International Journal of *Geo-Information*

Article

The Non-Linear Influence of Built Environment on the School Commuting Metro Ridership: The Case in Wuhan, China

Jinming Yan [1], Qiuyu Wan [2], Jingyi Feng [1], Jianjun Wang [1], Yiwen Hu [2] and Xuexin Yan [2,3,*]

1 Guangzhou Urban Planning & Design Survey Research Institute, Guangzhou 510030, China
2 School of Urban Design, Wuhan University, Wuhan 430072, China
3 Department of City and Regional Planning, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA
* Correspondence: yan.xuexin902@whu.edu.cn; Tel.: +86-152-0848-0694

Abstract: Although many studies have investigated the non-linear relationship between the built environment and rail patronage, it remains unclear whether this influence is equally applicable to primary and secondary school students due to their physiological characteristics and cognitive limitations. This study applies the GBDT model to Wuhan student metro swipe data in order to investigate the relative importance and non-linear association of the built environment on the school-commuting metro ridership. The results show that the variable with the greatest predictive power is the number of living service facilities followed by the number of intersections, and the degree of land-use mixture. All of the built environment variables had non-linear associations with the school-commuting ridership, and the greatest attraction to the school-commuting metro ridership occurred when the number of living service facilities was 500, the number of intersections was 36, and the degree of land-use mixture was 0.8. These findings can help planners to prioritize land-use optimization and the effective range of land-use indicators when developing child-friendly rail transport policies.

Keywords: non-linear; built environment; school commuting; GBDT

check for updates

Citation: Yan, J.; Wan, Q.; Feng, J.; Wang, J.; Hu, Y.; Yan, X. The Non-Linear Influence of Built Environment on the School Commuting Metro Ridership: The Case in Wuhan, China. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 193. https://doi.org/10.3390/ijgi12050193

Academic Editors: Hartwig H. Hochmair and Wolfgang Kainz

Received: 28 January 2023
Revised: 25 April 2023
Accepted: 1 May 2023
Published: 6 May 2023

1. Introduction

Promoting active travel among children has important implications for children's health and for reducing traffic congestion, especially around schools. Active travel has been proved to help improve children's physical activity levels and overall health [1,2]; however, urban sprawl has been common across the globe over the past half century, which has seen cities grow in size. Many countries have adopted a policy, "Attend the school in the region where they live", in order to reduce commuting distances for primary and secondary school students; however, the uneven distribution of educational resources has forced some students to commute long distances, and the length of school commutes has increased significantly. Studies based on countries around the world have found that when walking distances exceed a certain threshold, students' travel patterns gradually shift from walking and cycling to taking their parents' private cars [3–5]. This reliance on cars has had significant negative effects on the environment and has negatively impacted urban traffic congestion and children's physical activity [6,7]. As an important measure to relieve traffic pressure in large cities, rail transportation has gradually become the backbone transportation system for passenger transportation in major cities around the world. Under the premise of ensuring the safety of primary and secondary school students' travel, it is crucial to encourage them to use rail transportation in order to reduce urban traffic congestion and enhance their physical health; however, there is a lack of existing studies on metro travel among primary and secondary school students.

In addition, the 5Ds; density, diversity, design, destination accessibility, and distance to public transport; are considered to be the core elements of the built environment, which has long been recognized as an effective factor influencing passenger flow at metro stations [8,9]; however, previous studies have focused on aggregated passenger ridership or commuters at the rail station level [10,11]. Children may be more sensitive to perceptions of the built environment than adults, so the applicability of these findings to children remains to be further validated. In addition, most existing studies assume a linear relationship between the built environment and metro ridership, ignoring the complex non-linear association between them and the relative importance of the built environment's influence on metro ridership [12–15]. Some recent studies have shown that the built environment has a significant non-linear effect on metro station flow, and the importance of the effect also varies significantly [16,17]. Therefore, further insight into the mechanisms of the non-linear effects of the built environment on the metro ridership of primary and secondary school students and the relative importance of the independent variables can provide more detailed and nuanced guidance for the development of child-friendly rail transit systems.

To fill the above research gap, this study applies the gradient boosting decision tree (GBDT) model in machine learning to the smart swipe data of Wuhan, China, in 2019. By exploring the key built environment factors affecting the rail transit school-commuting flow, we deeply analyze the relative importance of these factors and the complex non-linear relationships among them, and then assist urban planning and related policy formulation in a more targeted manner at the micro level. This study has two main contributions: First, it focuses on the rail transit travel of primary and secondary school students. The effects of the early built environment on metro ridership are mostly for all groups or commuters, while the special physiological characteristics of primary and secondary school students make the results of existing studies not necessarily applicable. Second, most existing studies assume a linear or logarithmic relationship between the built environment and metro ridership, and further in-depth exploration of the non-linear relationship and the relative importance of the influencing elements allows planners to better formulate land-use optimization policies.

The remainder of the article unfolds as follows: First, we review existing studies on the relationship between built environments of stations and school-commuting flow. Next, we present the study area, data, and methodology. In the results section, we first describe the relative importance of the built environment on the school-commuting metro ridership, and then, we further analyze the non-linear mechanism of the built environment's influence on the school-commuting metro ridership. Finally, we conclude with the policy implications and planning implications of this paper.

2. Literature Review

Previous researchers have identified some key variables that affect metro ridership. Since this study aims to examine the non-linear effects of the built environment of stations on the school-commuting metro ridership, we focus on reviewing existing research in the following areas: (1) the scope definition of metro station areas; (2) the effects of the built environment of stations on the metro ridership; (3) research methods; and (4) existing research inadequacies.

The metro station area coverage is often referred to as the pedestrian-intensive area, as walking is the most dominant mode of a rail connection; therefore, the coverage area of rail stations is usually defined based on the walking distance of residents. In the existing studies, 400–800 m (the range of a 5–10 min walk for residents) is used as the site impact area [15,17–19]; however, most studies recommend 800 m as the metro station area coverage, which is the range that would require a normal adult to walk for 10 min [20]. As the distance to the metro station increases further, residents may use bicycles, ground transportation, or private cars for rail transit connections. While considering the special characteristics of primary and secondary school students, their stride frequency and stride length are slightly weaker than those of adults [20]; therefore, referring to Peng et al.'s Wuhan-based study, the walking speed of primary and secondary school students is about 90% of that of adults, and thus, 720 m is set as the influence range of metro stations for primary and secondary school students in our study [21].

