

Research Article

Research on Nonlinear Associations and Interactions for Short-Distance Travel Mode Choice of Car Users

Mingwei He , Lijuan Pu , Yang Liu , Zhuangbin Shi , Chengfeng He ,
and Jiayou Lei 

Faculty of Transportation Engineering, Kunming University of Science & Technology, Jingming Road 727, Kunming 650500, China

Correspondence should be addressed to Yang Liu; kmluiyang@outlook.com

Received 11 August 2021; Revised 21 March 2022; Accepted 26 April 2022; Published 27 May 2022

Academic Editor: Dilum Dissanayake

Copyright © 2022 Mingwei He et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Encouraging car users who travel short distances to shift from car mode to active travel modes can effectively alleviate urban traffic congestion and reduce carbon emissions. However, few studies have examined the determinants of the travel mode choice of short-distance car users and ignored the nonlinear associations and interactions between variables. This paper conducts a questionnaire survey to investigate the short-distance travel mode choice of car users who travel less than 4 km in a specific city. A random forest (RF) model is applied to examine the influence of key variables on these three travel mode choices of short-distance car users and to explore the nonlinear associations and interactions of the variables. Compared with multinomial logic model, the results of RF show that significant threshold effects exist in the relationship between the car user's travel mode choice and the selected explanatory variables, mainly travel distance, road network density, distance to CBD, and number of bus stops. In particular, 1.2 km is a critical turning point for car and active travel mode choice, before which car users prefer to travel by walking and cycling and after which there is a significant increase in the car use probability. When the road network density was between 2.5 km/km² and 6.5 km/km², the proportion of car users who chose cycling showed an increasing trend, while car use showed a decreasing trend. These findings can provide a solid basis for planning managers to develop policy measures to encourage environmentally sustainable travel.

1. Introduction

In recent years, excessive and unreasonable car use among urban residents has resulted in a series of negative impacts on urban sustainable development, such as traffic congestion and environmental pollution [1]. This set of pressing issues highlights the need to reduce the car use of car owners by shifting their travel to a more low-carbon mode. Many studies have confirmed that car users are more likely to shift from car use to walking and cycling when they travel short distances [2, 3]. That is, under certain conditions, car users who travel a short-distance are likely to walk or cycle. To grasp the conditions of this change, it is necessary to understand the travel choice behavior of car users when choosing between cars and other active travel modes when they travel a short distance. At present, there is no unified definition of short-distance travel across research areas and

research contents; however, it is usually limited to within 5 km [4, 5].

Existing research on travel mode choice focuses mainly on the determinants of mode choice [6–8]. In contrast, few studies have examined the determinants of travel mode choice of short-distance car users, while ignoring the nonlinear associations and interactions between variables. The nonlinear association means that variables have differential threshold effects, while interaction indicates the joint influence of multiple variables. Many studies have confirmed that nonlinear associations and interactions are common in travel behavior [9–12]. For instance, Jixiang et al. [11] found that the built environment at the origin and destination of the trip has a threshold effect on an individual's active travel. For instance, for work trips, active travel was favored when the land use mix at the origin was within 0.6, or the land use mix at the end was within 0.45.

The land use mix indicates the number and proportion of all the land use types (e.g., municipal administration, education, and public open space), which is calculated based on the formula of land use entropy. The higher the entropy value, the greater the corresponding land use mix, that is, having a more complex land use type. Wang and Ozbilen [10] showed that there is an interaction between built environment and telecommuting duration, which can have a synergistic effect on the travel time share of active travel. Therefore, if we ignore the influence of nonlinear association and interaction in travel behavior, it is difficult to develop a refined strategy or policies to guide individual traveling with low-carbon modes.

Recently, some scholars have become aware of this problem and have begun to consider the nonlinear associations between variables when modeling individuals' travel behaviors [13–16]. They improved the existing research in two ways. On the one hand, to examine the influence proportion of the key influence variables, the relative importance of influencing factors is quantified. On the other hand, to illustrate the threshold effect of influencing factors, nonlinear associations between variables are explored. For instance, Tu et al. [13] used a gradient boosting decision tree to model the ridesplitting ratio. By ranking the relative importance of built environment factors, three key variables, the distance to the city center, land use diversity, and road density, are identified. Cheng et al. [14] used survey data from elderly people in Nanjing (China) and explored the nonlinear associations and threshold effect between the built environment and walking time through the random forest method. The results show that the land use mix and population density affect the walking time of the elderly only within a certain range. However, none of them have explored the possible nonlinear associations between individual travel mode choices and the corresponding influencing factors. In addition, the interaction between variables is rarely discussed, although the existence of this interaction has been confirmed by some studies.

To fill the above research gap, we need to explore both the nonlinear associations and interactions in the short-distance travel mode choices of car users. A random forest (RF) model is adopted for modeling in this paper for the following reasons: (1) compared to multinomial logit models (MNL), random forest models as a machine learning method have the advantage of capturing nonlinear associations and interactions among determinants; and (2) compared with other machine learning methods, random forest models have superior predictive power in studying travel mode choice [17–19].

Taking Kunming, China, as a case study, we collected data on short-distance trips of car users in 2019 through questionnaires as well as corresponding built environment data and used random forest for modeling. The research involves the following three aspects: (1) identify the key factors that affect the short-distance travel mode choice of car users; (2) explore the complex relationship between these factors; and (3) propose strategies to guide the transfer of car users' travel modes. The results can provide important planning insights.

The remainder of the paper is organized as follows. Section 2 provides a literature review of research on individual travel mode choice. Section 3 outlines the study area and data sources and provides a preliminary statistical analysis of the data. Section 4 describes the random forest modeling approach. Section 5 provides the results of the study, and the final section provides further discussion and draws final conclusions.

2. Literature Review

At present, research on the choice of individual travel mode is gradually increasing. These studies mainly focused on the factors that affect the choice of travel mode, which can provide a valuable reference for this study on the selection of influencing factors. In addition, the existing research on the nonlinear associations of travel behavior also provides important enlightenment for our study.

A large amount of literature shows that economic demographic attributes, travel characteristic attributes, psychological attributes, and built environment are important variables that affect individual travel mode decisions. In terms of economic demographic attributes, age, sex, occupation, and income are important factors. For instance, through a survey of short-distance travel in Queensland, Cole et al. [20] found that women, middle-aged and older people, and those without a job were more likely to use cars for travel. In contrast to those findings, Dèdelè et al. [21] pointed out that car users are more likely to be young, male, and employed. This difference may be related to other characteristics of the individuals studied; for example, Cheng et al. [22] found that, among low-income commuters, women were more likely to choose public transportation options. The above findings suggest that interactions between variables may exist. The travel characteristic attributes include mainly travel purpose and travel distance. Among them, travel distance has a more significant impact on the choice of travel mode. In general, short-distance travel is always associated with walking and cycling [23], but short-distance travel for commuting, carrying heavy objects, and multiple travel destinations may increase the use of cars [4]. In terms of the built environment, Cao et al. [24] indicate that measures such as upgrading public transportation services and enhancing accessibility can reduce the use of cars for short-distance trips. Indicators such as employment density, population density, street connectivity, and land use mix were shown to impact travel mode choice decisions [25, 26]. In addition to the abovementioned objective attributes, subjective psychological factors such as attitude are also considered factors that affect the choice of travel mode [27].

The above research shows that the choice of travel mode may be affected by multiple factors. In this context, it is particularly important to identify the key variables of travel mode choice [19, 28, 29]. From the perspective of planners, it is necessary to grasp the priority levels of determinants to formulate the right interventions [30, 31]. Recently, Cheng et al. [17] quantified the relative importance of the effect of explanatory variables on the choice of travel mode through

