

## Research Article

# Design of a Cultural Tourism Passenger Flow Prediction Model in the Yangtze River Delta Based on Regression Analysis

Jian Xu 

*School of Economics and Management, Hefei University, Hefei 230601, Anhui, China*

Correspondence should be addressed to Jian Xu; xujian@hfu.edu.cn

Received 15 March 2021; Revised 5 April 2021; Accepted 15 April 2021; Published 24 April 2021

Academic Editor: Shah Nazir

Copyright © 2021 Jian Xu. 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.

Cultural tourism has gained much attention in the last decade and has promoted the preservation of a variety of tangible and intangible assets of culture. In order to accurately predict the cultural tourism passenger flow in the Yangtze River Delta and improve its economic benefits, this paper designs the prediction model of cultural tourism passenger flow in the Yangtze River Delta based on regression analysis. Taking the competitiveness of passenger flow as the core, this paper selects 28 indexes from four aspects of cultural tourism brand resources, cultural tourism support and protection, and urban tourism market income to build the evaluation index system of influencing factors of passenger flow. The principal component analysis method is used to simplify many related factors into a few uncorrelated factors to eliminate the multicollinearity caused by too many dependent variables; on this basis, the principal component regression model is constructed, and the determination coefficient is used to test the model fitting. Taking 15 cultural tourism cities in the Yangtze River Delta as the research object, the results show that the designed model has a good fitting degree, and the average error is only 0.41%, which can meet the needs of the prediction of cultural tourism passenger flow in the Yangtze River Delta. After the application of the prediction model, the foreign exchange earning amount of each cultural tourism city can be increased by more than 12%. The study has revealed good results.

## 1. Introduction

The improvement of living standards makes the tourism industry gradually rise. Generally speaking, tourism cities have developed commercial economy, and there are many tourist attractions and classic snacks, which have attracted a large number of tourists. The development and prosperity of cultural tourism in the near future will be due to the dual role of industry consciousness and system design [1]. Promoting the consciousness, self-confidence, and self-improvement of regional cultural construction is not only the main tone of current cultural construction but also the fundamental driving force for the rapid development of cultural tourism [2]. Different cities may have the same mode and path of tourism development, but the connotation of urban cultural tourism is unique and competitive. The tourism development mode dominated by cultural tourism often has more vitality and durability.

The Yangtze River Delta region is an important growth pole of China's economic development and also one of the regions with the highest level of urbanization in China [3, 4]. As an important strategic base of China's tourism industry, the Yangtze River Delta region is not only the main force and vanguard of China's tourism development but also the most successful region of China's cross regional tourism cooperation and the region with great potential to become a world-class tourism destination. In the process of implementing the national strategy of regional integration in the Yangtze River Delta, cultural tourism needs to play a greater role. The Yangtze River delta needs a complete image in the process of representing China to participate in global competition and cooperation. Cultural tourism is the advance force and core backbone of regional international image publicity. At present, regional integration has become a consensus of China's future development. Through resource sharing, complementary advantages, and market

interaction, we can break the regional, spatial, and institutional barriers, create barrier-free tourism areas, and realize the all-round development of social economy. At the same time, the Yangtze River Delta region is rich in tourism resources, with a complete range of natural and cultural landscapes. It is an important concentrated distribution area of tourism resources in China and has the inherent advantages of developing regional tourism integration. Therefore, it has attracted a large number of tourists to travel all the time.

Tourist flow refers to the number of people and the flow pattern of tourists from the source to the destination [5]. According to the different travel time, it can be divided into four types: daily passenger flow, monthly passenger flow, quarterly passenger flow, and annual passenger flow. Tourist flow is an important indicator of the development level of tourism industry, an important part of tourism planning by national or regional tourism authorities, an effective guarantee to improve the quality of tourism products, and an important basis for the development of tourism resources and the construction of reception facilities such as hotels. Accurate prediction of tourist flow is related to the successful operation of an international and regional tourism project, which will directly affect the scientific decision making of the tourism project and is an important part of urban tourism development planning [6, 7].

