The Impact of Spatio-Temporal Constraints on Tourists' Travel Mode Choice Behaviour – Case of Sijiao Island, China

Yazao YANG[1], Hui WEN[2], Tangzheng WENG[3]

[1] yyzjt@cqjtu.edu.cn, Chongqing Key Laboratory of Intelligent Integrated and Multidimensional Transportation System, Chongqing Jiaotong University, Chongqing, China
[2] 9099177155@qq.com, School of Traffic and Transportation, Chongqing Jiaotong University, Chongqing, China
[3] 2421910999@qq.com, School of Traffic and Transportation, Chongqing Jiaotong University, Chongqing, China

ABSTRACT

With the continuous rise in tourism consumption demand, the spatio-temporal aggregation effects are strengthening, leading to an imbalance in travel structure within tourist destinations. To address these issues, this study focuses on SiJiao Island in Zhoushan, China. It empirically studies tourists' mode choice behaviour from the perspectives of personal social attributes, temporal attributes and spatial attributes. This study gathers tourist characteristics through surveys and multidimensional data mining. It employs a multinomial logistic regression method to identify influencing factors and constructs a gradient boosting decision tree model to explore the relationship between spatio-temporal factors and tourists' mode choice behaviour. The results show that the cumulative importance of temporal and spatial attributes is much greater than that of personal social attributes. Among the many spatiotemporal influencing factors, travel distance, transfer time, departure time and the number of bus stops have a significant impact on tourists' mode choice behaviour, accounting for 26.25%, 16.57%, 14.65% and 11.87% of the total importance, respectively. Additionally, there are complex non-linear relationships between these influencing factors and tourists' mode choice behaviour. The interaction effects of spatiotemporal combinations on tourists' mode choice behaviour exhibit correlations. The sensitivity of mode choice to different spatiotemporal factors varies.

KEYWORDS

spatio-temporal factors; machine learning; travel mode choice; active travel.

1. INTRODUCTION

Tourism, as a strategic pillar industry of the national economy, has long been driving economic growth. However, the outbreak of the COVID-19 pandemic in 2020 led to a "cold winter" for domestic tourism. Following the reclassification of COVID-19 as a Category B infectious disease in 2023, the demand for tourism consumption surged. According to the 2023 National Economic and Social Development Statistical Bulletin released by the National Bureau of Statistics, the total number of domestic tourists reached 2.89 billion, representing a year-on-year increase of 93.3%. Domestic tourism revenue reached 4.9 trillion yuan, marking a year-on-year increase of 140.3%, as shown in *Figure 1*.

Observations of tourists' digital footprints reveal strong spatio-temporal clustering in their travel activities [1, 2]. This means tourists often engage in many travel activities simultaneously, heading to the same destination. During the COVID-19 pandemic, tourists strongly avoided clustered activities. However, after COVID-19 was reclassified as a Category B infectious disease, the domestic tourism industry recovered, and tourists became more willing to engage in clustered activities such as tourism [3, 4]. This has further intensified the spatio-temporal clustering effect, leading to traffic congestion and parking difficulties around tourist spots.

Research indicates that the purpose of travel largely determines travellers' tolerance for time and space, with leisure activities reflecting a stronger demand for both. For example, Chinese residents have an average commuting time of over 36 minutes and a commuting radius of 25 kilometres [5]. This indicates that travellers exhibit high spatio-temporal tolerance during commuting activities. However, when they assume the role of tourists, they tend to limit their travel time and distance for leisure activities to within 15 minutes and 8 kilometres, often remaining in areas around their hotels [6, 7]. Given these spatio-temporal demands, tourists tend to prefer travel modes that offer shorter waiting times, faster speeds and greater flexibility, such as driving and taxis. Furthermore, researchers have found that built environment factors, such as the duration of activities in the travel chain, higher road network density near residences and accessibility to bus stops, significantly influence travel mode choices [8]. In summary, time and space are crucial factors influencing tourists' travel behaviour, significantly affecting their travel experience and mode choice. Therefore, studying the impact of spatio-temporal factors on tourists' travel mode choices is highly significant.
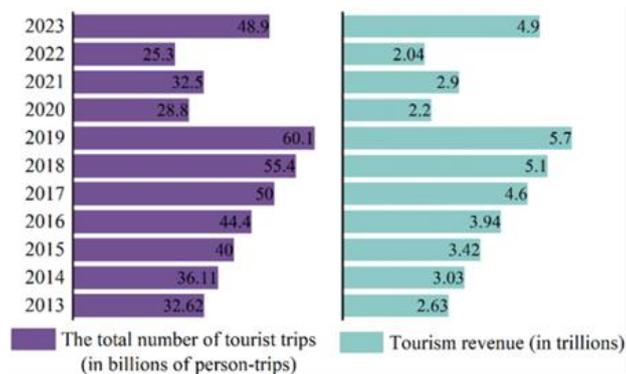

Figure 1 – The total number of tourist trips and tourism revenue from 2013 to 2023

The paper is structured as follows. Section 1 introduces the growing spatio-temporal concentration of tourist travel behaviour in the post-COVID era and highlights the need to understand how temporal and spatial constraints affect tourists' mode choice, taking SiJiao Island as a case study. Section 2 reviews existing literature on travel mode choice, categorising influencing factors into personal-social, temporal and spatial attributes, and notes the increasing application of machine learning methods to capture complex, non-linear relationships. Section 3 describes the data collection and presents key descriptive statistics from 858 valid questionnaires, followed by the construction of a gradient boosting decision tree (GBDT) model to analyse how various attributes affect travel mode choice. Section 4 discusses the model results, showing that spatial and temporal attributes play a more significant role than personal-social attributes in influencing mode choice, particularly travel distance, transfer time, departure time and the number of bus stops. Section 5 further interprets these findings by analysing non-linear and interaction effects by using partial dependence and elasticity analyses, and offers practical recommendations for destination traffic planning and operational strategies. Section 6 concludes the paper by reaffirming the effectiveness of GBDT in modelling tourists' mode choice under spatio-temporal constraints and suggests that future research should explore entire travel chains and the broader influence of the built environment.