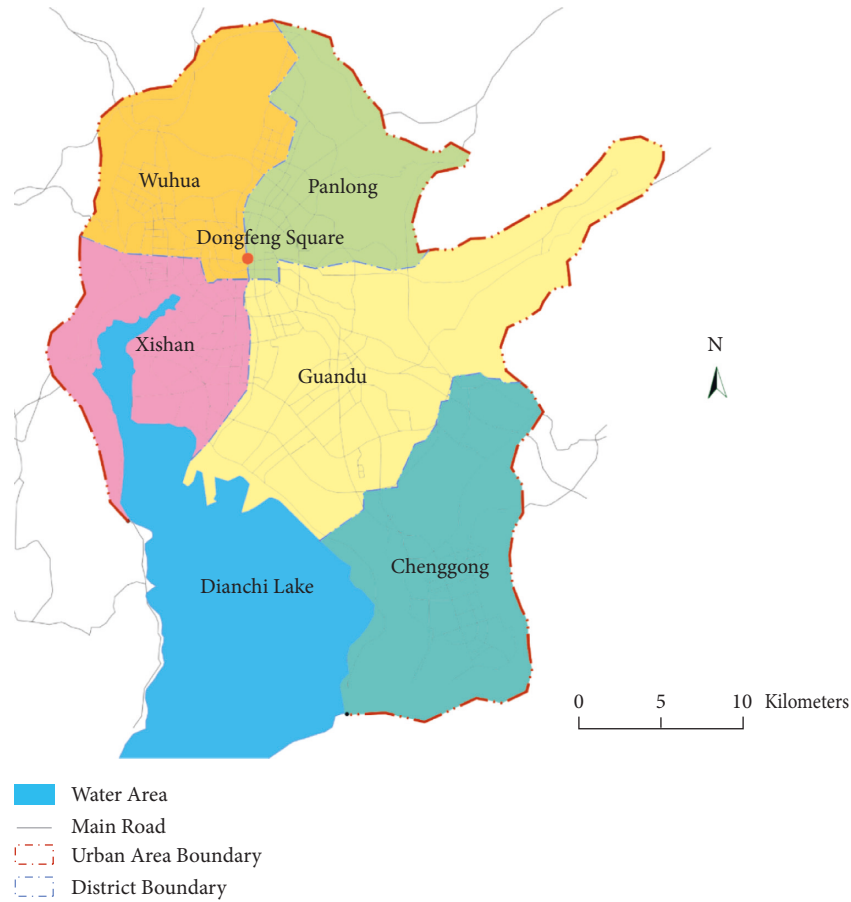


FIGURE 1: Map of the main urban area of Kunming.

the random forest method. They found that, overall, the built environment is more important than other attributes, and the land use mix is the most important variable. The key variable may vary by research area and research object. Therefore, it is necessary to explore the key variables in the short-distance travel mode choice of car users. On the other hand, in recent years, an increasing number of studies have found nonlinear associations between travel behavior and its determinants [11, 32–35]. For instance, Jixiang et al. [11] found that when population density reached 3,000 per square kilometer, the additional density had a negligible impact on driving reduction. This provides planners with a more detailed reference basis [36]. However, there is still a gap in the exploration of the nonlinear associations between the short-distance travel mode choice of car users and its key variables. In addition, few studies have pointed out that there are interactions between determinants, and the resulting synergistic effects will affect individual travel behavior [10, 12, 37]. In fact, travel choice behavior is affected not only by variables. Traffic demand management strategies and land use policies can produce synergistic effects [37]. Therefore, the interaction between key variables in travel mode choice also needs to be further explored.

Therefore, we will study the key determinants of car users' travel mode choice in the case of short-distance travel and use the random forest model to explore the nonlinear

associations and interaction between the variables. This research will contribute to the research on travel mode choice in two ways. On the one hand, we research the short-distance travel choice of car users. On the other hand, we explore the complex nonlinear associations between travel mode choice and its determinants and examine whether there is any interaction between determinants. The findings of our study provide planning managers with the priority and effective range of key variables, which will help them allocate resources effectively.

3. Data

3.1. Study Area and Short Distance Definition. Kunming is the capital of Yunnan Province, located in southwestern China, and is an important central city in the western region. The urban area of Kunming consists of five administrative districts (Wuhua District, Panlong District, Xishan District, Guandu District, and Chenggong District), as shown in Figure 1. The total area is 21,473 km², and the built-up area is 483.52 km². The total registered population of the city was 5.83 million at the end of 2020, and the number of cars was approximately 2,970,100, an increase of 5.2% compared to 2019 [38]. The increase in the use of cars has led to traffic congestion and environmental pollution [39].

For the definition of short-distance travel, this study will determine the specific distance value by considering the study content and the daily travel situation of residents in the study area. First, we mainly study the short-distance travel mode choice of car users. As shown in Section 2, walking and cycling are important transportation modes for short-distance trips. Second, the results of the 2016 Kunming resident travel survey show that the average distance of residents' cycling is 2.9 km, and 75.6% of the cycling distance and more than 99% of the walking distance are concentrated within 4 km [39]. Therefore, to effectively examine car users, the 4 km threshold is selected in this paper as the definition of short-distance travel in the study.

3.2. Survey and Data Collection. We surveyed a sample of car users' short-distance trips in Kunming in May 2019. The survey took the form of a questionnaire that considered factors that may influence travel mode choice, including economic demographic attributes and short-distance travel attributes. Following the research of Jakobsson et al. [40], the car user in this study is defined as someone who has at least one car in the household and has used car for travel once or more. Based on this definition of car users, we first asked respondents about their household car ownership and car use experience before doing a formal survey and only survey respondents who meet the criteria. We asked respondents to recall their past trips of less than 4 km (excluding hanging out and running), and the trip characteristics included trip purpose, trip mode, and trip distance. To improve the quality of the questionnaire, surveys were conducted in areas such as parking lots, car sales stores, and car washes within the study area, and at the same time, respondents who completed the questionnaires were rewarded with 10 RMB (approximately 1.55 dollars). Finally, a total of 1,835 travel samples were collected. After eliminating invalid samples (for example, unclear home address, incomplete travel information, etc.), 1,578 short-distance travel samples were obtained, with an 86% effective sample rate.

To explore the impact of the built environment, we used ArcGIS software to extract the built environment data for a 500 m area around the home address of the car user. There were three aspects: land use, transportation facility layout, and accessibility. The land use mix is calculated according to the formula of land use entropy $-\sum_i (P_i \ln(P_i))$, with P_i denoting the proportion of land in category i [17]. Road network density is expressed by the ratio of the total length of all roads in the region to the total area of the region; intersection density is measured by the number of intersections in the region. The transportation facilities layout includes the number of bus lines and bus stops; accessibility includes the distance from the residential area to the nearest bus stop, the distance to the nearest subway station, and the distance to the central business district.

3.3. Descriptive Statistics. In this study, short-distance travel mode choice was the dependent variable, which mainly included five travel modes: car, cycling, walking, bus, and subway. The share of each travel mode was 51.29%, 26.29%,

13.33%, 6.13%, and 2.96%, respectively. The proportion of bus and subway trips in the sample was low (less than 10.00% in total). Therefore, we focused on only three modes of travel: car, cycling, and walking. There were 1,425 short-distance trips retained.

The travel share of the three modes is shown in Table 1. We found that car trips accounted for the highest proportion (53.05%), followed by cycling and walking. This shows that, even in the case of short-distance travel, car users still like to travel by car, so it is necessary to guide car users to shift to the active mode of transportation.

Table 2 presents a descriptive statistical analysis of the explanatory variables. Economic demographic attributes mainly include age, sex, education level, occupation, and annual household income. In China, citizens under 18 and over 70 years old are not allowed to drive cars, so the age range for the survey was 18–70 years old. The ages were divided into 5 categories. A total of 69.68% of the respondents were 18–40 years old. Car user sex is a binary variable, with 1 assigned to males and 0 assigned to females. The mean of 0.62 indicated that the proportion of male car users in the selected sample was higher than that of female car users. Educational attainment was divided into three levels: high school or junior college and below, college or university, and graduate and above. The mean value was 2.10, with a high percentage (81.33%) of respondents having completed higher education. Similarly, annual household income was divided into 3 categories. The mean value was 2.16, indicating that the majority (75.16%) of respondents had an annual household income of 50,000 RMB (approximately 7,730 dollars) or more. In addition, occupations were divided into four categories. More than half (60.00%) of the respondents were employed, where the employed mainly consisted of corporate employees, institutional employees, or civil servants.

In terms of the built environment variables, the average road network density was 6.64 km/km², which approximately reflects the current situation of road network density in Kunming urban administrative districts (6.78 km/km² in the 2020 annual road network density test report of major cities in China [41]). The average intersection density was 3.30 count/km². The average land use mix was 0.57. The central business district of Kunming is Dongfeng Square, which is marked in Figure 1. According to the statistics, the average Euclidean distance from respondents' residence to Dongfeng Square was 9.89 km. The average numbers of bus stops and bus routes were 11.22 and 13.67, respectively. The average distances to the nearest bus station and subway station were 0.2 km and 1.43 km, respectively. These values reflect the supply and accessibility of public transportation near the respondent's place of residence.

Travel characteristics mainly include travel purpose and travel distance. Travel purpose was divided into four categories: commute, entertainment, official business, and other. Commuting (36.42%) and entertainment (35.93%) were the main purposes of travel reported by the respondents. Travel distance has a continuous value, and the average travel distance was 2.73 km, with a standard deviation of 1.09. Table 3 shows the trip sharing rates for various modes by

TABLE 1: Travel mode choice distribution.

Mode choice category	Frequency	Percentage
Automobile	756	53.05
Bicycle	445	31.23
Walk	224	15.72
Total	1425	100.00

TABLE 2: Explanatory variable descriptive statistics.