There are many factors that affect the cultural tourism passenger flow in the Yangtze River Delta, among which economy, politics, education level, resources, and transportation will have an impact on it [8]. Taking transportation as an example, the places with convenient transportation with the Yangtze River Delta will have a large tourist flow to the Yangtze River Delta. In addition, psychological factors and personal preferences will also have an impact on the tourist flow of the Yangtze River Delta. Therefore, many factors must be fully considered in the process of analyzing the tourist flow in the Yangtze River Delta.

Scientific and reasonable prediction of tourist flow is of great guiding significance for the efficient use of tourism resources and local economic development. It can also help the government to formulate tourism development planning and tourism emergency plan, improve the quality and level of tourism service, and then, improve the tourist satisfaction and sense of experience [9]. The interdisciplinary research on "tourism passenger flow" and "passenger flow prediction" has developed rapidly. It has gone deep into systems science, computer science and technology, and other disciplines and has derived a number of interdisciplinary themes.

At present, the methods of predicting tourist flow are mainly divided into quantitative prediction and qualitative prediction. Qualitative prediction is generally based on qualitative analysis combined with empirical judgment [10], with low prediction accuracy. Quantitative prediction is mainly to establish a quantitative prediction model through mathematical methods. It is a widely used prediction method with high prediction accuracy. There are three kinds of models for tourism passenger flow prediction by using quantitative prediction methods, namely, the permeability

model, gravity model, and GM (1, 1) model. The penetration model is an intuitive model with strong subjectivity [11], which is based on the interviewees' willingness to visit as the main data, combined with the population base and carrying coefficient to make an intuitive inference of the passenger flow of tourist attractions. However, the model has its own regional bias due to the willingness to visit, so the permeability model is only used for interval estimation of the willingness to visit and the passenger flow. The gravity model is a commonly used international method to predict passenger flow under normal conditions [12]. However, in the actual prediction process, this method only considers a single or a small number of factors affecting passenger volume, which is lack of comprehensiveness, resulting in biased passenger volume prediction. In the practical application of the GM (1, 1) model, too many factors may be considered, resulting in multicollinearity among factors [13], and the regression coefficient cannot pass the significance test, and even the sign of some regression coefficients is inconsistent with the actual economic significance.

In view of the problems existing in the abovementioned models, a prediction model of cultural tourism passenger flow in the Yangtze River Delta based on regression analysis is designed on the basis of the gravity model and permeability model. In the parameter selection of the multiple regression model, the model comprehensively considers many factors that affect the cultural tourism passenger flow in the Yangtze River Delta, so compared with the gravity model and GM (1, 1) model, it considers the factors affecting passenger flow more comprehensively, overcomes the influence of multicollinearity, and improves the accuracy of passenger volume prediction.

The organization of the paper is as follows: the materials and methods of the paper are presented in Section 2 with details. Section 3 of this paper shows the results of the paper. The conclusion of this paper is given in Section 4.

## 2. Materials and Methods

The following sections briefly present the materials and methods used in this study.

*2.1. Index Selection.* Compared with other traditional types of industries or industries, cultural tourism has its own unique characteristics in the development connotation, development context, and social effects [14, 15]. First of all, in the world, the government vigorously promotes the cultural tourism industry, making it an increasingly important way to cultivate national cultural identity and national identity. Cultural tourism shoulders the sacred function of nurturing citizens. Secondly, the system structure and operation mechanism of cultural tourism embody the characteristics of the binary compound system of culture and tourism integration. The correct understanding of the concept of culture and the reasonable mining of tourism elements in the cultural field are the primary problems of cultural tourism transformation from resources to products, which provides an important reference for identifying the

elements of cultural tourism competitiveness. Thirdly, cultural tourism reflects the close interaction between the host and the guest. Cultural tourism space is the combination of geographical space, cultural space, and social space of cultural tourism destination. It is a field of interaction between the host and the guest with clear geographical space, which also defines a clear spatial scope for clarifying the system of cultural tourism competitiveness. Therefore, the development of theory and practice shows that the competitiveness of urban cultural tourism passenger flow in the Yangtze River Delta can carry out more scientific index screening and system construction.