2. LITERATURE REVIEW

In our review of factors influencing tourists' travel mode choices, we classify influencing factors into three categories: personal social attributes, temporal attributes and spatial attributes. This classification is based on established frameworks in transportation research that highlight the importance of these dimensions in understanding travel behaviour. Personal social attributes reflect individual demographics and socio-economic status, temporal attributes encompass factors related to time of travel and its flexibility, and spatial attributes involve geographical and environmental aspects. This tripartite framework allows for a comprehensive analysis of the diverse influences on mode choice, aligning with our research objectives to examine how these varied factors interact.

Firstly, on the level of individual social attributes, Zhang et al. [9] conducted a survey on the travel mode choice behaviour among tourists in Beijing. This survey indicated that factors such as tourists' age, private car ownership and the number of companions significantly influence their choice of travel mode. Meanwhile, Liu

[10], through field investigations, found that monthly income and car ownership are the key factors affecting tourists' travel mode choices. In addition, Limtanakool et al. [11] observed that the impact of individual social attributes on tourists' travel mode choice behaviour varies with different travel distances.

Time factors have multifaceted impacts on travel behaviour, primarily involving departure time, travel duration and arrival time. These temporal factors significantly influence individuals' travel decision-making processes. Consequently, many scholars opt to base their research on commuting activities, constructing models to delve deeply into the mechanisms by which these time factors exert their effects. Guo et al. [12] conducted a study on the departure time choices of public transit commuters, in which they developed utility functions for normal, early and delayed departures. It also calculated the choice probabilities for these three departure times under different dynamic information scenarios. On the other hand, Yu et al. [13], based on the travel activity bottleneck model proposed by Li [14], examined the behaviour of commuters choosing to travel by metro in terms of their choice of departure time under the influence of bike sharing. Researchers typically do not isolate the impact of travel time on behaviour. Instead, they treat it as a characteristic variable influencing travel behaviour, constructing models to understand its effects [15, 16].

The spatial factors influencing travel behaviour are diverse. Based on current research, these factors can generally be divided into two main categories. The first category is travelling distance, which typically refers to the actual distance between the origin and destination. The second category is the built environment, which includes the road network within buffer zones, transportation facilities, building areas, and the surrounding geographic and natural landscapes. Recent research shows that as travel distance increases, there is a significant shift in the modal split among different travel modes [17]. Specifically, travel distance is positively correlated with the choice probabilities of private cars and buses, but negatively correlated with active travel [18]. Compared to travel distance, the built environment is a spatial factor that travellers can tangibly perceive. Researchers often quantify aspects of the built environment to explore their impact on travel behaviour. For instance, increasing green spaces and the density of apartment buildings can effectively enhance the use of active travel modes [19, 20], Improving the accessibility of bus stops increases the share of public transportation use, as some travellers switch from active modes to buses [21], and this also reduces the share of private cars and taxis [22]. Wali et al. [23] and He et al. [24] explored non-linear relationships and transition thresholds between travel modes and the built environment using generalised additive models and random forest models, respectively.

Due to travel mode being a discrete variable, existing literature tends to focus on constructing logit models based on utility theory to analyse tourists' choice behaviour. For example, Hatamzadeh et al. [25] modelled the choice of walking patterns of elderly men and women in Rasht city, Gilan province, Iran, by means of a discrete choice model. Tang et al. [26] analysed the influencing factors of travel destinations as well as travel mode choice behaviour of tourists in Hangzhou based on a multi-indicator logit model. Heinen et al. [27] employed multiple regression models to assess the influence of social, economic, transportation and destination factors on various travel modes.

Meanwhile, as data volumes increase and computational capabilities improve, machine learning models are gradually being applied in the field of transportation. In recent years, machine learning models like random forests and gradient boosting decision trees have gained popularity for quantifying non-linear and threshold effects between response variables (traffic volume or behavioural decisions) and explanatory variables (such as age, travel distance and built environment factors) [28-31]. Machine learning models do not require strict a priori assumptions, thus partially overcoming the limitations of traditional parametric statistical models. Their ability to derive impact relationships and threshold ranges can effectively guide decision-making.

Therefore, this study will introduce various spatio-temporal variables, such as free activity time, transfer time, building area density and the number of bus stops, on the basis of traditional influencing factors. It will construct a model of tourists' travel mode choice behaviour using machine learning algorithms. This model aims to investigate the impact of multidimensional influencing factors on tourists' travel mode choices, analyse the non-linear relationships between influencing factors and travel mode choice behaviours and clarify the spatio-temporal thresholds for each travel mode.

3. DATA

This section focuses on the collection and analysis of data related to factors that influence tourists' choice of travel modes.

3.1 Details of the study area data

Shengsi County, also known as Shengsi Archipelago, is under the jurisdiction of Zhoushan City, Zhejiang Province. It is located in the easternmost part of Zhejiang Province, the northernmost part of the Zhoushan Archipelago, between 121°30′E-123°25′E and 30°24′N-31°04′N.

This study focuses on researching the travel mode choice behaviour of tourists within Sijiao Mountain Island. Considering the constraints of natural conditions such as terrain within the study area, the determination of study units referred to administrative regions and divided traffic zones based on administrative boundaries.