Variables	Description	Mean	SD
<i>Demographic characteristics</i>			
Age	Belonging to the age. {1 = 18–30, 2 = 31–40, 3 = 41–50, 4 = 51–60, 5 = 61–70}	2.07	1.16
Sex	Male or female. {0 = female, 1 = male}	0.62	0.48
Education	Education level. {1 = low, 2 = medium, 3 = highly}	2.10	0.68
Occupation	Type of occupation. {1 = self-employed, 2 = employed, 3 = retirement, 4 = other}	2.28	0.91
Income	Annual family income (10,000 RMB). {1 = below 5, 2 = 5–20, 3 = above 20}	2.10	0.62
Car use	Car domination rights, it also means car usage in daily life. {1 = totally at my disposal, 2 = used primarily by me, 3 = used primarily by family}	1.85	0.79
Travel cost	Daily acceptable transportation expenses (RMB). {1 = below 10, 2 = 10–40, 3 = above 40}	1.94	0.63
<i>Built environment</i>			
Road density	Road network density (km/km ²).	6.64	2.51
Intersection density	Number of road intersections in the area (count/km ²).	3.30	3.01
Land use mix	Land use diversity index, calculated by the land use entropy formula.	0.57	0.15
CBD distance	Distance from residence to central business district (km).	9.89	7.65
Bus stop	Number of bus stops in the area (count).	11.22	5.95
Bus route	Number of bus lines in the area (count).	13.76	10.14
Bus distance	Distance from residence to nearest bus stop (km).	0.20	0.10
Subway distance	Distance from residence to the nearest subway station (km).	1.43	1.29
<i>Travel information</i>			
Travel distance	Travel distance (km).	2.73	1.09
Trip purpose	Trip purpose. {1 = commute, 2 = entertainment, 3 = official business, 4 = other}	2.05	1.03

TABLE 3: The share of various travel modes at different travel distances.

Travel distance (km)	All modes of travel sharing rate (%)		
	Automobile	Bicycle	Walk
0 < travel distance ≤ 1	22.83	38.81	38.36
1 < travel distance ≤ 2	36.21	37.61	26.18
2 < travel distance ≤ 3	63.23	30.16	6.61
3 < travel distance ≤ 4	71.86	23.67	4.77

distance. This table illustrates that the share of cars increases with trip distance, while the relationship between active travel modes (cycling and walking) and trip distance is reversed.

In addition, as shown in Figure 2, household car dominance and acceptable transportation costs can impact travel mode choice. Figure 2(a) shows that when the family car is mainly controlled by oneself, the car becomes the main mode of travel. In contrast, when the family car is mainly used by family members, cycling becomes a good choice. This indicates that car users who can use the car at any time are likely to travel by car more often. Figure 2(b) shows that

the average acceptable daily cost of transportation also influences travel decisions to some extent. When the acceptable cost was below 10 RMB, the respondents were more likely to use bicycles to travel.

4. Methodology

4.1. Random Forest Method. To explore the nonlinear associations and interactions in the short-distance travel mode selection of car users, we adopt the random forest (RF) model for modeling. RF is an algorithm that integrates multiple trees through the idea of integrated learning. Its basic unit is a decision tree, where each tree is formed by randomly selecting observations (rows) and variables (columns). For the classification model, the final classification tree results are aggregated according to the voting method. This method helps optimize the fitting and prediction of the model [42]. RF has two important parameters [30]: (1) the number of variables preselected by the tree nodes (m) and (2) the number of trees in the RF (n). When using the bootstrap method to randomly sample the dataset, the probability p that each sample is not selected is $(1 - (1/n))^n$,

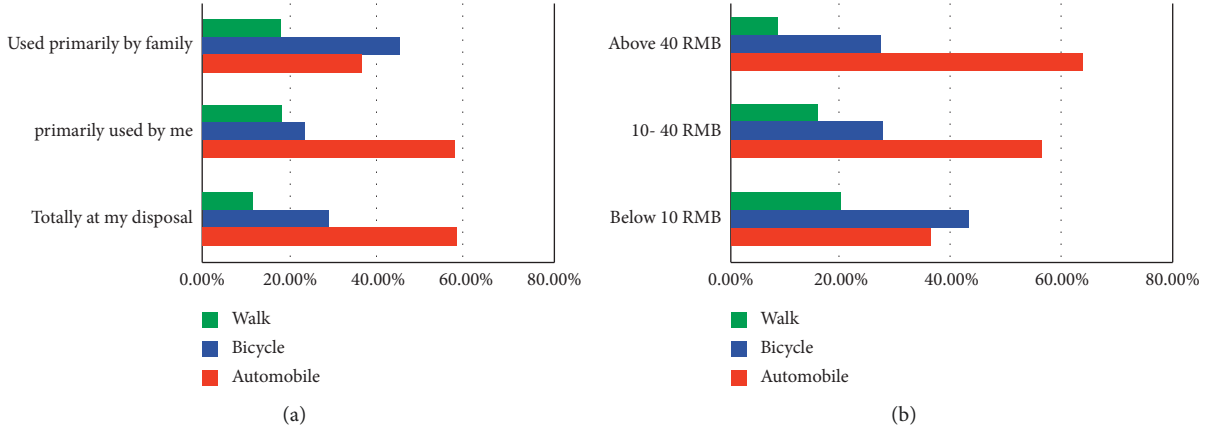


FIGURE 2: The choice of travel mode by situation. (a) Car domination rights. (b) Acceptable transportation costs.

indicating that some samples in the dataset will not appear in the training sample set, that is, out of bag (OOB) [42]. In order to ensure the stability of the OOB error in the modeling process, the model parameters need to be calibrated to make the model more robust and obtain better model performance. This estimation can be used to replace the test set for error estimation (OOB estimation). RFs can quantify the relative importance of explanatory variables in predicting outcomes, and partial dependency plots visualize the relationship between the target and explanatory variables.

4.2. Relative Importance. Mean Decrease Gini (MDG) estimation can be used to measure the relative importance of the explanatory variables [43]. It is measured by the decrease in Gini impurity due to the variable v . According to formulas (1) and (2), the reduction in Gini impurity for each characteristic variable is calculated. Finally, normalize all the obtained importance scores. In this method, the larger the value, the more important the explanatory variable.

$$\text{MDG}(v) = \sum_{i=1}^n \sum_{m \in M} (GI_m - GI_l - GI_r). \quad (1)$$

$$GI_j = 1 - \sum_{k=1}^K P_{jk}^2, \quad (2)$$

where $\text{MDG}(v)$ is the relative importance score of the variable; n is the number of decision trees; m is the node; M is the set of node m ; GI is the Gini impurity index; k represents the total number of classes of the target variable; P_{jk} represents the conditional probability that the target variable is the k th class in node j ; and GI_l and GI_r are the Gini indices of the two new nodes after branching.

4.3. Partial Dependence Plots. Partial dependence (PD) can indicate the complex relationship between explanatory variables and travel mode choices. Partial dependence plots (PDPs) show the marginal effect of a feature on the predicted

outcome of a previously fitted model, with the prediction function fixed at a few values of the selected feature and averaged over the other features [41]. For example, when a feature value increases to a certain extent, the average predicted value decreases, and the prediction tends to be other categories. PDPs can reflect the influence of variables on the prediction results in a visual way. The abscissa represents the value range of the explanatory variable, the ordinate represents the probability of choosing a certain travel mode, and the shape of the curve reflects the magnitude and direction of the impact. The PD function is defined as follows:

$$\hat{f}_{(X_s)}(X_s) = \frac{1}{m} \sum_{j=1}^m \hat{f}(X_s, X_c^{(j)}), \quad (3)$$

where X_s is the target variable for plotting the partial correlation function, $X_c^{(j)}$ is the feature value in the dataset other than the target variable, m is the number of instances in the dataset, and $\hat{f}_{(X_s)}(X_s)$ is the prediction of the target variable.

5. Results

5.1. Random Forest Model Calibration. We use the “randomForest” toolkit of the R language to implement the RF algorithm. The entire sample set is divided into a training set and a test set at a ratio of 8 : 2, referring to Friedman [44]. To determine the optimal splitting value of model training, we set the loop function with m taking values from 1 to 17 and tested n according to the default value of 500 to obtain the OOB error corresponding to each m value. The test results show that the OOB error is lowest when $m = 7$; thus, 7 is adopted as the value of parameter m . Accordingly, an experimental model with n values from 100 to 800 in increments of 100 is established. The test results are shown in Figure 3. Four OOB error curves are generated from the three-classification model. Since the reduction in the OOB error is close to zero, when $n = 600$, the maximum number is determined to be 600. The AUC value corresponding to the ROC curve of the RF model was 0.924, and the prediction accuracy was 82.75%.