The built environment of the station area is considered to be an effective factor in attracting metro ridership and is reflected by the "5Ds", which are density, diversity, design, destination accessibility, and distance [8]. Among them, density (including population density and spatial density) is a major influencing factor on metro ridership. Existing studies show that the higher the population density and the higher the floor area ratio, the more beneficial it is to use the metro for travel [19,22]. Diversity is mainly expressed by the degree of land-use mixture. A higher degree of land-use mixture indicates that the area has better public service facilities and is more conducive to residents walking or using public transportation [17,23]. At the design-plan level, convenient and appropriate neighborhood scales are more conducive to residents using public transportation, but too many intersections are not conducive to residents using public transportation [24,25]. In terms of destination accessibility, public service facilities, such as amenities, catering facilities, and companies around the rail transit stations, have a remarkably positive effect on metro ridership [12,26]. In terms of distance from public transport, the longer the length of the roads and the more bus stops around a metro station, the better the accessibility of the metro, and the more conducive metro travel is for primary and secondary school students. Rail transit station characteristics also have an impact on metro ridership [27–29]. Some studies have shown that the number of transfer stations and entrances has a significant positive influence on the metro ridership [28,29]; however, most of the existing research is based on commuters or all groups. In contrast, primary and secondary school students are more sensitive to the perception of the built environment, and the influence of the built environment on them may be significantly different from other groups; therefore, this study aims at the specific group of primary and secondary school students to explore the effects of the built environment on the school-commuting metro ridership. In addition to this, residents' income is also an influential factor on metro ridership. Usually, the higher the income of residents, the more willing they are to use private cars to travel [30]. This study uses house prices as a measure of residents' income to explore the effect of residents' income on the school-commuting metro ridership.

The ordinary least squares (OLS) model is the most common method used in the studies of the effects of the built environment of a stations on metro passenger flow [12,31]. Some studies have also introduced linear models, such as geographically weighted regression (GWR) and stepwise regression, to explore the relationship between them [32]; however, most studies assume a linear relationship between them. Some recent studies have found that there may be complex non-linear effects of the built environment of stations on metro ridership. For example, Liu et al., using the GBDT model, found that almost all built environment variables had a significant non-linear effect on the metro ridership in Chongqing [33]. They found that the distance between the departure point and the city center had a significant positive effect on the metro ridership between 0 and 2.8 km, and no positive effect beyond 2.8 km. Similarly, the number of companies in the departure area only has a positive effect on the metro patronage between 0 and 400 companies. [16]. The number of bus stops has a significant effect on the metro ridership only if it is from 11 to 25 per kilometer, and the degree of land-use mixture influences metro ridership only if it is from 0.5 to 0.65. Studies based on Washington, Nanjing, and other locations show similar non-linear impact results [17,19]. In addition, the GBDT model can determine the relative importance of the predictor variables. For example, Ding et al. found that the average of car ownership, the number of bus stops, and employment density were the three largest contributing factors on metro ridership [17]. Determining the relative importance of the influencing factors plays an important role in better prioritization and planning with limited resources; therefore, this study uses the GBDT model to explore the non-linear influence of the built environment of the station on the school-commuting metro ridership. Analyzing the specific influence queues of each variable and comparing the relative importance of each fact helps to determine which built environment elements play a more important role in predicting the school-commuting metro ridership. Determining the effective range of influence of these important variables will, in turn, help to provide a more refined rail service [26,34].

Although existing studies have identified the relationship between metro ridership and built environment factors; such as density, diversity, and design; studies have mainly focused on commuters or all groups [16,19]. Whether the existing results are equally applicable to primary and secondary school students needs to be further explored due to their specific physiological characteristics. In addition, most existing studies assume a linear relationship between the built environment and metro ridership, but some recent studies suggest that a complex non-linear relationship may exist between them. Whether this non-linear effect applies to the effect of the built environment on school-commuting metro ridership also needs further verification.

3. Research Design

3.1. Study Area

Our study is based in Wuhan, China, which is the educational center of the central region of China. The study area is the central urban area within the third ring road of Wuhan, which is a cluster of primary and secondary schools with over 500,000 primary and secondary school students (Figure 1). In this area, the need to travel to and from school for primary and secondary school students is an issue that cannot be ignored in urban development. In addition, as the construction of rail transport in Wuhan continues to accelerate, a rail transport system covering the main primary and secondary schools and residential areas within the third ring road has been formed. Influenced by the advantages of rail transport in terms of comfort, safety, and convenience, and being supported by Wuhan's policy for actively implementing preferential public transport rides for primary and secondary school students, the attractiveness of rail transport to primary and secondary school students has gradually increased. Based on existing research [12,15], we selected a circular area of 800 m around the metro station to measure the built environment.

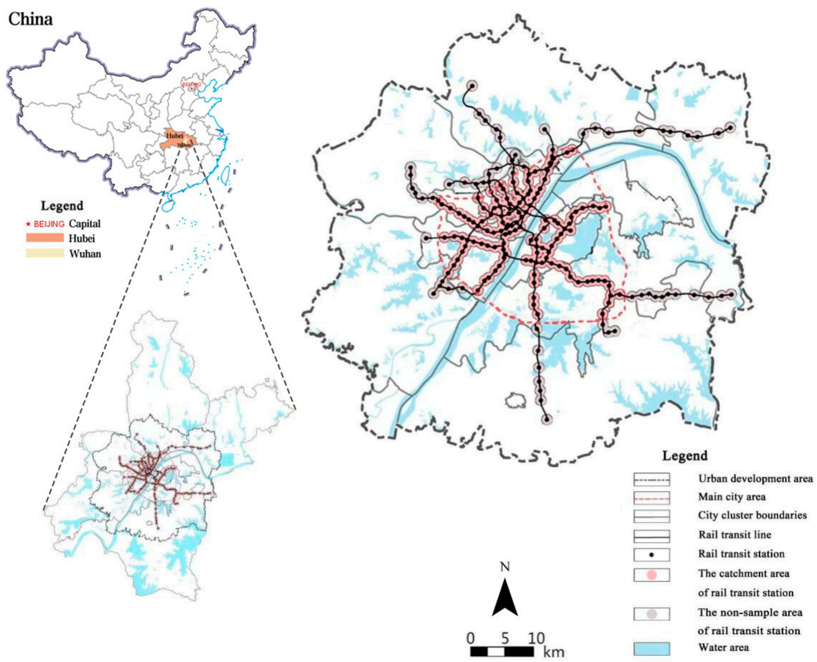


Figure 1. Study area.

3.2. Data and Variable Settings

This study uses the passenger flow data of 128 stations from 9 metro lines in the central urban area of Wuhan provided by the Wuhan Institute of Strategic Transport Development. The number of riders was determined from the smart card OD data of urban rail transit in Wuhan for one consecutive week in March 2019. The raw data records the card number of the cardholder, the nature of card use (including disabled, elderly, and student cards, etc.), the number of stations entered and exited, and the time of swiping the card. After excluding some invalid data, a total of 147,828 records of primary and secondary school students swiping their cards were extracted by the nature of the card use. The average passenger flow per station for primary and secondary school students was 104.