3.2 Questionnaire design and distribution

A questionnaire survey was used to investigate the travel behaviour of tourists. The questionnaire includes three aspects: personal social attributes, temporal attributes and spatial attributes. In terms of personal social attributes, key factors include age, gender, occupation, educational level, disposable income, travel purpose and travel mode. For temporal attributes, the main data collected are the tourists' departure time, free activity time and transfer time. Although spatial attributes do not directly provide the travel distance of tourists, it is possible to obtain this information by asking tourists about the coordinates of their departure and destination points and using mapping services like Baidu Maps or Amap to calculate the travel distance.

The distribution of this survey questionnaire was carried out using a combination of online surveys and on-site interviews. Online surveys were generated using the professional online survey platform Wenjuanxing (https://www.wjx.cn/), while offline surveys involved visiting selected study units and randomly distributing questionnaires to tourists in the area. The formal survey of the questionnaire began on 25 July 2022 and lasted until 22 August 2022. A total of 962 questionnaires were collected in this survey. After screening and cleaning the questionnaires, 104 invalid questionnaires were removed, resulting in 858 valid questionnaires. The effective response rate was 89.19%.

3.3 Descriptive statistics

The personal data and travel behaviour data obtained online and offline are assigned corresponding logical values, and the valid questionnaires collected this time are uniformly organised. Partial statistical results are shown in *Table 1*.

Table 1 – Basic statistics of the questionnaire

Variables	Category	Sample size	Percentage	Variables	Category	Sample size	Percentage
Sex	Male	462	53.38%	Education	Primary school and below	73	8.51%
	Female	396	46.62%		Middle school	127	14.80%
Age	<16	112	13.05%		Vocational school/high school	152	17.72%
	16~30	285	33.22%		College/Undergraduate	471	54.90%
	31~45	273	31.82%		Master's and above	35	4.07%
	46~60	105	12.24%	Disposable income	<3000Yuan	178	20.75%
	>60	81	9.67%		3000~5000Yuan	144	16.78%
Occupation	Student	147	17.13%		5000~7000Yuan	179	20.86%
	Enterprise staff	297	34.62%		7000~9000Yuan	118	13.75%
	Government officials	120	13.99%		>9000Yuan	237	27.86%
	Self-employed	132	15.38%	Travel mode	Independent travel	585	68.18%
	Retirees	66	7.69%		Semi-independent travel	67	7.81%
	Others	96	11.19%		Group travel	206	24.01%
Travel purposes	Tourist attractions	657	76.57%				
	Others	201	23.43%				

In this survey, a total of 858 tourists were effectively surveyed, including 456 male tourists, which is 6.76% more than the number of female tourists. In terms of age, the tourists more willing to participate in the survey are mainly distributed between 16 and 30 years old and 31 and 45 years old, with a combined proportion of 65.04%. Regarding occupation, the highest proportion of tourists is employed in enterprises and institutions, while the lowest proportion is retirees. Concerning education level, the surveyed tourists generally have a college degree or above, accounting for 54.90%. In terms of disposable income, there is not much difference in the proportion of tourists in each income bracket; however, the highest proportion is of tourists with an income of more than 9,000 yuan.

Within tourist destinations, common travel modes include walking, (electric) bicycles, motorcycles, private cars, ride-hailing services, taxis, urban buses and subways, among others. Although there are a variety of travel options, the availability of these modes varies due to the differing socio-economic conditions at each tourist destination. To succinctly describe the travel modes within tourist destinations, this paper categorises the available transportation means based on those present in the study area: for example, walking, bicycles and electric bicycles are grouped under active travel; ride-hailing and taxis are classified together as taxi services; private and rental cars are categorised under self-driving; and medium-capacity transport such as tourist charters, rural-urban buses and urban buses are classified as public buses. The study found that male tourists are more likely to choose driving, whereas female tourists have a higher preference for buses. Tourists under 30 are more inclined to choose flexible and active travel options, while buses are mainly favoured by those under 16 and over 60. Self-employed individuals are more likely to choose driving, whereas retired individuals tend to prefer buses. As education level increases, the proportion of tourists choosing driving and taxis rises, along with an increased acceptance of buses. Tourists with lower disposable income are more likely to choose active travel and buses. When visiting scenic spots, tourists prefer active travel and taxis. For other travel purposes, active travel is chosen more frequently than other modes. The study found that in the morning and afternoon, active travel is the most popular choice among tourists, followed by buses. At noon, the preference shifts towards driving and taxis. For long-term activities at scenic spots or other destinations, tourists tend to choose highly motorised modes of transport such as driving, taxis and buses.

3.4 Collecting spatial elements

The transportation network data mentioned in this article consist of road network line distribution data, bicycle lane distribution data, pedestrian street distribution data, etc., within the study area. Road data for Sijiao Island in 2022 were exported from the OpenStreetMap website.

Building base data can depict the shape and appearance of buildings within the research area, providing a visual understanding of the characteristics of buildings. The building base data used in this article were obtained by downloading building outline images from the Tianditu map and performing image analysis using the ArcGIS software.

4. METHODOLOGY

This section primarily focuses on developing a predictive model for tourists' travel mode choice behaviour using the gradient boosting decision tree algorithm.

4.1 Model basic assumptions

Tourists are easily influenced by multidimensional factors when choosing transportation modes, making it particularly important to explore the correlation between influencing factors and modes of transportation. Currently, researchers use various methods such as machine learning and discrete choice models to analyse the relationship between them. However, before constructing the model, it is necessary to make assumptions about the model in advance to simplify the complexity of the problem and clarify the direction of analysis. The basic assumptions of this paper are as follows:
1) Although tourists typically travel as families during trips, the choice of transportation mode is solely determined by the tourists themselves and does not involve family decisions. Therefore, the basic unit for tourists' choice of transportation mode is the individual tourist.
2) Within the study area, the number of tourists will not increase significantly in a short period, and relevant parameters such as local policies, transportation means and socioeconomic development levels will not undergo drastic changes in the near future.