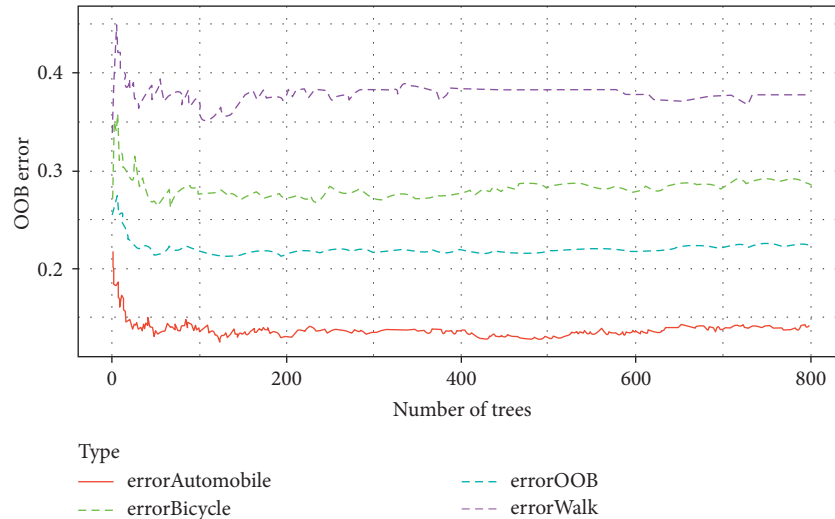


FIGURE 3: Evaluation of the number of trees.

TABLE 4: Relative importance of determinants.

Categories	Variable	Relative importance (%)	Rank	Total (%)
Travel information	Travel distance	14.94	1	21.33
	Trip purpose	6.40	9	
Built environment	CBD distance	8.53	2	57.75
	Road density	8.41	3	
	Bus distance	7.78	4	
	Subway distance	7.59	5	
	Land use mix	7.50	6	
	Bus route	7.26	7	
	Bus stop	6.63	8	
	Intersection density	4.04	10	
Demographic characteristics	Travel cost	3.76	11	20.92
	Age	3.73	12	
	Car use	3.55	13	
	Occupation	3.26	14	
	Education	2.45	15	
	Income	2.21	16	
	Sex	1.96	17	
Total relative importance				100.00

5.2. Relative Importance Analysis of Explanatory Variables

5.2.1. Relative Importance of Determinants of Travel Mode Choice. Table 4 shows the relative importance of economic demographic attribute variables, built environment variables, and travel attribute variables on the impact of travel mode choice of car users who travel short distances. The relative importance is quantified by the MDG value, and the sum of the relative impacts of all explanatory variables is 100%.

Table 4 shows that the built environment variable has the largest total relative influence among the three influencing factor categories, accounting for 57.75%, which is consistent with previous study results [11, 14, 31]. This shows that the built environment has the greatest influence on the short-distance travel mode choice of car users. Travel attribute variables are more important than economic demographic attribute variables. The sum of the

relative importance of the former two explanatory variables exceeds the sum of the relative importance of the latter's seven explanatory variables, and the travel distance is absolutely leading. Specifically, the relative importance of travel distance accounted for 14.94%, ranking 1st. This shows that even in short trips, travel distance can influence travel mode decisions. Therefore, it is necessary to explore the distance thresholds for different travel mode choices within 4 km. In terms of the built environment, the relative importance of the distance to the city center (Dongfeng Square) was the largest (8.53%), followed by road network density (8.41%), the distance to the nearest bus station (7.78%), the distance to the nearest metro station (7.59%), and the land use mix (7.50%). These are the five key variables that need to be the focus of researchers and planners. In terms of economic demographic attributes, acceptable travel costs, age, and car domination rights had a relatively

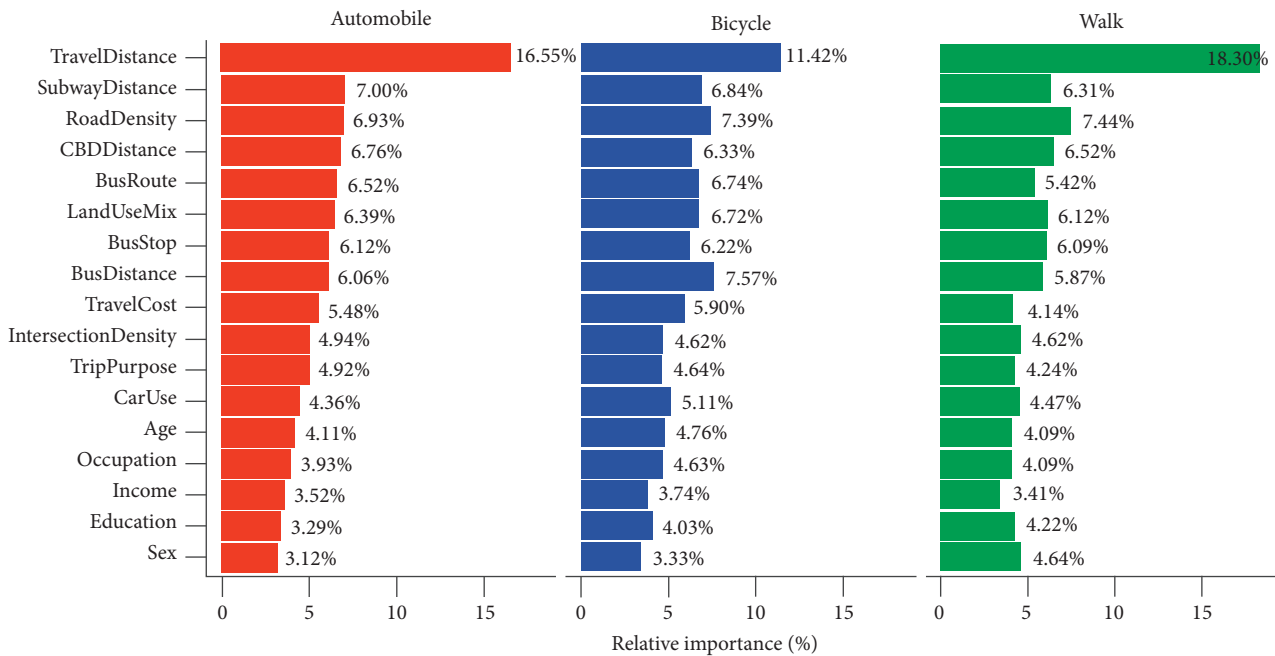


FIGURE 4: Importance of characteristic variables.

strong influence on the travel mode choice of short-distance car users, with influences of 3.76%, 3.73%, and 3.55%, respectively.

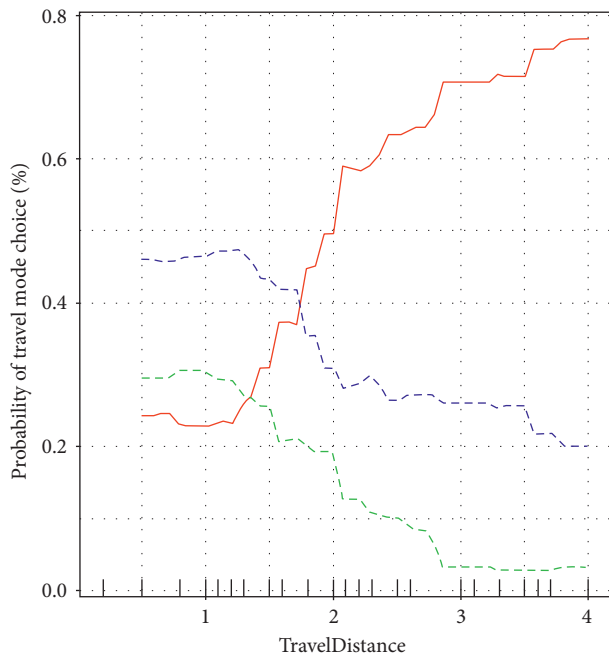
5.2.2. Differences in the Relative Importance of Determinants in the Three Modes. Figure 4 shows that the relative importance of the explanatory variables was not entirely consistent across car use, cycling, and walking. Taking the car as the reference group in the model, the ranking characteristics of the relative importance for each explanatory variable were observed. For cycling, the relative importance of the distance to the nearest bus stop is more prominent. The possible reason is that bicycles are likely to be used as a feeder mode to access/egress transit, taking on the role of transfer connections with public transit [45]. For walking, the concern is not only road density, but also the distance from the residence to the nearest metro station and the central business district. This indicates that good environmental supply and the accessibility of nearby public transportation facilities can promote the adoption of walking travel [46].

5.3. Nonlinear Association Analysis of Determinants and Travel Mode Choice. To investigate the nonlinear associations of the short-distance travel mode choice of car users, we visualized the marginal effects of the top ten key variables in terms of relative importance for predicting travel mode choice through partial dependency plots. Among them, travel distance, road density, distance to CBD, and number of bus stops have significant nonlinear effects, as shown in Figure 5.