Based on the theory of cultural tourism passenger flow competitiveness in the Yangtze River Delta region, the following three principles are considered and followed in the design process of the specific factor index:

- (i) The combination of comprehensiveness and operability: this paper selects and analyzes the connotation of cultural tourism development, the structure of tourist flow competitiveness, and the influencing factors in the Yangtze River Delta, so as to present the overall situation of tourist flow competitiveness of cultural tourism in the Yangtze River Delta as far as possible. At the same time, the feasibility and reliability of the index data sources are fully considered. According to the different grades of the Yangtze River Delta, in the design of the index system, besides quoting some necessary total indexes reflecting the scale effect, the strength indexes are considered as much as possible.
- (ii) Systematic and hierarchical: by fully combining the competitiveness of cultural tourism passenger flow in the Yangtze River Delta with the regional reality, the differences and emphases of cultural and tourism industry in different regions of the Yangtze River Delta are very different. In the evaluation, we should consider the problems from the actual situation of different regions as far as possible, so that the differences can be reflected in the evaluation index system. The hierarchy is reflected in that the competitiveness of cultural tourism passenger flow is mainly composed of three-level indexes. Besides the general goal and target decomposition level, it is explained and evaluated by element indexes [16].
- (iii) The evaluation objectives and methods are consistent. This paper focuses on the theme of the evaluation of cultural tourism passenger flow competitiveness in the Yangtze River Delta region, designs the indexes around the center as far as possible, and always defines the direction and goal of the evaluation system. In terms of evaluation methods, after solving the basic problems through qualitative methods, we use scientific mathematical methods to calculate and screen the indexes needed by the research according to the quantitative relationship of each indicator and then carry out comprehensive measurement and evaluation.

Based on the abovementioned competitiveness theory and index system design principles, this paper decomposes the evaluation objectives of cultural tourism passenger flow competitiveness in the Yangtze River Delta region into four subobjectives: cultural tourism brand resources, cultural performance and creativity, cultural tourism support and protection, and urban tourism market income. At the same time, 28 quantitative indexes are designed to form the Yangtze River Delta region's cultural tourism. The evaluation index system of influencing factors of passenger flow is shown in Table 1.

**2.2. Construction of the Prediction Model.** Before the establishment of the model, the passenger volume is  $y$ , the sample size is  $n$ , and the observation value of the  $i$ -th index is  $x_i$  ( $i = 1, 2, \dots, 27$ ).

**2.2.1. Multicollinearity Diagnosis.** The dependent variable  $y$  and the independent variable  $x_1, x_2, \dots, x_{27}$  are used to establish the regression model. SPSS statistical software is used to select the variables by the backward regression method, and  $p$  variables are set to enter the regression analysis model, which are  $x_{s1}, x_{s2}, \dots, x_{sp}$  ( $p \leq n$ ), called initialization variables.

From the regression results to find the variance expansion factor (VIF), if  $VIF_i \geq 10$ , it means that there is a serious multicollinearity between the independent variables.

Multicollinearity has a great influence on the regression coefficient [17], which can be processed by principal component analysis.

**2.2.2. Principal Component Analysis.**

(1) Data standardization

In order to eliminate the influence of different orders of magnitude and dimensions, it is necessary to standardize the original data. The standardized formula is as follows:

$$x'_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}, \quad (1)$$

where  $x'_{ij}$  is the standardized data and  $\bar{x}_j$  and  $(j = 1, 2, \dots, n)$  represent the mean value and standard deviation of the  $j$ -th index sample, respectively.

(2) Calculation of the correlation coefficient matrix: after processing the original data, the standardized data matrix  $(x'_{ij})_{p \times n}$  is obtained, and the corresponding correlation coefficient matrix is calculated.

$$R = (r_{ij})_{n \times n} \quad (2)$$

In formula (2),  $R$  is a symmetric matrix of order  $n$ .

(3) The eigenvalues and eigenvectors of correlation coefficient matrix  $R$  are calculated.

The eigenvalue  $\lambda_i$  ( $i = 1, 2, \dots, n$ ) of  $R$  and its corresponding eigenvector  $u_i$  ( $i = 1, 2, \dots, n$ ) are solved,

TABLE 1: Evaluation index system of influencing factors of cultural tourism passenger flow in the Yangtze River Delta.