3) When tourists face the choice of transportation modes, it is only related to the influencing factors considered in this paper and has no correlation with other influencing factors.

Once the prerequisite assumptions are established and the basic travel information of tourists is obtained through surveys and big data platforms, these data can be utilised to train machine learning models, ultimately enabling the prediction of tourists' chosen modes of travel.

4.2 Model construction

This paper uses the gradient boosting decision tree (GBDT) model to explore tourists' choice of transportation modes under temporal and spatial constraints. The gradient boosting decision tree (GBDT) method was selected for its ability to model complex, non-linear relationships and its robustness in handling large datasets with multiple features. GBDT has been shown to outperform traditional methods like logistic regression or decision trees in predictive accuracy, particularly when interactions between features are not linear. Compared to other machine learning methods, such as random forests or support vector machines, GBDT offers better interpretability, allowing for a clear understanding of the contribution of individual features to the prediction. This makes it an ideal choice for analysing the travel mode choices of tourists, where multiple factors interact in complex ways.

This paper quantifies each factor influencing tourists' choice behaviour and maps them to independent variables. When constructing the model, it is usually necessary to create a sample for each tourist, inputting the travel data of each tourist, using various attribute values as input variables, and the chosen mode of transportation by the tourist as the output variable to study the correlation between the two. The format of sample data can be represented as *Equation 1*.

$$\vec{P}_j = (x_1, \cdots, x_n, y_j) \tag{1}$$

x_i represents the i-th factor influencing tourists' choice of transportation mode, where $i = 1, 2, \cdots, n$; n is the total number of factors influencing tourists' choice of transportation mode; y_j represents the mode of transportation chosen by the j-th sample, where $j = 1, 2, \cdots, m$. The expression of input variables is shown as *Equation 2*.

$$\vec{X}_j = (x_1, \cdots, x_n) \tag{2}$$

Similarly, the expression of the output variables is shown in *Equation 3*.

$$\vec{Y}_j = y_j \tag{3}$$

Suppose a total of m tourist travel samples are collected. If they are arranged and combined, the overall input variable matrix required by the model can be obtained, as shown in *Equation 4*.

$$\begin{bmatrix} \vec{X}_1 \\ \vec{X}_2 \\ \vdots \\ \vec{X}_m \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \tag{4}$$

According to the above equations, the mapping relationship of the model is shown in *Equation 5*.

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \xrightarrow{f(\theta_1, \theta_2, \cdots, \theta_L)} \begin{bmatrix} \vec{Y}_1 \\ \vec{Y}_2 \\ \vdots \\ \vec{Y}_m \end{bmatrix} \tag{5}$$

$f(\theta_1, \theta_2, \cdots, \theta_L)$ represents the mapping relationship from input variables to output variables; θ_i represents the internal parameters of the mapping relationship, where $i = 1, 2, \cdots, L$. Based on the above approach, the model for tourists' choice of transportation modes can be established using the gradient boosting decision tree machine learning algorithm. The corresponding expression for the model of transportation mode choice behaviour is shown in *Equation 6*.

$$Y_{GBDT} = F_{GBDT}(x_1, \cdots, x_n, \theta_1, \cdots, \theta_{LG}) \tag{6}$$

F_{GBDT} represents the mapping relationship between the gradient boosting decision tree machine learning model and influencing factors θ_{LG} represents the internal parameters of the gradient boosting decision tree

machine learning model; Y_{GBDT} represents the mode of transportation chosen by tourists predicted by the gradient boosting decision tree machine learning model.

4.3 Model calculation process

After the model is constructed, it needs to be tested with sample data to validate its performance. The testing process generally consists of two parts: one part is training and the other part is testing. To improve the final predictive performance of the model, continuous adjustments need to be made to the model's internal parameters. The machine learning model with adjusted parameters is represented as *Equation 7*.

$$Y_{GBDT} = F_{GBDT}(x_1, \cdots, x_n, \theta_1^*, \cdots, \theta_{LG}^*) \tag{7}$$

5. RESULTS

This section uses the travel mode choice behaviour model based on gradient boosting decision trees and explores tourists' travel mode choices through three approaches: feature importance ranking, partial dependence plots and elasticity analysis.

5.1 Calibration of the gradient boosting decision tree model

This study employs grid search to optimise the hyperparameters of the gradient boosting decision tree (GBDT) model. Prior to training and testing the model, the dataset is partitioned into two subsets: a training set and a test set. In this study, the test set is allocated at a ratio of 0.3, meaning that the system randomly splits the 858 travel data points, with 600 used for training and 258 for evaluating the model's performance. To identify the optimal hyperparameters, the process involves four steps. First, the parameter range is adjusted, and a parameter grid is generated. Second, a grid search model is constructed, specifying the number of cross-validation folds, with accuracy as the evaluation metric. Third, the model is fitted, and the optimal parameter values and accuracy are recorded. Fourth, based on the optimal parameters, the parameter range is further refined, and a new parameter grid is generated. Steps one through four are repeated, comparing the accuracy obtained in each iteration. Finally, when the model's accuracy reaches its optimal value, parameter tuning ends, and the optimal hyperparameters are output. As a result, the gradient boosting decision tree model achieves an accuracy of 87.6%, a precision of 0.873 and a recall of 87.2% in predicting tourists' travel mode choice behaviour.

5.2 Feature importance

The gradient boosting decision tree model has made preliminary predictions about tourists' choice of transportation modes. Therefore, this section will further analyse the influencing factors of tourists' choice of transportation modes. To demonstrate the impact of each attribute on the choice of transportation mode, this section has created a table of feature variable importance and ranked the influencing factors within each attribute from highest to lowest importance, as shown in *Table 2*.