Figure 5(a) shows the nonlinear associations of travel distance on the choice of travel mode. As the travel distance increases, the probability of car users choosing cars gradually

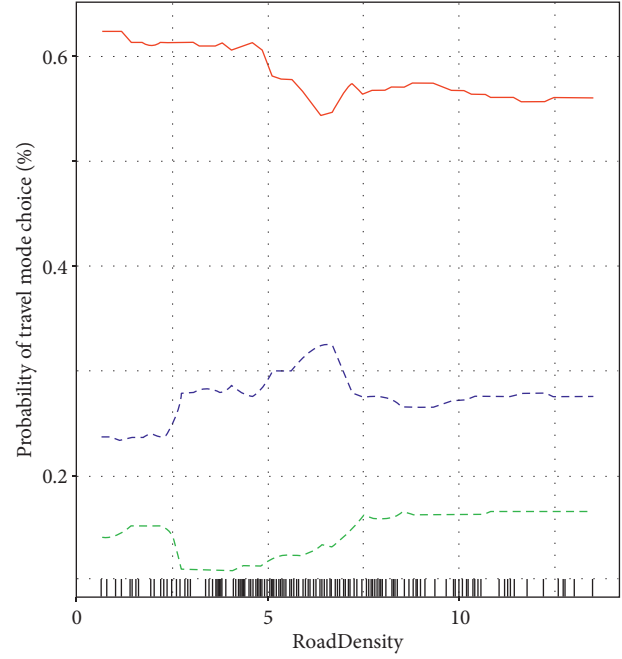
increases, while the probability of choosing walking and cycling shows a decreasing trend. This indicates that the willingness of car users to choose a travel mode varies significantly with the travel distance. This finding is consistent with that of Jonas, who pointed out that an increase in travel distance has a negative effect on walking and cycling and a positive effect on driving [47]. However, these associations are nonlinear, as shown in the PDP. Specifically, the presentation of modal splitting varies with distance threshold. A distance of 1.2 km seems to be a turning point, before which we find that car users prefer to travel by walking and cycling and after which a significant increase in the probability of car use is observable. Similar findings can be found in Neves and Brand [3], which pointed out that travelers are likely to switch from walking to other travel modes when their walking distance exceeds 1.6 km. That is, car users are most likely to travel by active travel modes when the travel distance is less than 1.2 km. Cycling has a competitive advantage over walking. When the distance travel exceeds 3 km, car users will hardly consider using walking to travel.

Figure 5(b) shows that there are nonlinear associations between road network density and travel mode choice. We found that an increase in road network density within a certain range can reduce car use while increasing the share of walking and cycling trips. In more detail, when the road network density was between 2.5 km/km² and 6.5 km/km², the proportion of those who chose cycling was on an increasing trend, while car use was on a decreasing trend. The probability of walking choice increased when the road network density was between 5 km/km² and 7.5 km/km². This suggests that, under reasonable road network density conditions, there will be an increased tendency to choose alternative modes of travel such as walking and cycling. In



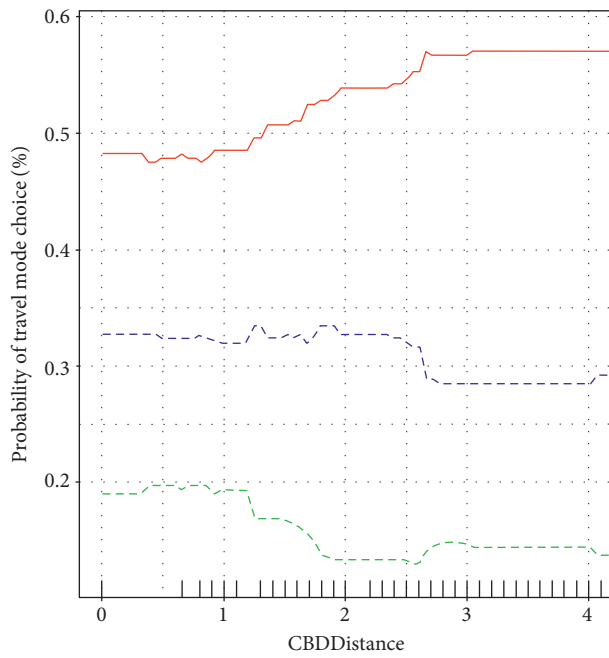
TravelMode
 — Automobile
 - - Bicycle
 - - Walk

(a)



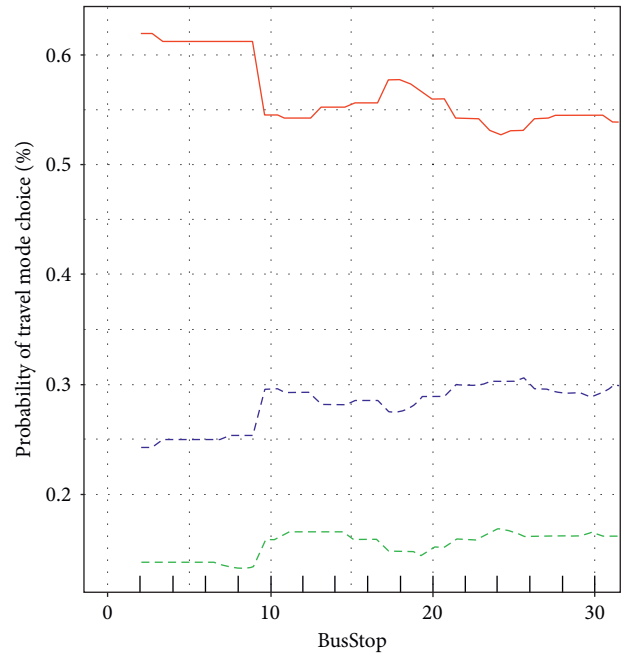
TravelMode
 — Automobile
 - - Bicycle
 - - Walk

(b)



TravelMode
 — Automobile
 - - Bicycle
 - - Walk

(c)



TravelMode
 — Automobile
 - - Bicycle
 - - Walk

(d)

FIGURE 5: Nonlinear associations of determinants on travel mode choice. (a) Partial dependence of travel distance. (b) Partial dependence of road density. (c) Partial dependence of CBD distance. (d) Partial dependence of bus stop.

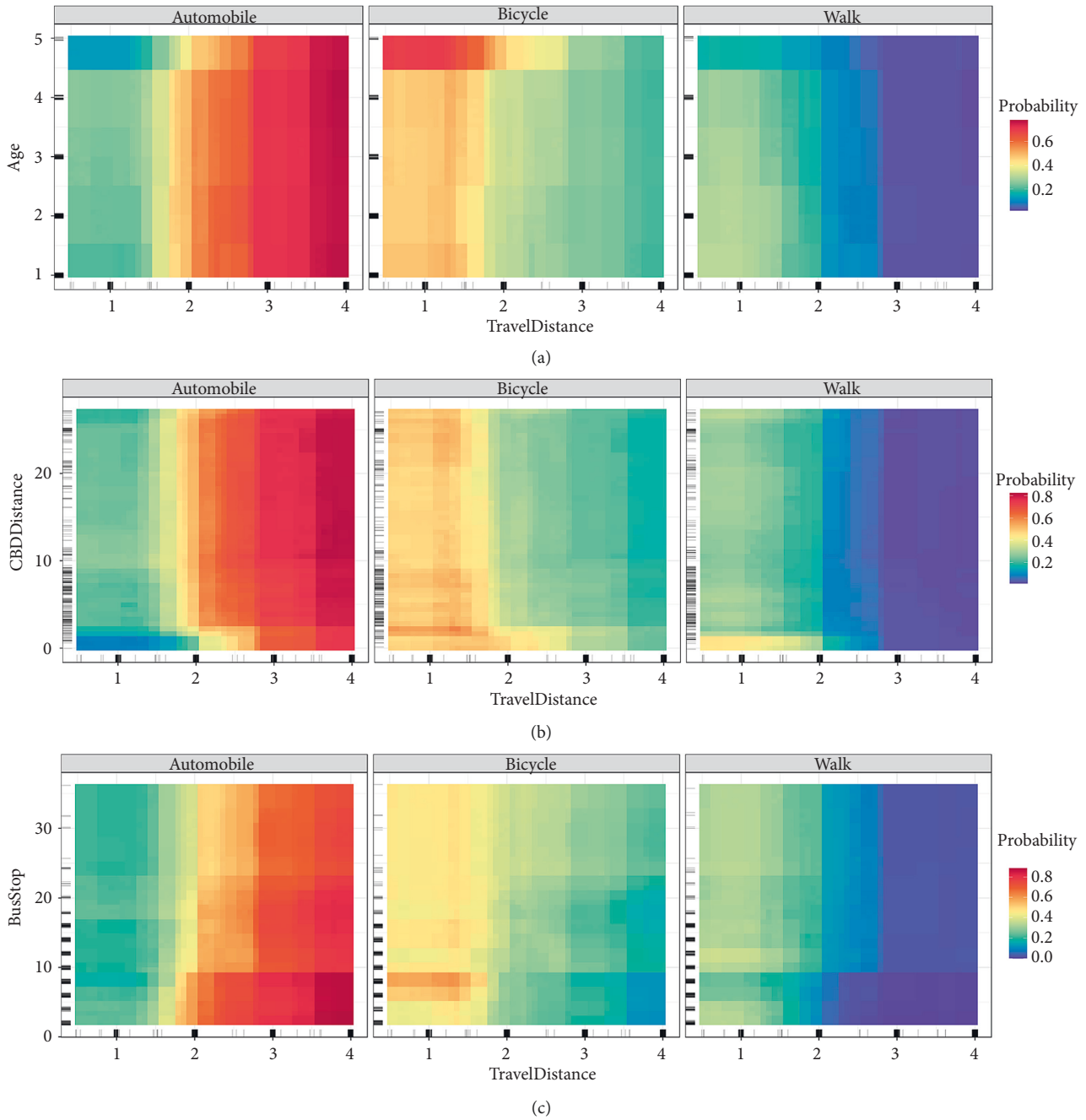


FIGURE 6: Interaction between explanatory variables. (a) The interaction between travel distance and age. (b) The interaction between travel distance and CBD distance. (c) The interaction between travel distance and bus stop.

addition, when the road network density was in the range of 2.5–5 km/km², the choice probability of walking was low; however, the probability of choice cycling was on the rise, highlighting the competitive complementary role of the two.