This paper examines the relationship between the built environment and the metro school-commuting flow and constructs a system of built environment indicators based on the 5Ds (density, diversity, design, destination accessibility, and distance) with reference to public transport stations [8]. Density includes floor area ratio and student density. The entropy index is used to measure land-use diversity. Design is represented by the number of intersections and parking lots, which measure street connectivity. Accessibility to destinations was chosen to include living service facilities, dining facilities, companies, and shopping centers, all of which are relevant to residents' daily travel purposes. The distance of the station from the city center and the number of buses and road lengths within the station areas are used to measure the accessibility of the station. Station characteristics are expressed in terms of the number of entrances and exits and the presence of interchange stations, where the presence of interchange stations is a dummy variable for non-interchange stations. The socio-economic attributes of residents are expressed in terms of house prices. Built environment data was obtained through the 2019 Baidu API, and house price data was obtained from a real estate transaction website (Lianjia). Table 1 summarizes these variables.

Table 1. Variable definitions and data summary.

Variables	**Calculation and Interpretation**	**Mean**	**SD**
Dependent variable			
School commuting metro ridership	Average number of boarding per station calculated from Wuhan Metro card data	103.523	114.328
Independent variable			
Density			
Floor area ratio	Floor area ratio within the catchment area, ratio of total floor area to site area	1.067	0.580
Student density	Student density within the catchment area (persons/km^2)	1445.835	1068.308
Diversity			
The degree of land-use mixture	$\text{Landuse} = \frac{-\sum_{i=1}^{k} P_{ki} \ln(P_{ki})}{\ln k}$ where k is the number of land-use types in the area around station i. P_{ki} is the proportion of the area of k and the type of land use in the catchment area.	0.643	0.153
Design			
Number of intersections	Number of road junctions within the catchment area	24.781	14.599
Number of parking lots	Number of car parks within the catchment area	78.828	63.500
Destination			
Number of primary and secondary schools	Number of primary and secondary schools within the catchment area	1.820	1.856
Living service facilities	Number of living service facilities within the catchment area	389.484	350.595
Dining facilities	Number of dining facilities within the catchment area	323.414	297.502
Companies	Number of companies within the catchment area	168.484	158.348
Shopping centers	Number of shopping centers within the catchment area	737.070	893.208
Distance			
Distance from the city center	Straight line distance of the station from the city center (m)	8005.988	3617.339

Table 1. *Cont.*

Variables	Calculation and Interpretation	Mean	SD
Number of bus stops	Number of bus stops within the station area	9.547	4.786
Road length	Total length of roads within the catchment area (m)	9818.032	3091.995
Characteristics of rail transit stations			
Transfer station	Dummy variables, where 1 means the station is a transfer station, and 0 means the station is not a transfer station	0.156	0.363
Exit quantity	Number of entrances and exits within the station area	5.188	3.391
Social attributes			
House prices	Average house price within the catchment area (CNY/m^2)	17,245.90	5552.74

3.3. Methods

In this study, a gradient boosting decision tree (GBDT) model was constructed to better analyze the non-linear effects of the built environment within the station area on the metro school-commuting flow and the degree of influence among the elements. We assume that x is a set of independent variables (including built environment characteristics, station characteristics, and residential socio-economic attribute characteristics), $F(x)$ is an approximation function for the dependent variable y (metro school-commuting flow), and GBDT estimates the function $I(x;\varepsilon_m)$ based on the accumulation of the basis function $F(x)$ after multiple rounds of iteration. Referring to existing studies [35], the GBDT model can be expressed as follows:

$$F(x) = \sum_{m=1}^{M} f_m(x) = \sum_{m=1}^{M} \alpha_{jm} I(x;\varepsilon_m) \tag{1}$$

where $I(x;\varepsilon_m)$ denotes the decision tree, ε_m denotes the parameters of the tree, and M is the number of trees [36–38].

In this study, to suppress possible overfitting problems during GBDT operations, we qualified the residual learning results for each regression tree by introducing a learning rate factor φ $(0 < \varphi < 1)$ (i.e., a shrinkage algorithm):

$$f_m(x) = f_{m-1}(x) + \phi \cdot \textstyle\sum_{j=1}^{J} \varepsilon_{jm} I\left(x \in A_{jm}\right), 0 < \phi \leq 1 \tag{2}$$

We multiply each tree by a learning rate factor, a, to minimize the loss function. However, this generates more regression trees and significantly increases the number of learners. Another parameter is the complexity of the regression tree (i.e., the number of leaf nodes). In order to capture the complex interrelationships between the variables, it is often necessary to increase the number of leaf nodes of the regression tree. Therefore, the optimal fit of the GBDT model depends on the combination of the learning rate, the number of regression trees, and the complexity. The GBDT model in this study was estimated using the "gbm" package in the *R* platform.

4. Results

4.1. Relative Importance of Independent Variables

Table 2 illustrates the relative importance of the independent variables in predicting the school-commuting metro ridership. The number of amenities within the station area is the most contributing independent variable for predicting the school commuting metro ridership with a relative contribution of 15.57%. Places with a higher density of amenities tend to be more centrally located and have a higher level of security and are, therefore, more conducive to promoting active school-commuting for primary and secondary school students [39,40]. The higher the number of intersections, the better the street connectivity, and the more conducive it is to rail transit travel [41–43]. The degree of land-use mixture is the third most important independent variable with a contribution of 9.93%. This is consistent with the results of most studies [40,44]. The higher the degree of land-use mixture, the better the supporting public service facilities in the area, which makes riders more likely to choose metropolis rail transit. The number of primary and secondary schools only ranked 13th in terms of contribution to metro ridership, with a contribution of 2.55%. This is because Wuhan actively promotes the policy of enrolling primary and secondary school students in schools close to their neighborhoods, and the percentage of primary and secondary school students who use the metro to travel to school in Wuhan is still low compared to other modes of transportation, which is in line with the proportion of primary and secondary school students' metro ridership.

Table 2. Relative importance and ranking of the independent variables in predicting the school commuting metro ridership.