Table 2 – Feature importance

Attributes	Variable name	Feature importance	Importance ranking	Cumulative importance
Personal social attributes	Tour	0.0411	6	10.94%
	Income	0.0272	10	
	Education	0.0226	11	
	Purpose	0.0185	12	
Temporal attributes	T_transit	0.1657	2	41.04%
	T_departure	0.1465	3	
	T_free	0.0982	5	
Spatial attributes	Distance	0.2625	1	48.02%
	D_facility	0.1187	4	
	D_roadnetwork	0.0347	7	
	D_bilit	0.0345	8	
	D_landuse	0.0298	9	

As a whole, spatial attributes have the highest cumulative importance on tourists' travel mode choice behaviour, with 48.02%; followed by temporal attributes, with 41.04% cumulative importance; and lastly, personal social attributes, with 10.94% cumulative importance. Through observation, it is found that the cumulative importance of spatial attributes is the largest, but the characteristic variables that play a key role are travel distance and the number of bus stops. The cumulative importance of temporal attributes is the second largest, and there is not much difference in the importance of the variables; and personal social attributes have the smallest cumulative importance on the behaviour of travel mode choice, which proves that when tourists choose travel modes, the influence of personal attributes on the choice of behaviour is far less than the spatial and temporal constraints.

5.3 Intrinsic mechanisms between influencing factors and mode choice

This part will investigate the nonlinear relationships between variables and travel modes, clarifying the spatiotemporal threshold characteristics for different travel modes.

Univariate analysis

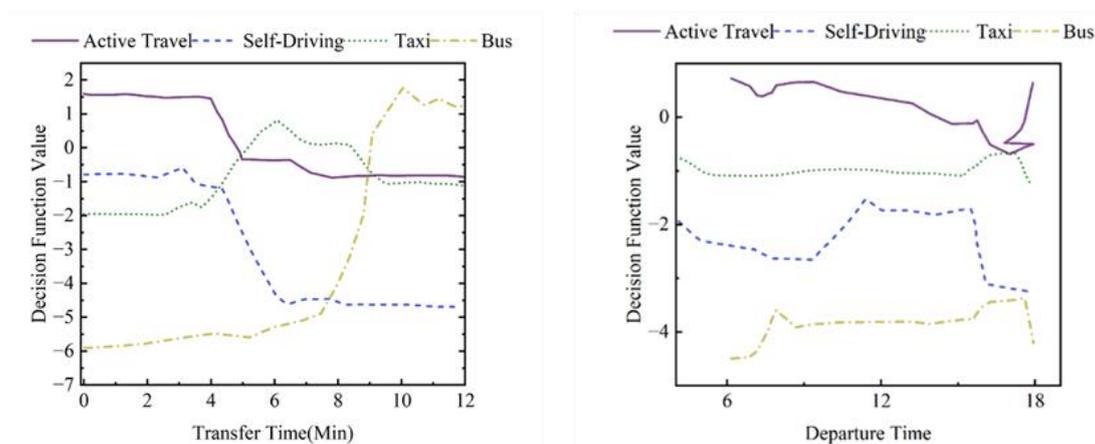1) Temporal factors and tourists' travel mode choice

Figure 2 – Non-linear relationship between temporal factors and tourists' travel mode choice

Figure 2 illustrates that when tourists choose active travel and driving as their modes of transportation, the decision function value is negatively correlated with transfer time. This means that as transfer time increases, tourists' inclination to choose these modes gradually decreases. Conversely, taxis and buses show a positive correlation with the decision function value at around 6 minutes and 10 minutes, respectively. This indicates that the transfer times for these two modes of transportation in the study area are concentrated around 6 minutes and 10 minutes. As transfer time increases, the probability of tourists choosing these modes of transportation also increases.

We also found that in the morning (06:00-11:00), the decision function values of the four travel modes are positively correlated with departure time, indicating that tourists prefer to travel in the morning. Between 17:00 and 18:00, the decision function values of driving, taxi and public transportation are negatively correlated with departure time, while the decision function value of active travel is positively correlated with departure time. This is because the temperature during this period is comfortable, increasing tourists' preference for active travel. Although there are some similarities in the change patterns of decision function values for each travel mode, differences can still be observed. For example, during the 06:00-11:00 period, the growth starting point of the decision function for active travel and public transportation is earlier than that for driving and taxis. During the 17:00-18:00 period, the decay starting point of the decision function for public transportation is later than that for driving and taxis. It is also found that tourists choosing active travel tend to travel during the day, with their departure times mainly concentrated between 05:00-10:00 and 17:00-18:00. Delaying departure time within these periods can continuously increase tourists' inclination towards active travel. Through the above analysis, it can be seen that although there is no obvious positive or negative correlation between departure time and travel mode, each travel mode has its most suitable departure time.

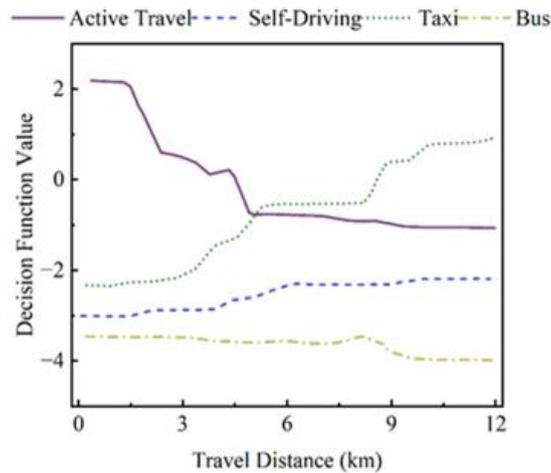2) Travel distance and tourists' choice of transportation mode

Figure 3 – Non-linear relationship between travel distance and tourists' travel mode choice

In *Figure 3*, there is a negative correlation between travel distance and the decision function values of active travel and public transportation, while there is a positive correlation with driving and taxis. Specifically, the decision function values for active travel continue to decline notably at travel distances between 1.5-2 km and 4-5 km. Observing the partial dependence plot for buses reveals that tourists prefer buses for short-distance travel. In general, there is a noticeable positive or negative correlation between travel mode and travel distance. Each travel mode exhibits different distance thresholds; for example, active travel shows a decline starting at around 5 km, while public transportation's preference is concentrated around 10 km. Moreover, driving and taxi services are preferred for medium to long distances, with tourists showing a preference for these options for longer-distance travel needs.