Figure 5(c) shows the threshold effect between the distance to the city center and travel mode choice. The results show that living in an area 1 km away from the city center decreased the probability of walking. Conversely, the tendency to use cars increased. When the distance exceeded 2.5 km, car users were less inclined to choose cycling.

Figure 5(d) illustrates the nonlinear effects of the number of bus stops on travel mode choice. We found that the number of bus stops had a significant effect on the travel mode choice of car users. When the number of bus stops near a place of residence exceeded 9, it attracted car users to shift to walking and cycling. This suggests that, for short-distance trips, a sufficient number of transit stops can encourage car users to walk or cycle. In other words, when the number of nearby bus stops meets the travel demand of car users, it can effectively reduce car travel and thus promote active modes.

TABLE 5: Interactions between travel distance and other determinants.

Variable	Interaction results	
	0 km < travel distance < 2 km	2 km < travel distance < 4 km
Trip purpose	Commuters rely more on cars for travel	—
Road density	2.5–7.5 km/km ² is more suitable for the development of cycling	—
Intersection density	—	More than 10 can reduce the probability of car trips and promote cycling
Travel cost	—	Preference for cycling within 10 RMB
Car use	—	When families use more, they cycle
Occupation	Retirees/other occupations are more likely to cycling	Daily commuters have a stronger need for car travel

Note: “—” indicates no interaction.

5.4. Interaction between Determinants. As mentioned in the literature section, there is interaction between different determinants. In other words, travel mode choices are driven not only by decisions due to a single variable. Thus, we further discussed the interaction of factors on the choice of travel mode. Since the findings in Sections 5.2 and 5.3 emphasized that travel distance is the most important factor in travel mode choice, in this section, we will further explore the interaction between travel distance and other explanatory variables by visualizing two-dimensional partial dependencies. Figure 6 mainly shows the interaction between travel distance and age, the distance to the CBD, and the number of bus stops.

Figure 6(a) shows the results of travel mode choice under the interaction of two characteristics, travel distance and age. The figure shows that, for trips less than 2 km, seniors over 60 years old are clearly more likely to travel by cycling than those in other age groups. The possible reason is that cycling, as a healthy travel mode that can exercise physical functions and improve immunity, is widely used by the elderly [48]. The interaction with age was not significant when the travel distance was more than 2 km. This suggests that when planning and designing, it is necessary to provide convenient cycling conditions and safety measures for the elderly to travel within 2 km.

Figure 6(b) shows that there is a significant interaction between travel distance and distance to the CBD. When both the travel distance and the distance to the CBD are less than 2 km, car users are willing to travel by cycling or walking. Otherwise, car trips are more likely to be used. This finding suggests that enhanced land use diversification and the development of land that basically meets the needs of travel activities in areas where walking and cycling are acceptable to car users benefit the development of active travel modes.

Figure 6(c) also shows a significant interaction between travel distance and the number of bus stops. We found that when the travel distance is greater than 2 km, cycling and driving are more advantageous than walking; however, when the number of bus stops exceeds 9, the probability of choosing a car is reduced, whereas that of choosing walking and cycling is increased. This indicates that when the travel distance exceeds 2 km, increasing public transportation facilities such as bus stops can help reduce car use. These findings are consistent with the findings of the nonlinear relationship in Section 5.3.

The results of the interaction analysis of travel distance with travel purpose, road network density, intersection density, acceptable transportation cost, household car use, and occupation are shown in Table 5. When the travel distance is within 2 km, commuting is more likely to drive car users to adopt car travel than travel purposes such as entertainment. Retirees and people in other occupations tend to choose cycling. On the other hand, when the travel distance is 2–4 km, cycling travel would be a good choice for car users if their cars are mainly used by their families or if the acceptable cost is less than 10 RMB. To provide good cycling conditions, a road network density of 2.5–7.5 km/km² and an intersection density greater than 10 are more appropriate.

5.5. Model Comparisons. Since the MNL model has been used widely to study travel mode choice behavior, we select it as the reference model to explore the differences between the linear and nonlinear models. Before establishing the MNL, we first tested the collinearity of the variables. The test results show that the variance inflation factor (VIF) of each variable was less than 5, indicating that there was no multicollinearity between the variables. Then, we built an MNL model based on maximum likelihood estimation. The evaluation indicators of model performance are the size of the area under the ROC curve (AUC), accuracy, recall, and precision. The latter three indicators can be calculated according to the confusion matrix. The calculation criteria are as follows:

$$\left\{ \begin{array}{l} \text{Accuracy} = \sum_k T_K \times \frac{100\%}{N}, \\ \text{Recall} = \sum_K \frac{T_K}{A_K} \times \frac{100\%}{K}, \\ \text{Precision} = \sum_K \frac{T_K}{P_K} \times \frac{100\%}{K}, \end{array} \right. \quad (4)$$

where T_K is the number of correct predictions in class K ; N is the total number of samples tested; A_K is the number of samples for the actual results in class K ; and P_K is the number of samples for the predicted results in class K .

TABLE 6: Model evaluation index results.

	Evaluation indicators			
	AUC	Accuracy (%)	Recall (%)	Precision (%)
RF	0.924	82.75	77.98	79.80
MNL	0.705	63.86	55.41	60.33

TABLE 7: MNL model estimation results.

Mode (automobile)	Bicycle		Walk	
	Coef.	$P > Z $	Coef.	$P > Z $
Constant	2.68	0.00**	0.58	0.60
	Travel information			
Travel distance	-0.63	0.00**	-1.27	0.00**
<i>Trip purpose (other)</i>				
Commuter	-0.36	0.09	-0.42	0.15
Entertainment	-0.07	0.74	0.08	0.78
Official business	-0.64	0.02*	-1.01	0.01*
	Built environment			
Bus stop	0.05	0.00**	0.10	0.00**
Bus route	-0.01	0.34	-0.01	0.63
Subway distance	-0.05	0.33	-0.20	0.02*
CBD distance	-0.02	0.11	0.01	0.42
Road density	0.02	0.47	0.10	0.03*
Intersection density	0.05	0.07	-0.01	0.76
Bus distance	-0.63	0.00**	1.77	0.06
Land use mix	-0.73	0.34	0.80	0.30
	Demographic characteristics			
<i>Travel cost (above 40 RMB)</i>				
Below 10 RMB	0.54	0.03*	0.96	0.01*
10–40 RMB	-0.04	0.83	0.54	0.08
<i>Age (61–70 years old)</i>				
18–30 years old	-1.50	0.00**	-1.43	0.02*
31–40 years old	-1.28	0.00**	-1.30	0.03*
41–50 years old	-1.04	0.01*	-0.85	0.16
51–60 years old	-1.62	0.00**	-1.76	0.00**
<i>Car use (used primarily by family)</i>				
Totally at my disposal	-0.64	0.00**	-0.24	0.36
Used primarily by me	-0.98	0.00**	0.02	0.93
<i>Occupation (other)</i>				
Self-employed	-0.62	0.03*	-0.39	0.29
Employed	-0.75	0.00**	-0.37	0.19
Retirement	-0.37	0.22	-0.13	0.78
<i>Education (highly)</i>				
Low	-0.57	0.03*	0.83	0.01*
Medium	-0.18	0.36	0.29	0.28
<i>Income (above 200,000 RMB)</i>				
Below 50,000 RMB	0.60	0.03*	0.45	0.24
50,000–200,000 RMB	0.54	0.01*	0.82	0.00**
Male (female)	0.11	0.49	-0.92	0.00**
	Pseudo $R^2 = 0.33$			
	LR $\chi^2 = 579.15$			

Note: *significance at the 10% level; **significance at the 5% level. The contents in parentheses are the reference group.

