlinear TOD-Metro Ridership Relationship: A Novel Machine Learning Approach Embedding Spatiotemporal Heterogeneity. *J. Transp. Geogr.* **2025**, *126*, 104222. [CrossRef]
8. Li, L.; Zhong, L.; Ran, B.; Du, B. Analysis of the Relationship between Metro Ridership and Built Environment: A Machine Learning Method Considering Combinational Features. *Tunn. Undergr. Space Technol.* **2024**, *144*, 105564. [CrossRef]
9. Ding, C.; Cao, X.; Liu, C. How Does the Station-Area Built Environment Influence Metrorail Ridership? Using Gradient Boosting Decision Trees to Identify Non-Linear Thresholds. *J. Transp. Geogr.* **2019**, *77*, 70–78. [CrossRef]
10. Sun, Y.; Shi, J.; Schonfeld, P.M. Identifying Passenger Flow Characteristics and Evaluating Travel Time Reliability by Visualizing AFC Data: A Case Study of Shanghai Metro. *Public Transp.* **2016**, *8*, 341–363. [CrossRef]
11. Lee, E.H.; Kim, K.; Kho, S.Y.; Kim, D.K.; Cho, S.H. Estimating Express Train Preference of Urban Railway Passengers Based on Extreme Gradient Boosting (XGBoost) Using Smart Card Data. *Transp. Res. Rec.* **2021**, *2675*, 64–76. [CrossRef]

12. Long, Y.; Thill, J.-C. Combining Smart Card Data and Household Travel Survey to Analyze Jobs-Housing Relationships in Beijing. *Comput. Environ. Urban Syst.* **2015**, *53*, 19–35. [CrossRef]
13. Yang, H.; Shi, J.; Tao, T. Where Do Built Environment Attributes Most Effectively Influence Bike Sharing Usage? *Transp. Res. Part D Transp. Environ.* **2025**, *143*, 104717. [CrossRef]
14. Gan, Z.; Yang, M.; Feng, T.; Timmermans, H.J.P. Examining the Relationship between Built Environment and Metro Ridership at Station-to-Station Level. *Transp. Res. Part D Transp. Environ.* **2020**, *82*, 102332. [CrossRef]
15. Chen, E.; Ye, Z.; Wang, C.; Zhang, W. Discovering the Spatio-Temporal Impacts of Built Environment on Metro Ridership Using Smart Card Data. *Cities* **2019**, *95*, 102359. [CrossRef]
16. Li, S.; Lyu, D.; Liu, X.; Tan, Z.; Gao, F.; Huang, G.; Wu, Z. The Varying Patterns of Rail Transit Ridership and Their Relationships with Fine-Scale Built Environment Factors: Big Data Analytics from Guangzhou. *Cities* **2020**, *99*, 102580. [CrossRef]
17. Gao, F.; Yang, L.; Han, C.; Tang, J.; Li, Z. A Network-Distance-Based Geographically Weighted Regression Model to Examine Spatiotemporal Effects of Station-Level Built Environments on Metro Ridership. *J. Transp. Geogr.* **2022**, *105*, 103472. [CrossRef]
18. Li, R.; Wang, X.; Lovreglio, R.; Ding, H.; Wang, Q.; Chen, J.; Jiang, E.; Ma, J. Influence of Subsequent Path and Pressure on Pedestrian Route Choice in Emergency Evacuations. *Phys. A* **2025**, *657*, 130252. [CrossRef]
19. Liu, S.; Yao, E. Holiday Passenger Flow Forecasting Based on the Modified Least-Square Support Vector Machine for the Metro System. *J. Transp. Eng. Part A Syst.* **2017**, *143*, 04016005. [CrossRef]
20. Wang, Z.; Zheng, X.; Meng, F.; Wang, K.; Wu, X.; Yu, D. Exploring the Joint Influence of Built Environment Factors on Urban Rail Transit Peak-Hour Ridership Using DeepSeek. *Buildings* **2025**, *15*, 1744. [CrossRef]
21. Zhao, P.; Li, S. Suburbanization, Land Use of TOD and Lifestyle Mobility in the Suburbs: An Examination of Passengers' Choice to Live, Shop and Entertain in the Metro Station Areas of Beijing. *J. Transp. Land Use* **2018**, *11*, 195–215. [CrossRef]
22. Zhou, Y.; Wang, H.; Chen, S.; Jiang, J.; Wang, Z.; Liu, W. Investigating Holiday Subway Travel Flows with Spatial Correlations Using Mobile Payment Data: A Case Study of Hangzhou. *Sustainability* **2025**, *17*, 5873. [CrossRef]
23. Taran, I.; Szarata, M.; Kowalski, K. The relationship between peak-hour passenger demand and daily station ridership: A case study of the rzeszów agglomeration. *Transp. Probl. Int. Sci. J.* **2026**, *21*, 231–244. [CrossRef]