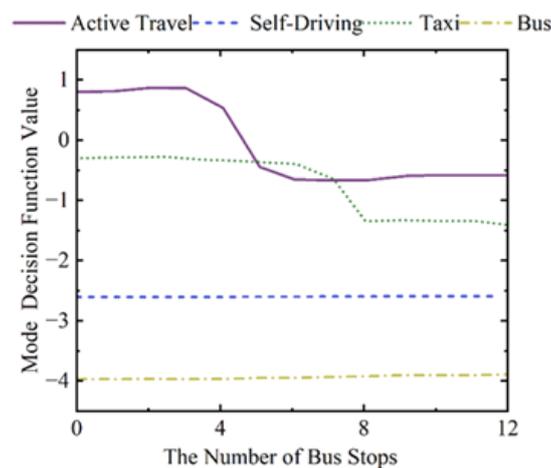3) The number of bus stops and tourists' choice of transportation decision function value

Figure 4 – Non-linear relationship between the number of bus stops and tourists' travel mode choice

In *Figure 4*, the impact of the number of bus stops on travel mode choices is evident. Tourists' preference for driving shows a slight overall growth trend in the decision function, with minimal fluctuations due to changes in the number of bus stops. There is a positive correlation overall between the number of bus stops and tourists choosing to travel by bus. Conversely, there is a negative correlation between the number of bus stops and the decision function values for active travel and taxis, with distinct thresholds observed.

From the analysis above, it is evident that buses exhibit an overall positive correlation with the number of bus stops, while active travel and taxis show a negative correlation. The number of bus stops causes slight fluctuations in the preference for driving, and it restricts the choice of active travel and taxis to varying degrees.

Dual-factor interaction analysis

    Partial dependence plots not only describe the non-linear relationship between feature variables and the dependent variable but also visually demonstrate how different feature variables correlate and interact with travel mode choice behaviour. This section focuses on selecting specific spatio-temporal factors to explore their combined interactive effects on travel mode choice.

Figure 5 – Cross-sectional dependency graph of departure time and travel distance

Figure 6 – Dependency graph of the intersection between the number of stations and departure time

1) The impact of the combination of departure time and travel distance on tourists' choice of travel mode

The first step involves combining departure time and travel distance to examine how tourists' travel mode preferences evolve under the combined influence of time and space. Using cross-partial dependence plots, different colours represent varying decision function values. Darker colours indicate lower preference for a travel mode, while lighter colours signify higher preference, reflecting a continuous increase in preference as the colour shifts from dark to light.

Upon observing *Figure 5*, it is evident that the decision function values of driving, taxi and public transportation show a positive correlation with the spatio-temporal combination of departure time and travel distance, while they exhibit a negative correlation with active travel. Additionally, the decision function values for public transportation and driving demonstrate smaller changes with increasing departure time and travel distance. In contrast, active travel and taxis display larger changes, indicating that the interaction between departure time and travel distance has a greater impact on preferences towards active travel and taxis. Furthermore, it is observed that travel distance has a more significant influence on the decision function values for active travel and taxis compared to departure time.

2) The combined effect of the number of bus stops and departure time on tourists' choice of travel mode

Upon observing *Figure 6*, it is evident that the number of bus stops and the increase in departure time show varying correlations with the decision function values of different travel modes. Specifically, they are negatively correlated with the decision function values of active travel and taxis, positively correlated with public transportation, and show no clear positive or negative correlation with driving. From the decision function values, it is apparent that public transportation is most influenced by this spatio-temporal combination, followed by active travel, while the impact on taxis is minimal. This combination also highlights the varying strengths of factors influencing travel mode preferences. For instance, the number of bus stops has a greater impact on active travel and taxis compared to departure time, whereas departure time has a stronger influence on driving.

5.4 Analysis of travel mode choice behaviour flexibility

The importance ranking of feature variables output by the gradient boosting decision tree model highlights that travel distance, departure time, transfer time and the number of bus stops significantly influence tourists' choice of travel mode. To further analyse the impact of these temporal and spatial features, these continuous variables were discretised. Using the elasticity analysis method, we explore how the probability of tourists choosing different travel modes changes for each unit increase in these feature variables. This study utilises Nlogit software to obtain the average elasticity values of temporal and spatial feature variables, as detailed in *Table 3*.

Table 3 – Average elasticity values of spatio-temporal influencing factors

Variable	Choice set	Elasticity value
Distance	Active travel	-1.1525
	Self-driving	1.2205
	Taxi	0.9601
	Bus	-1.1752
Departure time	Active travel	0.3472
	Self-driving	0.3105
	Taxi	0.1554
	Bus	0.6918
Transfer time	Active travel	-0.7289
	Self-driving	-2.4046
	Taxi	-0.3748
	Bus	1.2161
Number of bus stops	Active travel	-0.5939
	Self-driving	0.2764
	Taxi	-1.1929
	Bus	1.6701

According to the data in *Table 3*, it is observed that for every 1 km increase in travel distance, 1 min increase in transfer time and 1 unit increase in the number of bus stops within the study unit, the average selection probability of active travel decreases by 1.1525%, 0.7289% and 0.5939%, respectively. Conversely, for every 1-hour delay in departure time, the average selection probability of active travel increases by 0.3472%. The absolute elasticity value of travel distance is greater than 1, indicating that active travel is more sensitive to increases in travel distance. As transfer time and the number of bus stops within the study unit increase, the selection probability of active travel decreases. Furthermore, the average elasticity value of departure time suggests that a delay in departure time increases the selection probability of active travel to a certain extent, particularly during daytime hours.