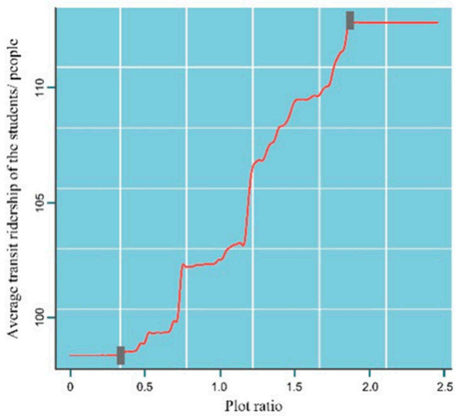
Variables	Rank	Relative Importance (%)
Density		
Floor area ratio	10	4.87
Student density	16	0.47
Diversity		
The degree of land-use mixture	3	9.93
Design		
Number of intersections	2	11.79
Number of parking lots	12	2.64
Destination		
Number of primary and secondary schools	13	2.55
Living service facilities	1	15.57
Dining facilities	7	5.44
Companies	9	4.97
Shopping center	8	5.09
Distance		
Distance from the city center	4	9.66
Number of bus stops	6	9.25
Road length	11	4.76
Characteristics of subway stations		
Whether it is a transfer station	5	9.61
Exit quantity	15	1.59
Social economy attributes		
House prices	14	1.79

Station characteristics also play a significant role in influencing the school-commuting metro ridership. Whether or not it is an interchange station has a greater impact on the school-commuting metro ridership, and it ranks fifth among all variables examined with a relative contribution of 9.61%. This is in line with the findings in Seoul [45,46]. With higher network accessibility, interchange stations are more attractive to passengers than non-interchange stations. The exit quantity also has an influence on the school-commuting metro ridership, but the contribution is relatively small. Socio-economic attributes also have a relatively small impact on the school-commuting metro ridership at 1.79%. Considering Wuhan's policy of offering a 30% discount to primary and secondary school students who ride the metro, we believe it is reasonable to assume that the economic impact on the school-commuting metro ridership is weaker than the built environment.

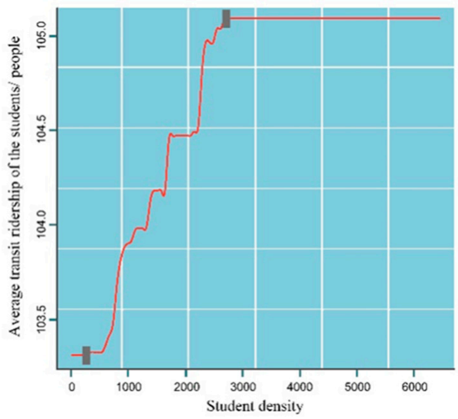
4.2. Non-Linear Association between Built Environment and School Commuting Metro Ridership

Existing studies usually assume a linear or logarithmic relationship between the built environment and metro ridership, but this does not necessarily apply to all built environment variables. Based on the dependency diagram derived from the GBDT, we found that built environment elements show significant non-linear and threshold effects on the school-commuting metro ridership.

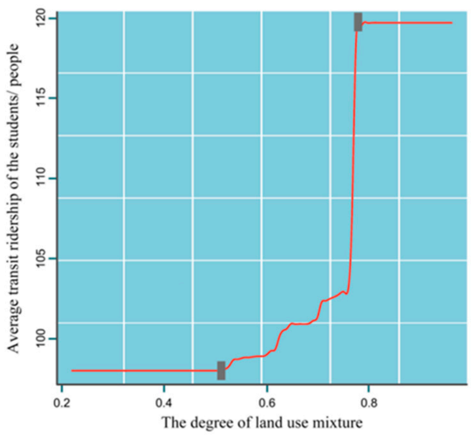
Figure 2a,b depicts the effects of the floor area ratio (FAR) and student density on the average school-commuting metro ridership. In line with most existing studies, the FAR exhibits a significant positive impact on the school-commuting metro ridership but shows a significant non-linear effect. When the FAR is 0.4 or less, it has little impact on the average school-commuting metro ridership. When the FAR increases from 0.4 to 1.9, the average school-commuting metro ridership increases from 98 to 113, and when the FAR increases further, it no longer has an impact on the average school-commuting metro ridership. Similar to the FAR, there is a significant positive effect of the number of students within the station area on the average school-commuting metro ridership. This variable only has an impact between the values of 500 and 2600, but the increase was smaller. This phenomenon is likely to be influenced by the proximity to school policy. Areas with more children of school age tend to be equipped with more primary and secondary schools. Students can walk or ride to school, so there is relatively little impact on school commuting-metro ridership.



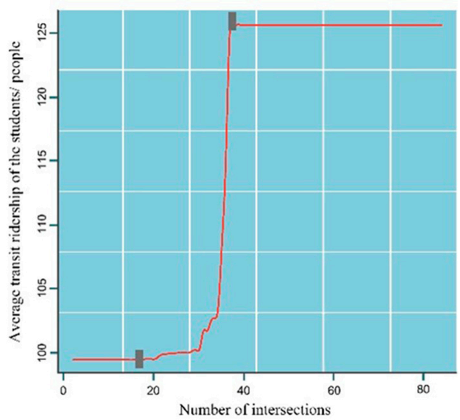
(**a**)



(**b**)

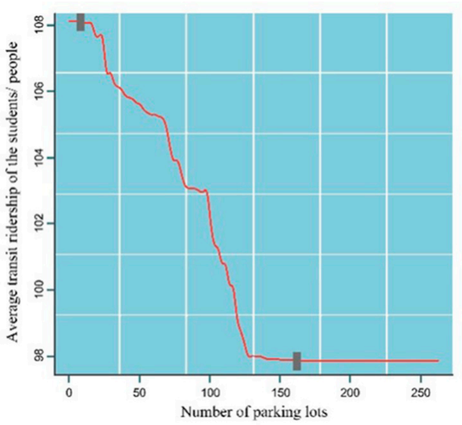


(**c**)

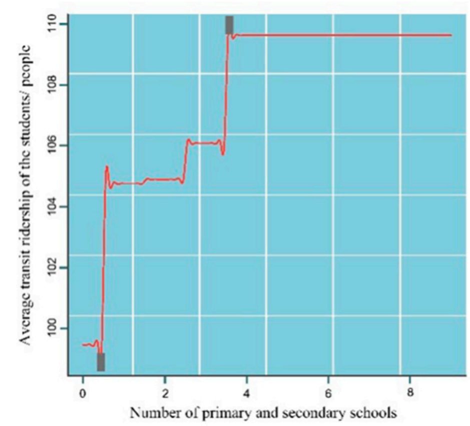


(**d**)

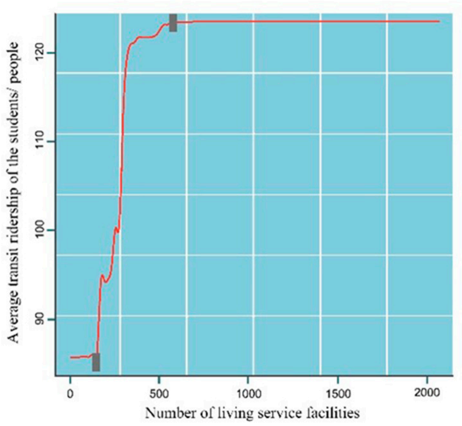
Figure 2. *Cont.*



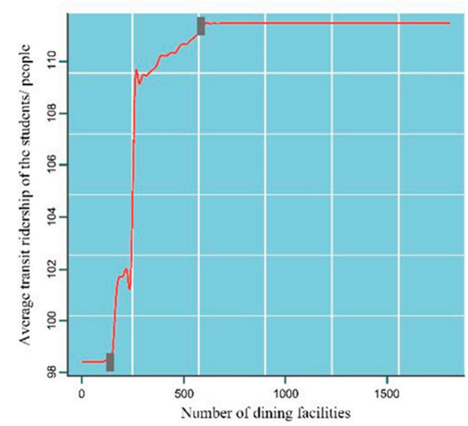
(**e**)