24. Yin, C.; Wen, R.; Zhou, Y.; Shao, C.; Wang, X. How Last-Mile Amenities Affect Metro Ridership? Evidence from Shanghai. *Transp. Plan. Technol.* **2025**, 1–23. [CrossRef]
25. Liu, X.; Chen, J.; Du, B.; Yan, X.; Peng, Q.; Shen, J. Resilience Assessment of Urban Rail Transit Stations Considering Disturbance and Time-varying Passenger Flow. *Comput.-Aided Civ. Infrastruct. Eng.* **2025**, *40*, 1627–1651. [CrossRef]
26. Ge, Y.; Xu, Z.; Yin, C.; Wang, X. Revealing Nonlinear and Spatial Interaction Effects of Built Environment on Ride-Hailing Demand in Nanjing, China. *Buildings* **2025**, *15*, 2967. [CrossRef]
27. Zhang, T.; Tang, F.; Hu, Y.; Zhang, L.; Guo, Y. Shaping Greener Mobility: Impact of Urban Greening Structure on Time-Dependent Bike-Sharing Usage. *Transp. Res. Part D Transp. Environ.* **2025**, *141*, 104657. [CrossRef]
28. An, R.; Wu, Z.; Tong, Z.; Qin, S.; Zhu, Y.; Liu, Y. How the Built Environment Promotes Public Transportation in Wuhan: A Multiscale Geographically Weighted Regression Analysis. *Travel Behav. Soc.* **2022**, *29*, 186–199. [CrossRef]
29. Du, Q.; Zhou, Y.; Huang, Y.; Wang, Y.; Bai, L. Spatiotemporal Exploration of the Non-Linear Impacts of Accessibility on Metro Ridership. *J. Transp. Geogr.* **2022**, *102*, 103380. [CrossRef]
30. Xi, Y.; Hou, Q.; Duan, Y.; Lei, K.; Wu, Y.; Cheng, Q. Exploring the Spatiotemporal Effects of the Built Environment on the Nonlinear Impacts of Metro Ridership: Evidence from Xi'an, China. *IJGI* **2024**, *13*, 105. [CrossRef]
31. Wu, W.; Yao, Y.; Wang, R. Green Space Exposure at Subway Stations, Transportation Mode Choice and Travel Satisfaction. *Transp. Res. Part D Transp. Environ.* **2023**, *122*, 103862. [CrossRef]
32. Wang, R.; Lu, Y.; Wu, X.; Liu, Y.; Yao, Y. Relationship between Eye-Level Greenness and Cycling Frequency around Metro Stations in Shenzhen, China: A Big Data Approach. *Sustain. Cities Soc.* **2020**, *59*, 102201. [CrossRef]
33. Dong, S.; Wang, Y.; Wang, C.; Dou, M.; Gong, J. Discovering the Nonlinear Association between the Built Environment and Metro Ridership: Global and Local Perspectives. *Cities* **2026**, *169*, 106561. [CrossRef]
34. Larkin, A.; Hystad, P. Evaluating Street View Exposure Measures of Visible Green Space for Health Research. *J. Expo. Sci. Environ. Epidemiol.* **2019**, *29*, 447–456. [CrossRef]
35. Li, T.; Zheng, X.; Wu, J.; Zhang, Y.; Fu, X.; Deng, H. Spatial Relationship between Green View Index and Normalized Differential Vegetation Index within the Sixth Ring Road of Beijing. *Urban For. Urban Green.* **2021**, *62*, 127153. [CrossRef]
36. Hua, J.; Cai, M.; Shi, Y.; Ren, C.; Xie, J.; Chung, L.C.H.; Lu, Y.; Chen, L.; Yu, Z.; Webster, C. Investigating Pedestrian-Level Greenery in Urban Forms in a High-Density City for Urban Planning. *Sustain. Cities Soc.* **2022**, *80*, 103755. [CrossRef]
37. Huang, Y.; Sanatani, R.P.; Liu, C.; Kang, Y.; Zhang, F.; Liu, Y.; Duarte, F.; Ratti, C. No "True" Greenery: Deciphering the Bias of Satellite and Street View Imagery in Urban Greenery Measurement. *Build. Environ.* **2025**, *269*, 112395. [CrossRef]
38. Li, S.; Lyu, D.; Huang, G.; Zhang, X.; Gao, F.; Chen, Y.; Liu, X. Spatially Varying Impacts of Built Environment Factors on Rail Transit Ridership at Station Level: A Case Study in Guangzhou, China. *J. Transp. Geogr.* **2020**, *82*, 102631. [CrossRef]

39. Zhu, Z.; Zhang, Y.; Qiu, S.; Zhao, Y.; Ma, J.; He, Z. Ridership Prediction of Urban Rail Transit Stations Based on AFC and POI Data. *J. Transp. Eng. Part A Syst.* **2023**, *149*, 04023077. [CrossRef]
40. Dong, W.; Wang, N.; Dong, Y.; Cao, J. Examining the Nonlinear and Interactive Effects of Built Environment Characteristics on Travel Satisfaction. *J. Transp. Geogr.* **2025**, *123*, 104111. [CrossRef]
41. Jiang, W.; Li, Q.; Li, J.; Zhu, Z. Differentiated Impacts of Urban Streetscapes on Green Mobility Experiences. *J. Asian Archit. Build. Eng.* **2025**, 1–17. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.