Table 6 shows the evaluation index results of the RF and MNL. The results showed that all evaluation indicators of RF were higher than those of MNL model, indicating that considering nonlinear associations between variables can greatly improve the predictive power of the model. The

estimation results of the MNL model are provided in Table 7. In this model, the relationship between the explanatory and dependent variables is prespecified. In contrast, the RF approach explores the relationship between the explanatory and dependent variables based on the association of the data, avoiding the a priori assumption of a specific function to fit the data [49]. Thus, using RF to explore the nonlinear associations and interactions of short-distance travel mode choices of car users not only improves the predictive performance of the model, but also allows us to further determine the effective range of variables that influence travel mode choices. In addition, the effect of variable interactions on travel mode choice can be obtained.

6. Discussion and Conclusions

Taking Kunming, China, as a case study, car users with a travel distance of less than 4 km are defined as short-distance car users. A questionnaire survey of car users was conducted, and a total of 1,425 valid short-distance travel samples were collected. Statistical descriptive results show that the share of car use exceeds 50% for short-distance trips by car users, reflecting a high level of car dependence. To effectively transfer car users to other modes of travel, this paper adopts the RF model to study the determinants of short-distance car users' travel mode choice and the complex relationships among them. Specifically, three main steps were taken: (1) assessing the relative importance of determinants; (2) revealing the nonlinear associations between the determinants and travel mode choice; and (3) exploring the interactions between travel distance and other variables. Overall, it fills a gap in the current research on travel mode choice in terms of nonlinear associations and exploring variable interactions. The results of the study provide recommendations for developing measures to reduce car trips over short distances.

Our study shows that built environment variables have a larger collective influence than economic demographic attribute variables and travel attribute variables. This is consistent with most current findings [14, 16, 50]. However, the relative importance of travel distance dominates overwhelmingly among all explanatory variables because each mode of travel has its own competitive range (more on this later). This finding suggests that we should improve the built environment within a certain distance to maximize the function of active travel modes. Going a step further, we identified characteristic variables that have a greater impact on cycling and walking than on car use. For cycling, the relative influence of distance to the nearest bus stop is more prominent. The meaning of this is that bicycle and bus feeder functions are considered in rational planning and design. This will not only increase the willingness to choose cycling, but also promote the codevelopment of green transportation modes. From a pedestrian perspective, the density of the road network is important. In other words, a pleasant and walkable environment enhances the possibility of walking trips. It is also necessary to focus on the distance to the nearest subway station and the distance to bus stops because convenient transportation facilities can promote the development of walking. The results of the above study help

provide planning managers with a reference point to guide car users toward green transportation while rationalizing the allocation of limited resources.

After identifying the key determinants of short-distance travel mode choice for car users, it is more important to specify the effective range of planning and design interventions. In our study, a 1.2 km travel distance is a turning point in travel mode choice, within which car users favor both walking and cycling. However, some studies have proposed 3 km as the dividing point between motorized and active modes of transportation [51]. In comparison, in terms of travel distance, car users have more stringent requirements regarding the alternative of active travel modes. In terms of the built environment, a road network density of 2.5–7.5 km/km² can promote the development of active travel modes. However, it is important to note that the distance to the city center is unfavorable for cycling when it exceeds 2.5 km. Therefore, group development should be considered when planning land use to ensure a balance between commercial, residential, and other work and housing [52]. In addition, the deployment of public transportation infrastructure deserves attention. The results of the study indicate a positive impact on active travel modes when the number of bus stops within 500 m of a residence exceeds 9. Therefore, to improve the efficiency of travel, for the coordinated development of various green transportation modes, it is recommended that connections between modes of travel be fully considered. These suggested values can help planners find the best design threshold or range. It is beneficial for active traffic development and can improve the cost effectiveness of construction to a certain extent.

On the other hand, the interactions between the determinants are more consistent with real-life situations where travel mode choice is affected by more than a single variable in the decision. Therefore, this study explores the interaction between travel distance and other determinants. For trips within 2 km, compared with younger people, people over 60 years of age tend to use cycling. This is a noteworthy phenomenon. To improve the quality of life of the elderly, it is recommended that planning and design should be carried out with attention to providing convenient cycling conditions and safety measures for the elderly. In addition, commuting is the main purpose of travel for car use, probably because cars can improve the convenience of commuting [53]. For 2 to 4 km trips, the choice of active travel modes increases when the car is used mainly by family members or when the acceptable transportation cost is less than 10 RMB. Therefore, the interaction between key variables should also be taken into account in planning management and policy development.

The above research results can provide detailed guidance for specific planning and management work. However, this paper also has some shortcomings. First, the impact may vary with spatial scale [50]. The scope of data collection for the built environment in this study is a 500 m area around the car users' residence. It is necessary to explore different spatial scales and collect more comprehensive information about the characteristics of the built environment. Second, this study focuses on the nonlinear associations of short-

distance travel mode choice of car users, focusing mainly on guiding car users to active travel modes. Therefore, it equally ignores green public transportation modes such as buses, subways, and trams, which can be further expanded in future studies. Third, random forests fit the available data well, but the reliability of predictions for regions without data will be reduced, as random forest is a data-driven method. Thus, forecasting in certain ranges where data are lacking will produce abnormal fluctuations. Therefore, the amount and density of data in further studies are crucial for random forest prediction.

Data Availability

All the data included in this study are available upon request to the corresponding author (kmluiyang@outlook.com).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (Grant nos. 71861017) and 52102378) and Yunnan Fundamental Research Projects (Grant nos. 202001AT070030 and 202201AU070109).

References

- [1] M. Migliore, G. D'Orso, and D. Caminiti, "The environmental benefits of carsharing: the case study of Palermo," *Transportation Research Procedia*, vol. 48, pp. 2127–2139, 2020.
- [2] C. Park and S. Y. Sohn, "An optimization approach for the placement of bicycle-sharing stations to reduce short car trips: an application to the city of Seoul," *Transportation Research Part A: Policy and Practice*, vol. 105, pp. 154–166, 2017.
- [3] A. Neves and C. Brand, "Assessing the potential for carbon emissions savings from replacing short car trips with walking and cycling using a mixed GPS-travel diary approach," *Transportation Research Part A: Policy and Practice*, vol. 123, pp. 130–146, 2019.
- [4] J. Lee, S. Y. He, and D. W. Sohn, "Potential of converting short car trips to active trips: the role of the built environment in tour-based travel," *Journal of Transport & Health*, vol. 7, pp. 134–148, 2017.
- [5] M. Li, G. Song, Y. Cheng, L. Yu, and B. S. D. Sagar, "Identification of prior factors influencing the mode choice of short distance travel," *Discrete Dynamics in Nature and Society*, vol. 2015, Article ID 795176, 9 pages, 2015.
- [6] X. Ma, J. Yang, C. Ding, and J. Q. Liu, "Joint analysis of the commuting departure time and travel mode choice: role of the built environment," *Journal of Advanced Transportation*, vol. 2018, Article ID 4540832, 13 pages, 2018.
- [7] S. Gandhi and G. Tiwari, "Sociopsychological, instrumental, and sociodemographic determinants of travel mode choice behavior in Delhi, India," *Journal of Urban Planning and Development*, vol. 147, no. 3, 2021.
- [8] M. Enayat, K. Reza, and M. Dominique, "Exploring the effect of the built environment, weather condition and departure time of travel on mode choice decision for different travel