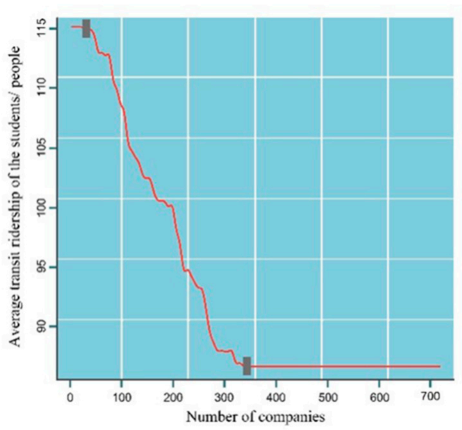
(**f**)



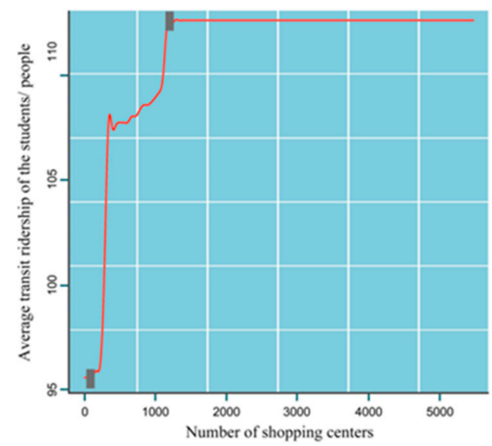
(**g**)



(**h**)

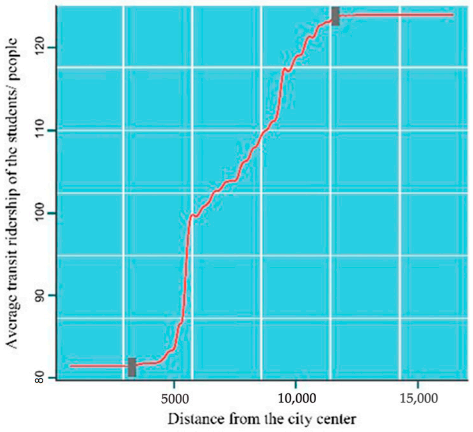


(**i**)

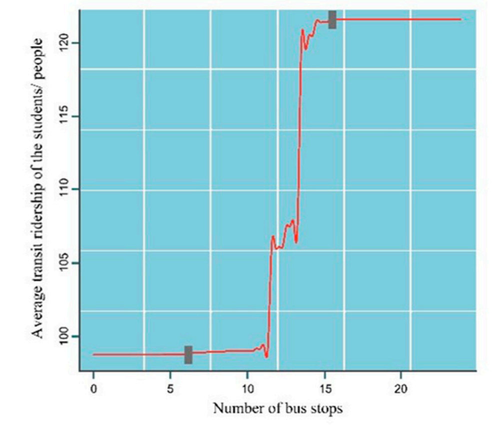


(**j**)

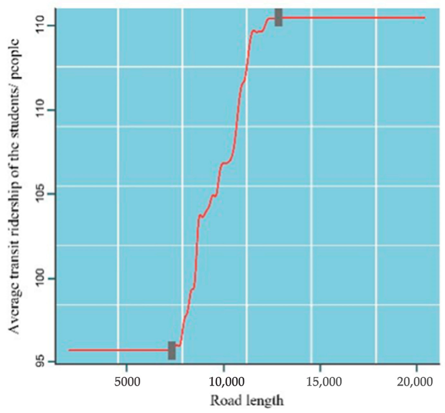
Figure 2. *Cont*.



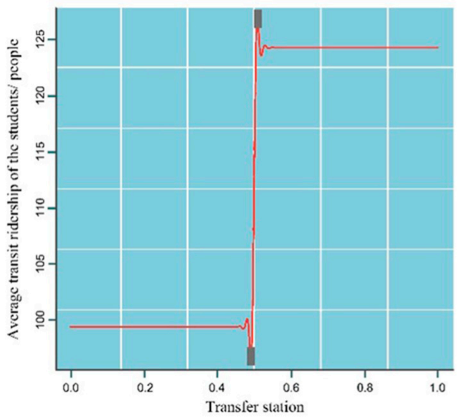
(**k**)



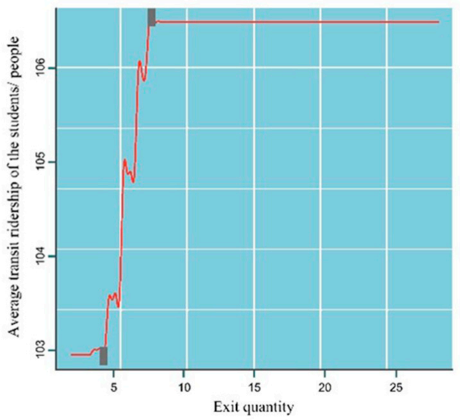
(**l**)



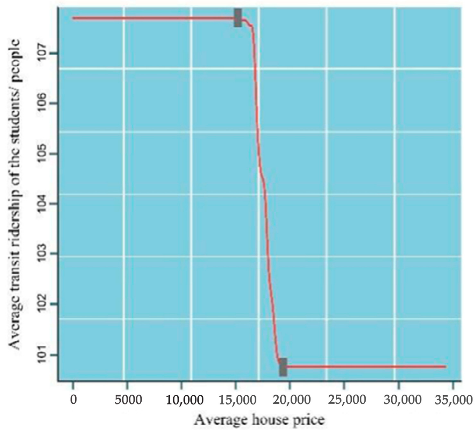
(**m**)



(**n**)



(**o**)



(**p**)

Figure 2. The non-linear influence of built environment on the school commuting-metro ridership. (**a**) Plot ratio. (**b**) Student density. (**c**) The degree of land-use mixture. (**d**) Number of intersections. (**e**) Number of parking lots. (**f**) Number of primary and secondary schools. (**g**) Number of living service facilities. (**h**) Number of dining facilities. (**i**) Number of companies. (**j**) Number of shopping centers. (**k**) Distance from the city center. (**l**) Number of bus stops. (**m**) Road length. (**n**) Transfer station. (**o**) Exit quantity. (**p**) Average house price.

Figure 2c shows the effect of the degree of land-use mixture on the average school commuting-metro ridership. In line with most existing studies, the degree of land-use mixture also has a dramatically positive effect on the average school commuting metro ridership. Further, when the degree of land-use mixture is within 0.5, the impact on the average school-commuting metro ridership is small. When the degree of land-use mixture increases further to around 0.76, the average school-commuting metro ridership gradually rises from 97 to around 103. When the degree of land-use mixture increases even further, to around 0.8, the average school-commuting metro ridership rises rapidly to around 120. However, when the degree of land-use mixture is further increased, there is no longer an impact. This indicates that, in the planning and construction of the areas around the metro stations, the degree of land-use mixture of 0.8 can attract the maximum number of primary and secondary school students to the rail transit.