General objective	Target decomposition	Element description
Influencing factors of cultural tourism passenger flow in the Yangtze River Delta	Brand resources of cultural tourism	Number of world cultural heritages
		Number/item of abnormal intangible culture at the national level or above
		Number of national key cultural relics protection units
		Number of famous historical and cultural cities in China
		Number of famous historical and cultural towns and villages in China
		Number of 4A tourist attractions
		Number of red tourist attractions
		Number of local folk festivals and special activities
		Creative tourism performance project
		Number of provincial and national cultural industry demonstration bases
	Cultural performance and creative tourism	Number of art performing groups
		Number of cultural centers, art performance centers, and mass art centers
		Number of libraries
		Number of open museums
Cultural tourism support	Number of accommodation facilities	
	Number of beds	
	Number of rooms	
	Number of travel agencies	
	Number of food blocks	
	Number of passengers sent by civil aviation	
Regional tourism market revenue	Scale of highway passenger transport	
	The scale of railway passengers	
	Number of inbound tourists	
	Domestic tourists	
	Foreign exchange earned by tourism	
	Total tourism revenue	
	The proportion of total tourism revenue in total economic output	

and  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ , where  $\lambda_i$  is the variance of the main component  $F_i$ , and the greater the variance is, the greater the contribution to the total variance is [18].

- (4) The contribution rate is calculated, and the principal component is determined.

Formula (3) is defined as the contribution rate of main component  $F_i$ .

$$e_i = \frac{\lambda_i}{\sum_{i=1}^n \lambda_i} \times 100\%. \quad (3)$$

In formula (3),  $\sum_{i=1}^n e_i$  is the contribution rate of cumulative variance. Generally,  $m$  principal components with  $\sum_{i=1}^m e_i$  greater than or equal to 85% are selected for comprehensive analysis. Therefore,  $n$  factors are reduced to  $m$  principal components, and the main factors are selected [19].

2.2.3. *Principal Component Regression Model.* Multiple regression analysis is conducted between  $F_r$  and dependent

variable  $Y$  (standardized value of  $y$ ) to obtain the standardized regression equation.

$$Y = \sum_{r=1}^m B_r F_r, \quad (4)$$

In formula (4),  $B_r$  is the standardized partial regression coefficient of the  $r$ -th principal component  $F_r$ .

By synthesizing formula (2), (3), and (4), the standardized regression equation is obtained.

$$Y = \sum_{j=1}^p b'_j X_{s_j}, \quad (5)$$

In formula (5),  $b'_j$  is the  $j$ -th standardized partial regression coefficient of the standardized regression equation.

Through the abovementioned analysis, we can get the principal component regression model as follows:

$$y = \sum_{j=1}^p b_j x_{s_j} + b_0, \quad (6)$$

where

$$b_j = b'_j \sqrt{\frac{L_{yy}}{L_{x_{ij}x_{ij}}}}, \quad (7)$$

$$b_0 = \bar{y} - \sum_{j=1}^p b_j \bar{x}_{s_j}. \quad (8)$$

In the abovementioned formula,  $b_j$  is the  $j$ -th partial regression coefficient of the general linear regression equation;  $L_{yy}$  is the sum of squares of deviation of  $y$ ;  $L_{x_{ij}}$  is the sum of squares of deviation of  $x_{s_j}$ ;  $\bar{y}$  is the mean value of  $y$ ;  $\bar{x}_{s_j}$  is the mean value of  $x_{s_j}$ ; and  $b_0$  is the constant of the general linear regression equation.

By substituting the observed values of each index in the forecast year into formula (6), the cultural tourism passenger flow of the Yangtze River Delta in the forecast year can be obtained.

**2.3. Model Test.** The determination coefficient  $R^2$  is used to test the model fitting. The formula for determining coefficient  $R^2$  is as follows:

$$R^2 = 1 - \frac{\sum e_t^2}{\sum (y_t - \bar{y})^2}. \quad (9)$$

In formula (9),  $\sum e_t^2$  is the sum of squares of residuals.