- purposes: evidence from Isfahan, Iran,” *Case Studies on Transport Policy*, vol. 9, 2021.
- [9] Z. Gan, M. Yang, T. Feng, and H. J. P. Timmermans, “Examining the relationship between built environment and metro ridership at station-to-station level,” *Transportation Research Part D: Transport and Environment*, vol. 82, Article ID 102332, 2020.
- [10] K. Wang and B. Ozbilen, “Synergistic and threshold effects of telework and residential location choice on travel time allocation,” *Sustainable Cities and Society*, vol. 63, Article ID 102468, 2020.
- [11] L. Jixiang, W. Bo, and X. Longzhu, “Non-linear associations between built environment and active travel for working and shopping: an extreme gradient boosting approach,” *Journal of Transport Geography*, vol. 92, 2021.
- [12] J. Yang, J. Cao, and Y. Zhou, “Elaborating non-linear associations and synergies of subway access and land uses with urban vitality in Shenzhen,” *Transportation Research Part A: Policy and Practice*, vol. 144, pp. 74–88, 2021.
- [13] M. Tu, W. Li, O. Orfila, Y. Li, and D. Gruyer, “Exploring nonlinear effects of the built environment on ridesplitting: evidence from Chengdu,” *Transportation Research Part D: Transport and Environment*, vol. 93, Article ID 102776, 2021.
- [14] L. Cheng, J. De Vos, P. Zhao, M. Yang, and F. Witlox, “Examining non-linear built environment effects on elderly’s walking: a random forest approach,” *Transportation Research Part D: Transport and Environment*, vol. 88, Article ID 102552, 2020.
- [15] C. Ding and P. Næss, “Applying gradient boosting decision trees to examine non-linear effects of the built environment on driving distance in Oslo,” *Transportation Research Part A: Policy and Practice*, vol. 110, pp. 107–117, 2018.
- [16] E. Chen, Z. Ye, and H. Wu, “Nonlinear effects of built environment on intermodal transit trips considering spatial heterogeneity,” *Transportation Research Part D: Transport and Environment*, vol. 90, Article ID 102677, 2021.
- [17] L. Cheng, X. Chen, J. De Vos, X. Lai, and F. Witlox, “Applying a random forest method approach to model travel mode choice behavior,” *Travel Behaviour and Society*, vol. 14, pp. 1–10, 2019.
- [18] X. Chang, J. Wu, H. Liu, and X. H. Y. Yan, “Travel mode choice: a data fusion model using machine learning methods and evidence from travel diary survey data,” *Transportmetrica: Transportation Science*, vol. 15, no. 2, pp. 1587–1612, 2019.
- [19] X. Zhao, X. Yan, A. Yu, and P. Van Hentenryck, “Prediction and behavioral analysis of travel mode choice: a comparison of machine learning and logit models,” *Travel Behaviour and Society*, vol. 20, pp. 22–35, 2020.
- [20] R. Cole, G. Turrell, M. J. Koohsari, N. Owen, and T. Sugiyama, “Prevalence and correlates of walkable short car trips: a cross-sectional multilevel analysis,” *Journal of Transport & Health*, vol. 4, pp. 73–80, 2016.
- [21] A. Dėdelė, A. Miškinytė, S. Andrušaitytė, and J. Nemaniiūtė-Gužienė, “Dependence between travel distance, individual socioeconomic and health-related characteristics, and the choice of the travel mode: a cross-sectional study for Kaunas, Lithuania,” *Journal of Transport Geography*, vol. 86, 2020.
- [22] L. Cheng, X. Chen, S. Yang, and J. M. Wu, “Structural equation models to analyze activity participation, trip generation, and mode choice of low-income commuters,” *Transportation Letters*, vol. 11, no. 6, pp. 341–349, 2019.
- [23] T. H. de Sá, D. C. Parra, and C. A. Monteiro, “Impact of travel mode shift and trip distance on active and non-active transportation in the São Paulo Metropolitan Area in Brazil,” *Preventive medicine reports*, vol. 2, pp. 183–188, 2015.
- [24] X. Cao, S. L. Handy, and P. L. Mokhtarian, “The influences of the built environment and residential self-selection on pedestrian behavior: evidence from austin, TX,” *Transportation*, vol. 33, no. 1, pp. 1–20, 2006.
- [25] C. Ding, D. Wang, C. Liu, Y. Zhang, and J. Yang, “Exploring the influence of built environment on travel mode choice considering the mediating effects of car ownership and travel distance,” *Transportation Research Part A: Policy and Practice*, vol. 100, pp. 65–80, 2017.
- [26] L. Yu, B. Xie, and E. H. W. Chan, “Exploring impacts of the built environment on transit travel: Distance, time and mode choice, for urban villages in Shenzhen, China,” *Transportation Research Part E*, vol. 132, 2019.
- [27] J. De Vos, P. L. Mokhtarian, T. Schwanen, V. Van Acker, and F. Witlox, “Travel mode choice and travel satisfaction: bridging the gap between decision utility and experienced utility,” *Transportation*, vol. 43, no. 5, pp. 771–796, 2016.
- [28] M. Aghaabbasi, Z. A. Shekari, M. Z. Shah, and O. D. J. M. Olakunle, “Predicting the use frequency of ride-sourcing by off-campus university students through random forest and Bayesian network techniques,” *Transportation Research Part A: Policy and Practice*, vol. 136, pp. 262–281, 2020.
- [29] Y. Xiang, L. Xinyu, and Z. Xilei, “Using machine learning for direct demand modeling of ridesourcing services in Chicago,” *Journal of Transport Geography*, vol. 83, 2020.
- [30] Y. Xu, X. Yan, X. Liu, and X. Zhao, “Identifying key factors associated with ridesplitting adoption rate and modeling their nonlinear relationships,” *Transportation Research Part A: Policy and Practice*, vol. 144, pp. 170–188, 2021.
- [31] E. Chen and Z. Ye, “Identifying the nonlinear relationship between free-floating bike sharing usage and built environment,” *Journal of Cleaner Production*, vol. 280, Article ID 124281, 2021.
- [32] K. Wang and X. Wang, “Generational differences in auto-mobility: comparing America’s Millennials and Gen Xers using gradient boosting decision trees,” *Cities*, vol. 114, Article ID 103204, 2021.
- [33] Y. Linchuan, A. Yibin, K. Jintao, L. Yi, and L. Yuan, “To walk or not to walk? Examining non-linear effects of streetscape greenery on walking propensity of older adults,” *Journal of Transport Geography*, vol. 94, 2021.
- [34] C. Ding, X. Cao, and C. Liu, “How does the station-area built environment influence Metrorail ridership? Using gradient boosting decision trees to identify non-linear thresholds,” *Journal of Transport Geography*, vol. 77, pp. 70–78, 2019.
- [35] S. Qifan, Z. Wenjia, C. Xinyu, Y. Jiawen, and Y. Jie, “Threshold and moderating effects of land use on metro ridership in Shenzhen: implications for TOD planning,” *Journal of Transport Geography*, vol. 89, 2020.
- [36] X. Wu, T. Tao, J. Cao, Y. Fan, and A. Ramaswami, “Examining threshold effects of built environment elements on travel-related carbon-dioxide emissions,” *Transportation Research Part D: Transport and Environment*, vol. 75, pp. 1–12, 2019.
- [37] C. Ding, X. Cao, and Y. Wang, “Synergistic effects of the built environment and commuting programs on commute mode choice,” *Transportation Research Part A: Policy and Practice*, vol. 118, pp. 104–118, 2018.
- [38] SBO Kuming, *Kunming National Economic and Social Development Statistical Communiqué*, 2020.
- [39] M. He, Y. Fei, M. He, and J. Lee, “Exploring the factors associated with car use for short trips: evidence from kunming,

- China,” *Journal of Advanced Transportation*, vol. 2020, Article ID 3654130, 10 pages, 2020.
- [40] C. Jakobsson, S. Fujii, and T. Gärling, “Determinants of private car users’ acceptance of road pricing,” *Transport Policy*, vol. 7, no. 2, pp. 153–158, 2000.
- [41] TloCP Institute, *Annual Report on Road Network Density and Traffic Operation in Major Chinese Cities*, 2021.
- [42] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [43] A. B. Atkinson, “On the measurement of inequality,” *Academic Press*, vol. 2, no. 3, 1970.
- [44] J. H. Friedman, “Stochastic gradient boosting,” *Computational Statistics & Data Analysis*, vol. 38, no. 4, pp. 367–378, 2002.
- [45] R. Tamakloe, J. Hong, and J. Tak, “Determinants of transit-oriented development efficiency focusing on an integrated subway, bus and shared-bicycle system: application of Simar-Wilson’s two-stage approach,” *Cities*, vol. 108, Article ID 102988, 2021.
- [46] M. Attard, C. Cañas, and S. Maas, “Determinants for walking and cycling to a university campus: insights from a participatory Active Travel workshop in Malta,” *Transportation Research Procedia*, vol. 52, pp. 501–508, 2021.
- [47] D. V. Jonas, C. Long, M. D. Kamruzzaman, and W. Frank, “The indirect effect of the built environment on travel mode choice: a focus on recent movers,” *Journal of Transport Geography*, vol. 91, 2021.
- [48] J. Ryan, H. Svensson, J. Rosenkvist, S. M. Schmidt, and A. Wretstrand, “Cycling and cycling cessation in later life: findings from the city of Malmö,” *Journal of Transport & Health*, vol. 3, no. 1, pp. 38–47, 2016.
- [49] S. Sabouri, S. Brewer, and R. Ewing, “Exploring the relationship between ride-sourcing services and vehicle ownership, using both inferential and machine learning approaches,” *Landscape and Urban Planning*, vol. 198, Article ID 103797, 2020.
- [50] T. Tao, J. Wang, and X. Cao, “Exploring the non-linear associations between spatial attributes and walking distance to transit,” *Journal of Transport Geography*, vol. 82, Article ID 102560, 2020.
- [51] S. Li and P. Zhao, “The determinants of commuting mode choice among school children in Beijing,” *Journal of Transport Geography*, vol. 46, pp. 112–121, 2015.
- [52] D. Chetan, M. Manoj, and M. Yashasvi, “Geographical scale of residential relocation and its impacts on vehicle ownership and travel behavior,” *Journal of Transport Geography*, vol. 94, 2021.
- [53] F. Ahmed, J. Catchpole, and T. Edirisinghe, “Understanding young commuters’ mode choice decision to use private car or public transport from an extended theory of planned behavior,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2675, no. 3, pp. 200–211, 2021.