6. DISCUSSION AND CONCLUSIONS

Through a comprehensive analysis of domestic and international scholars' research findings on temporal and spatial factors influencing tourists' choice of travel modes, this study has identified its research focus and direction. Utilising survey questionnaires and Python web scraping techniques, data were collected on tourists' social attributes, time-related factors and spatial attributes during their travel activities. Machine learning methods were employed to analyse tourists' travel mode choices. The main research findings are as follows:

1) The gradient boosting decision tree (GBDT) model demonstrates a strong predictive performance in forecasting tourists' travel mode choice behaviour. Tourists' travel mode choice behaviour can be regarded as a classification problem. Therefore, the GBDT algorithm, which is well-suited for classification tasks, is chosen as the foundational model for predicting tourists' travel mode selection behaviour. In predicting the travel mode choice behaviour of tourists at travel destinations, the GBDT model achieves an accuracy of 87.6%, a precision of 0.873 and a recall rate of 87.2%.

2) Temporal and spatial factors significantly outweigh personal social attributes in influencing tourists' travel mode choice behaviour. According to machine learning interpretability methods, the cumulative importance of temporal factors and spatial factors amounts to 41.04% and 48.02%, respectively, far surpassing that of personal social attributes. Among these factors, key influencers on tourists' travel mode choices are travelling distance, transfer time, departure time and the number of bus stops, with importance percentages of 26.25%, 16.57%, 14.65% and 11.87%, respectively. This underscores travel distance as the most impactful factor on tourists, which aligns with existing research findings.

3) There exists a complex relationship between temporal and spatial factors influencing tourists' choice of travel mode. Using partial dependence plots – a method in machine learning interpretability – it becomes evident that travel distance, transfer time, departure time and the number of bus stops demonstrate significant non-linear associations with various travel modes. These factors also reveal temporal and spatial thresholds. Moreover, examining cross-partial dependence plots across different temporal and spatial combinations reveals specific correlations: departure time and travel distance positively correlate with driving, taxi and public bus choices, while negatively correlating with active travel. Conversely, the number of bus stops and departure time negatively correlate with active travel and taxi usage, exhibit a negative correlation with public bus choice, and show no clear positive or negative correlation with driving.

4) Temporal and spatial factors exert varying influences on tourists' choice of travel mode. Using elasticity analysis, this study quantified how changes in these factors affect the probability of tourists selecting a particular travel mode. Employing Nlogit software, we conducted elasticity analysis on travel distance, transfer time, departure time and the number of bus stops. The resulting elasticity values indicate distinct sensitivities of each travel mode to different temporal and spatial variables.

Based on the above research, this paper will provide relevant recommendations from the perspectives of both the government and operating enterprises. For government departments, it is recommended to coordinate and strategically plan the transportation infrastructure and service network within tourist destinations. Specific measures include: (1) improving infrastructure construction to enhance the travel satisfaction index; (2) strengthening the integration of travel modes and building an integrated network; (3) accelerating the transition from gasoline to electric vehicles, advocating for energy conservation and emission reduction; and (4) strictly controlling private vehicle use to alleviate road congestion. For operating enterprises, optimising the operational outcomes of travel modes and improving resource utilisation should be prioritised. Specific measures include: (1) analysing tourist demand to plan routes and schedules; and (2) purchasing new vehicles to enhance operational efficiency.

The paper investigates factors influencing tourists' mode choice behaviour, encompassing temporal, spatial and personal-social attributes. While some conclusions have been drawn, the research process reveals shortcomings that necessitate further detailed examination and enhancement. Specifically, this study focuses solely on single-trip scenarios, omitting consideration of how tourists' complete travel chains influence mode choice behaviour. Additionally, by analysing communities as the unit of study, it becomes evident that the geographical natural environment within the research area somewhat diminishes the impact of built environment factors on tourists' mode choice behaviour. As a result, the influence of built environment factors in this context appears relatively limited.

ACKNOWLEDGEMENTS

REFERENCES

[1] Deng Y, Wang F, Wang K. Spatio-temporal changes of intercity travel network in urban agglomerations during China's National Day Holiday under COVID-19: Based on Baidu migration big data in 2019 and 2021. *Geography & Geographic Information Science*. 2023;39(6). DOI: 10.3969/j.issn.1672-0504.2023.06.009

[2] Wei S, Pan J. Spatiotemporal characteristics of residents' intercity travel in China under the impact of COVID-19 pandemic. *Cities*. 2024;152:15. DOI: 10.1016/j.cities.2024.105206

[3] Stroom M, Eichholtz P, Kok N. Avoiding crowded places during COVID-19: Common sense or a complex strategic decision? *Front Psychol*. 2021;12:700640. DOI: 10.3389/fpsyg.2021.700640

[4] China Daily. World welcomes China's resumption of outbound tourism. 2023 Jan 10. Available from: https://global.chinadaily.com.cn/a/202301/10/WS63bcd98ca31057c47eba8b77.html. [Accessed 10th Jan. 2023].

[5] Commuting monitoring report of major cities in China in 2022. *Urban and rural construction*. 2023;(2):56-65.