Figure 2d,e shows the effect of the number of intersections and the number of car parks on the average school-commuting metro ridership. In line with the majority of studies, the number of intersections shows a positive correlation with the average school-commuting metro ridership. When the number of intersections was between 20 and 36, the average school-commuting metro ridership increased slowly, from 100 to 103, but when the number of intersections was increased to 40, the average school-commuting metro ridership increased significantly, to around 126. In line with the expected results, the number of parking lots is negatively correlated with the average school-commuting metro ridership. When the number of parking lots is between 10 and 125, the average school-commuting metro ridership decreases from 108 to 98.

Figure 2f,j shows the effect of the number of primary and secondary schools, living service facilities, dining facilities, companies, and shopping centers on the average school-commuting metro ridership. The graph of the number of primary and secondary schools shows that the number of schools is positively associated with the average school-commuting metro ridership. When the number of schools within the station area grows from 0 to 4, the average school-commuting metro ridership rises from 99 to about 110. The number of living service facilities, dining facilities, and shopping centers show similar trends. Specifically, when the number of living service facilities is between 200 and 500, the number of dining facilities is between 200 and 600, and the number of shopping centers is between 100 and 1100, and the average school commuting metro ridership increases by 25, 14, and 19, respectively. The number of companies showed a significant negative impact; when the number of companies increased from 0 to 340, the average school commuting metro ridership decreased from 115 to 84. This is because the proximity of students to their parents' firms makes it easier for parents to transport their children to school by private car.

Figure 2k–m show the effects of distance to the city center, number of bus stops, and road length on the average school-commuting metro ridership. Consistent with the expected results, the distance to the city center shows a significant positive effect on the average school-commuting metro ridership, which increases rapidly from 81 to 123 when the distance to the city center is from 5 to 12 km. The number of bus stops and the road length also shows a similar trend with the average school-commuting metro ridership. When the number of bus stops increases from 12 to 15, the average school-commuting metro ridership increases from 98 to 123; when the road length increases from 7.5 km to about 12 km, the average school-commuting metro ridership increases by about 20.

Figure 2n,o show the effects of two subway station characteristics, the number of transfer stations and stations exits, on the average school-commuting metro ridership. Consistent with the vast majority of results, both transfer stations and the number of entrances and exits show significant positive effects on the average school-commuting metro ridership, but there are differences in the magnitudes of the effects. The average school-commuting metro ridership at transfer stations is 25 persons higher than that at non-transfer stations, and the average school-commuting metro ridership only increases by about 3 persons when the number of exits at metro stations increases from 4 to 8.

Figure 2p shows the effect of house prices on average school-commuting metro ridership. Overall, house prices show a significant negative effect on the average school-commuting metro ridership, which is in line with most existing studies. Specifically, when the house price is below 15,000 CNY/m^2, its effect on the average school commuting metro ridership is minimal. When the house price increases from 15,000 CNY/m^2 to 19,000 CNY/m^2, the average school-commuting metro ridership decreases from 108 to 100. When the house price increases further, it no longer has an impact on the average school-commuting metro ridership.

5. Discussion

As large cities continue to expand rapidly in size, it leads to longer school-commuting distances. As the backbone of passenger transportation in large cities, rail transit has also gradually developed into an important mode of transportation for commuting to and from school. Our study applies the GBDT model to the subway swipe data of primary and secondary school students in Wuhan, China, and explores the non-linear effects on school travel from three aspects: the built environment, subway station characteristics, and socioeconomic attributes. The study yielded the following major findings worth exploring.

First, the relative importance of built environment factors on the impact of primary and secondary school metro ridership provides an order of priority for urban planning interventions. The GBDT-based results suggest that the number of amenities, the number of intersections, and the degree of land-use mixture are the three most important variables influencing metro school commuting. A combination of convenient amenities and higher land-use mixture with good rail accessibility is the most effective way to promote access to the metro for primary and secondary school students. This corresponds to previous studies that found that the combination of extensive amenities, entertainment facilities, and more land-use better promotes active travel of youth and increases the proportion of walking and public transportation school-commuting [40,41,43]. In addition, the number of intersections is an important evaluation indicator of street connectivity, and studies based on both developed and developing countries show that a higher number of intersections indicates better street connectivity, which significantly influences the adoption of walking and public transportation for school-commuting by primary and secondary school students [41].

Secondly, our analysis shows a non-linear association and threshold effect of all the examined factors on school-commuting metro ridership, which helps planners to determine the effective ranges of influence of planning interventions. Specifically, plot ratios only positively contribute to primary and secondary school student metro traffic between 0.4 and 1.9, with the most significant contribution to primary and secondary school student metro traffic at 1.9. Intersection densities between 36 and 40 have the most significant contribution to primary and secondary school student metro ridership, peaking at 40. Similarly, a degree of land-use mixture between 0.76 and 0.8 has the most significant effect on the promotion of metro ridership for primary and secondary school students. As the area covered by the rail transit station domain is delineated using a ten-minute walkable reach for residents, it is combined with the regulations for the planning and construction of a ten-minute living circle in the Urban Residential Area Planning and Design Code issued by China in 2018. Setting the floor area ratio within the rail transit station area at 2.0, the number of intersections at 40, and the degree of land-use mixture at 0.8 is most conducive to metro travel for primary and secondary school students. In addition, similar to the results of previous studies [21,46], school-commuting metro ridership at transfer stations is apparently higher than at non-transfer stations. Therefore, optimizing the built environment around the transfer stations is more beneficial in boosting school-commuting metro ridership. However, unlike the results of the study on other groups [16], the number of companies showed a significant negative effect on school-commuting metro ridership. This corresponds to previous research on school-commuting, where parental attitudes are a key factor influencing the children's travel to school, and the influence is greater than that of the built environment. When the parent's company is close to primary and secondary schools, they are more likely to use alternative means of school-commuting with the parent [40,41].

Third, house prices show a negative effect on metro travel for primary and secondary school students. This corresponds to previous research findings that families living in areas with higher housing prices tend to be better financially equipped, have higher rates of private car ownership, and have more options for teenagers commuting to school [39,41]. Therefore, even though Wuhan has a preferential policy for all primary and secondary school students to take the subway, housing prices still have a negative impact on primary and secondary school students' subway travel.