[6] He B, et al. Spatial and temporal characteristics of urban tourism travel by taxi—A case study of Shenzhen. *International Journal of Geo-Information*. 2021;10(7):445. DOI: 10.3390/ijgi10070445

[7] Lew A, McKercher B. Modeling tourist movements: A local destination analysis. *Annals of tourism research*. 2006;33(2):403-423. DOI: 10.1016/j.annals.2005.09.007

[8] Chen Y, Yin C, Sun B. Nonlinear associations of built environments around residences and workplaces with commuting satisfaction. *Transportation Research Part D: Transport and Environment*. 2024;133:17. DOI: 10.1016/j.trd.2024.104315

[9] Zhang M, He Y, Feng B. Study on tourist transportation mode choice behaviour in Beijing. *Transportation Research*. 2017;3(2):31-36. DOI: 10.1016/j.trs.2017.02.001

[10] Liu M, He Y, Sun X. Study on travel modes in historical districts based on the MNL model. *Transportation Research*. 2017;3(5):8-13. DOI: 10.1016/j.trs.2017.04.002

[11] Limtanakool N, Dijst M, Schwanen T. The influence of socioeconomic characteristics, land use and travel time considerations on mode choice for medium-and longer-distance trips. *Journal of Transport Geography*. 2006;14(5):327-341. DOI: 10.1016/j.jtrangeo.2006.01.005

[12] Guo S, et al. Analysis of commuting departure time choice behaviour under real-time bus information. *Journal of Transportation Information and Safety*. 2014;32(3):12-17.

[13] Yu J, et al. The issue of subway commuters' departure time choices under the influence of bike-sharing. *Journal of Advanced Transportation*. 2024;2024(1):2888275. DOI: 10.1155/2024/2888275

[14] Li ZC, Lam WHK, Wong SC. Bottleneck model revisited: an activity-based perspective. *Transportation research part B: methodological*. 2014;68:262-287. DOI: 10.1016/j.trb.2014.06.011

[15] Li X, Zhang Y, Du M. Analysis of travel decision-making for urban elderly healthcare activities under temporal and spatial constraints. *Sustainability*. 2018;10(5):1560. DOI: 10.3390/su10051560

[16] Fatima K, Moridpour S, Saghapour T. Spatial and temporal distribution of elderly public transport mode preference. *Sustainability*. 2021;13(9):4752. DOI: 10.3390/su13094752

[17] Ashalatha R, et al. Mode choice behaviour of commuters in Thiruvananthapuram city. *Journal of Transportation Engineering*. 2013;139(5):494-502. DOI: 10.1061/(ASCE)TE.1943-5436.0000526

[18] Liu S, Yamamoto T, Yao E. Joint modeling of mode choice and travel distance with intra-household interactions. *Transportation*. 2023;50(5):1527-1552. DOI: 10.1007/s11116-023-10394-9

[19] Ramezani S, et al. Shopping trip mode choice of older adults: An application of activity space and hybrid choice models in understanding the effects of built environment and personal goals. *Transportation*. 2021;48(2):505-536. DOI: 10.1007/s11116-020-10115-9

[20] Clark AF, Scott DM, Yiannakoulias N. Examining the relationship between active travel, weather, and the built environment: a multilevel approach using a GPS-enhanced dataset. *Transportation*. 2014;41:325-338. DOI: 10.1007/s11116-013-9476-5

[21] Wu J, et al. Active travel and the built environment: a theoretical model and multidimensional evidence. *Transportation research part D: transport and environment*. 2021;100:103029. DOI: 10.1016/j.trd.2021.103029

[22] Shi K, et al. The influence of ride-hailing on travel frequency and mode choice. *Transportation Research Part D: Transport and Environment*. 2021;101:103125. DOI: 10.1016/j.trd.2021.103125

[23] Wali B, et al. Developing policy thresholds for objectively measured environmental features to support active travel. *Transportation research part D: transport and environment*. 2021;90:102678. DOI: 10.1016/j.trd.2020.102678

[24] He M, et al. Research on nonlinear associations and interactions for short-distance travel mode choice of car users. *Journal of Advanced Transportation*. 2022;2022(1):8598320. DOI: 10.1155/2022/8598320

[25] Hatamzadeh Y, Hosseinzadeh A. Toward a deeper understanding of elderly walking for transport: An analysis across genders in a case study of Iran. *Journal of Transport & Health*, 2020, 19: 100949.

[26] Tang X, et al. Choice behaviour of tourism destination and travel mode: A case study of local residents in Hangzhou, China. *Journal of Transport Geography*. 2020;89:102895. DOI: 10.1016/j.jtrangeo.2020.102895

[27] Heinen E. Identity and travel behaviour: A cross-sectional study on commute mode choice and intention to change. *Transportation research part F: Traffic Psychology and Behaviour*. 2016;43:238-253. DOI: 10.1016/j.trf.2016.01.004

[28] Cheng L, et al. Examining non-linear built environment effects on elderly's walking: a random forest approach. *Transportation research part D: Transport and Environment*. 2020;88:102552. DOI: 10.1016/j.trd.2020.102552

[29] Gan Z, et al. Examining the relationship between built environment and metro ridership at station-to-station level. *Transportation Research Part D: Transport and Environment*. 2020;82:102332. DOI: 10.1016/j.trd.2020.102332

[30] Yang H, et al. Examining non-linear associations between built environments around workplace and adults' walking behaviour in Shanghai, China. *Transportation Research Part A: Policy and Practice*. 2022;155:234-246. DOI: 10.1016/j.tra.2021.11.019

[31] Friedman JH. Greedy function approximation: A gradient boosting machine. *Annals of Statistics*. 2001;29(5):1189-1232. DOI: 10.1214/aos/1013203451