Meanwhile, several issues need to be further discussed. First, family factors have an important influence on primary and secondary school students' subway travel [40,41]. Although this study uses house prices to characterize family income, factors such as family size, parental travel patterns, and parental perceptions may also have an impact on whether primary and secondary school students take the metro, which will be supplemented by questionnaires in the future. In addition, the age difference and cognitive differences between secondary school students and primary school students may make significant differences in the school-commuting mode choice; however, due to the limitation of data acquisition, the original data set obtained in this study could not distinguish secondary school students from primary school students. In the future, a distinction should be made between these two age groups to determine the difference between secondary school students and elementary school students with regard to metro travel. Third, like most studies [16,29,47], this study uses cross-sectional data, which can provide evidence for associations between variables but not causality, and we encourage the use of a longitudinal design to further explore the causal associations between variables. In addition, this study was based on Wuhan, a typical high-density city, and future evidence from low and medium-density cities should be added to validate our findings.

6. Conclusions

We use the GBDT model to explore the impact of the built environment on the subway ridership of primary and secondary school students. The results of the study showed that there are significant differences in the relative importance of the built environment on the impact of primary and secondary school students' metro ridership, and there are significant non-linear and threshold effects. Based on the results of the relative importance study, we recommend that a mix of leisure and recreational facilities with more land use should be prioritized around rail stations with a focus on enhancing accessibility to rail stations. Based on the non-linear and threshold effects of the built environment on primary and secondary school students' metro ridership, we suggest that the floor area ratio around the rail station should be set at 2.0, the number of intersections at 40, and the degree of land-use mixture at 0.8, which will optimally contribute to promoting school-commuting by metro for primary and secondary school students. In addition, as interchange stations are significantly more attractive to primary and secondary school students than non-interchange stations, prioritizing improvements to the built environment around interchange stations would be more helpful in boosting primary and secondary school students' metro ridership.

Author Contributions: Conceptualization, Xuexin Yan and Jinming Yan; methodology, Xuexin Yan, Jinming Yan and Yiwen Hu; formal analysis, Xuexin Yan, Jingyi Feng and Qiuyu Wan; validation, Xuexin Yan; writing—original draft preparation, Xuexin Yan, Jinming Yan and Qiuyu Wan; writing—review and editing, Xuexin Yan, Jinming Yan and Jianjun Wang. All authors have read and agreed to the published version of the manuscript.

Funding: This research is sponsored by the China Scholarship Council (File No. 202106270077).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Some data used during the study are confidential and may only be provided with restrictions.

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

References

1. Martínez-Gómez, D.; Ruiz, J.R.; Gómez-Martínez, S.; Chillón, P.; Rey-López, J.P.; Díaz, L.E.; Castillo, R.; Veiga, O.L.; Marcos, A. Active Commuting to School and Cognitive Performance in Adolescents: The AVENA Study. *Arch. Pediatr. Adolesc. Med.* **2011**, *165*, 300–305. [CrossRef] [PubMed]
2. Must, A.; Tybor, D.J. Physical Activity and Sedentary Behavior: A Review of Longitudinal Studies of Weight and Adiposity in Youth. *Int. J. Obes.* **2005**, *29*, S84–S96. [CrossRef]
3. Chica-Olmo, J.; Rodríguez-López, C.; Chillón, P. Effect of Distance from Home to School and Spatial Dependence between Homes on Mode of Commuting to School. *J. Transp. Geogr.* **2018**, *72*, 1–12. [CrossRef]
4. Mori, N.; Armada, F.; Willcox, D.C. Walking to School in Japan and Childhood Obesity Prevention: New Lessons from an Old Policy. *Am. J. Public Health* **2012**, *102*, 2068–2073. [CrossRef] [PubMed]
5. Van Dyck, D.; De Bourdeaudhuij, I.; Cardon, G.; Deforche, B. Criterion Distances and Correlates of Active Transportation to School in Belgian Older Adolescents. *Int. J. Behav. Nutr. Phys. Act.* **2010**, *7*, 87. [CrossRef]
6. Aranda-Balboa, M.J.; Chillón, P.; Saucedo-Araujo, R.G.; Molina-García, J.; Huertas-Delgado, F.J. Children and Parental Barriers to Active Commuting to School: A Comparison Study. *Int. J. Environ. Res. Public Health* **2021**, *18*, 2504. [CrossRef]
7. Wilson, K.; Clark, A.F.; Gilliland, J.A. Understanding Child and Parent Perceptions of Barriers Influencing Children's Active School Travel. *BMC Public Health* **2018**, *18*, 1053. [CrossRef]
8. Ewing, R.; Cervero, R. Travel and the Built Environment. *J. Am. Plan. Assoc.* **2010**, *76*, 265–294. [CrossRef]
9. Guo, J.Y.; Chen, C. The Built Environment and Travel Behavior: Making the Connection. *Transportation* **2007**, *34*, 529–533. [CrossRef]
10. 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]
11. Huang, J.; Chen, S.; Xu, Q.; Chen, Y.; Hu, J. Relationship between Built Environment Characteristics of TOD and Subway Ridership: A Causal Inference and Regression Analysis of the Beijing Subway. *J. Rail Transp. Plan. Manag.* **2022**, *24*, 100341. [CrossRef]
12. 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]
13. 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]
14. Choi, J.; Lee, Y.J.; Kim, T.; Sohn, K. An Analysis of Metro Ridership at the Station-to-Station Level in Seoul. *Transportation* **2012**, *39*, 705–722. [CrossRef]
15. Yang, H.; Ruan, Z.; Li, W.; Zhu, H.; Zhao, J.; Peng, J. The Impact of Built Environment Factors on Elderly People's Mobility Characteristics by Metro System Considering Spatial Heterogeneity. *ISPRS Int. J. Geoinf.* **2022**, *11*, 315. [CrossRef]
16. Shao, Q.; Zhang, W.; Cao, X.; Yang, J.; Yin, J. Threshold and Moderating Effects of Land Use on Metro Ridership in Shenzhen: Implications for TOD Planning. *J. Transp. Geogr.* **2020**, *89*, 102878. [CrossRef]
17. 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]
18. Gan, Z.; Yang, M.; Feng, T.; Timmermans, H. Understanding Urban Mobility Patterns from a Spatiotemporal Perspective: Daily Ridership Profiles of Metro Stations. *Transportation* **2020**, *47*, 315–336. [CrossRef]
19. 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. D Transp. Environ.* **2020**, *82*, 102332. [CrossRef]
20. Cardozo, O.D.; García-Palomares, J.C.; Gutiérrez, J. Application of Geographically Weighted Regression to the Direct Forecasting of Transit Ridership at Station-Level. *Appl. Geogr.* **2012**, *34*, 548–558. [CrossRef]
21. Peng, J.; Qi, J.; Cui, C.; Yan, J.; Dai, Q.; Yang, H.; Peng, J.; Qi, J.; Cui, C.; Yan, J.; et al. Research on the Impact of the Built Environment on the Characteristics of Metropolis Rail Transit School Commuting-Take Wuhan as an Example. *Int. J. Environ. Res. Public Health Artic. Public Health* **2021**, *18*, 9885. [CrossRef]
22. Aston, L.; Currie, G.; Kamruzzaman, M.; Delbosc, A.; Teller, D. Study Design Impacts on Built Environment and Transit Use Research. *J. Transp. Geogr.* **2020**, *82*, 102625. [CrossRef]
23. Mackett, R.L. Children's Travel Behaviour and Its Health Implications. *Transp. Policy* **2013**, *26*, 66–72. [CrossRef]
24. Timperio, A.; Ball, K.; Salmon, J.; Roberts, R.; Giles-Corti, B.; Simmons, D.; Baur, L.A.; Crawford, D. Personal, Family, Social, and Environmental Correlates of Active Commuting to School. *Am. J. Prev. Med.* **2006**, *30*, 45–51. [CrossRef]
25. Mitra, R.; Buliung, R.N. Built Environment Correlates of Active School Transportation: Neighborhood and the Modifiable Areal Unit Problem. *J. Transp. Geogr.* **2012**, *20*, 51–61. [CrossRef]
26. 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]
27. Panter, J.R.; Jones, A.P.; van Sluijs, E.M.F.; Griffin, S.J. Neighborhood, Route, and School Environments and Children's Active Commuting. *Am. J. Prev. Med.* **2010**, *38*, 268–278. [CrossRef]

28. Zhao, J.; Deng, W.; Song, Y.; Zhu, Y. Analysis of Metro Ridership at Station Level and Station-to-Station Level in Nanjing: An Approach Based on Direct Demand Models. *Transportation* **2014**, *41*, 133–155. [CrossRef]
29. Su, S.; Zhao, C.; Zhou, H.; Li, B.; Kang, M. Unraveling the Relative Contribution of TOD Structural Factors to Metro Ridership: A Novel Localized Modeling Approach with Implications on Spatial Planning. *J. Transp. Geogr.* **2022**, *100*, 103308. [CrossRef]
30. Liu, Y.; Wang, S.; Xie, B. Evaluating the Effects of Public Transport Fare Policy Change Together with Built and Non-Built Environment Features on Ridership: The Case in South East Queensland, Australia. *Transp. Policy* **2019**, *76*, 78–89. [CrossRef]
31. Tu, W.; Cao, R.; Yue, Y.; Zhou, B.; Li, Q.; Li, Q. Spatial Variations in Urban Public Ridership Derived from GPS Trajectories and Smart Card Data. *J. Transp. Geogr.* **2018**, *69*, 45–57. [CrossRef]
32. Zhao, P. The Impact of the Built Environment on Individual Workers' Commuting Behavior in Beijing. *Int. J. Sustain. Transp.* **2013**, *7*, 389–415. [CrossRef]
33. Liu, B.; Xu, Y.; Guo, S.; Yu, M.; Lin, Z.; Yang, H. Examining the Nonlinear Impacts of Origin-Destination Built Environment on Metro Ridership at Station-to-Station Level. *ISPRS Int. J. Geoinf.* **2023**, *12*, 59. [CrossRef]
34. Xiao, L.; Lo, S.; Liu, J.; Zhou, J.; Li, Q. Nonlinear and Synergistic Effects of TOD on Urban Vibrancy: Applying Local Explanations for Gradient Boosting Decision Tree. *Sustain. Cities. Soc.* **2021**, *72*, 103063. [CrossRef]
35. Saha, D.; Alluri, P.; Gan, A. Prioritizing Highway Safety Manual's Crash Prediction Variables Using Boosted Regression Trees. *Accid. Anal. Prev.* **2015**, *79*, 133–144. [CrossRef] [PubMed]
36. Friedman, J.H. Greedy Function Approximation: A Gradient Boosting Machine. *Ann. Statist.* **2001**, *29*, 1189–1232. [CrossRef]
37. Ding, C.; Cao, X.; Næss, P. Applying Gradient Boosting Decision Trees to Examine Non-Linear Effects of the Built Environment on Driving Distance in Oslo. *Transp. Res. Part A Policy Pract.* **2018**, *110*, 107–117. [CrossRef]
38. Zhang, Y.; Haghani, A. A Gradient Boosting Method to Improve Travel Time Prediction. *Transp. Res. Part C Emerg. Technol.* **2015**, *58*, 308–324. [CrossRef]
39. Corral-Abós, A.; Aibar, A.; Estrada-Tenorio, S.; Julián, J.A.; Ibor, E.; Zaragoza, J. Implications of School Type for Active Commuting to School in Primary Education Students. *Travel Behav. Soc.* **2021**, *24*, 143–151. [CrossRef]
40. Ding, D.; Sallis, J.F.; Kerr, J.; Lee, S.; Rosenberg, D.E. Neighborhood Environment and Physical Activity among Youth: A Review. *Am. J. Prev. Med.* **2011**, *41*, 442–455. [CrossRef]
41. Wang, X.; Liu, Y.; Zhu, C.; Yao, Y.; Helbich, M. Associations between the Streetscape Built Environment and Walking to School among Primary Schoolchildren in Beijing, China. *J. Transp. Geogr.* **2022**, *99*, 103303. [CrossRef]
42. Ikeda, E.; Mavoa, S.; Cavadino, A.; Carroll, P.; Hinckson, E.; Witten, K.; Smith, M. Keeping Kids Safe for Active Travel to School: A Mixed Method Examination of School Policies and Practices and Children's School Travel Behaviour. *Travel Behav. Soc.* **2020**, *21*, 57–68. [CrossRef]
43. Giles-Corti, B.; Wood, G.; Pikora, T.; Learnihan, V.; Bulsara, M.; van Niel, K.; Timperio, A.; McCormack, G.; Villanueva, K. School Site and the Potential to Walk to School: The Impact of Street Connectivity and Traffic Exposure in School Neighborhoods. *Health Place* **2011**, *17*, 545–550. [CrossRef] [PubMed]
44. Mitra, R.; Buliung, R.N. The Influence of Neighborhood Environment and Household Travel Interactions on School Travel Behavior: An Exploration Using Geographically-Weighted Models. *J. Transp. Geogr.* **2014**, *36*, 69–78. [CrossRef]
45. Sohn, K.; Shim, H. Factors Generating Boardings at Metro Stations in the Seoul Metropolitan Area. *Cities* **2010**, *27*, 358–368. [CrossRef]
46. Jun, M.J.; Choi, K.; Jeong, J.E.; Kwon, K.H.; Kim, H.J. Land Use Characteristics of Subway Catchment Areas and Their Influence on Subway Ridership in Seoul. *J. Transp. Geogr.* **2015**, *48*, 30–40. [CrossRef]
47. Zhao, J.; Deng, W.; Song, Y.; Zhu, Y. What Influences Metro Station Ridership in China? Insights from Nanjing. *Cities* **2013**, *35*, 114–124. [CrossRef]