In the multiple linear regression model, the number of variables in each regression model may not be the same [20]. It is not appropriate to use the size of  $R^2$  as a measure of fitting quality. Therefore, the coefficient of determination  $R^2$  of modified degrees of freedom is often used. The calculation formula is as follows:

$$\bar{R}^2 = 1 - \frac{((\sum e_t^2)/n - p)}{((\sum (y_t - \bar{y})^2)/n - 1)} = 1 - \frac{n-1}{n-p} (1 - R^2). \quad (10)$$

In formula (10),  $n$  is the sample size, and  $p$  is the number of regression coefficients.

### 3. Results

**3.1. Case Profile and Data Sources.** The Yangtze River Delta is an urban sprawling area with a high level of tourism development in China. It has rich cultural resources and profound cultural heritage. In view of the different regional coverage of the Yangtze River Delta, this paper takes the core area of the Yangtze River Delta as the research object in the reply of the State Council on the regional planning of the Yangtze River Delta (Guo Han (2010), No. 38), that is, taking Shanghai as the leader, and Nanjing, Suzhou, Wuxi, Changzhou, Zhenjiang, Nantong, Yangzhou, and Taizhou in Jiangsu Province, Hangzhou, Ningbo, and Huzhou in Zhejiang Province, and Jiaxing, Shaoxing, Zhoushan, and Taizhou, a total of 16 cities as both wings. This paper selects the 10-year statistical data from 2008 to 2018 as an example,

and the data come from the statistical yearbooks and statistical bulletins of various cities, including the China Urban Statistical Yearbook, China Regional Economic Statistical Yearbook, China Tourism Statistical Yearbook, China Tourism Yearbook, List of Top 100 Travel Agencies, and List of National Star Hotels. Data collection adheres to the principle of combining scientificity, authority, standardization, and data availability. When the data from different channels are not unified, the higher-level government department shall prevail.

**3.2. Data Processing.** Since the data collected are from the statistical yearbooks and statistical bulletins of various cities, the time span is from 2008 to 2018. Due to the snow disaster, earthquake, financial crisis, and other major events in 2008, the cultural tourism reception in the Yangtze River Delta region is seriously affected, and the number of tourists decreased significantly. In order to improve the accuracy of the prediction model, the data of cultural tourism passenger flow in the Yangtze River Delta in 2008 and 2009 are revised.

Linear interpolation is used to correct the data.

Firstly, the starting year  $a_1$  and the ending year  $a_2$  which are suitable for linear interpolation are selected, and the passenger flow in the starting year and the ending year is expressed by  $y_1$  and  $y_2$ , respectively. The tolerance  $d$  is determined by formula (11).

$$d = \frac{y_2 - y_1}{a_2 - a_1}. \quad (11)$$

The correction value is calculated by the interpolation equation, and the calculation formula is as follows:

$$y_n = y_1 + (a - a_1)d. \quad (12)$$

In formula (12),  $n$  is the year to be corrected and  $y_n$  is the correction value of the  $n$ -th year.

According to the abovementioned method, the starting year of linear interpolation is 2010, and the ending year is 2013. By substituting the data of these two years into formula (11) and formula (12), we can get the revised value of cultural tourism passenger flow in the Yangtze River Delta in 2008 and 2009. Table 2 shows the revised tourist flow of cultural tourism in the Yangtze River Delta from 2008 to 2018.

**3.3. Multicollinearity Diagnosis.** Taking the statistical data of the Yangtze River Delta from 2008 to 2018 as an example, this paper makes regression analysis on the cultural tourism passenger flow and 28 influencing factors in the Yangtze River Delta and uses SPSS statistical software to select variables by the backward regression method. Finally, there are 9 variables  $X_7, X_8, X_{15}, X_{16}, X_{17}, X_{18}, X_{20}, X_{22}$ , and  $X_{23}$  to enter into the regression model, and the VIF values are greater than 10, which indicates that there is still a serious collinear relationship between variables. Therefore, it is necessary to use principal component regression to simplify the analysis. On the premise of retaining all or most of the original information, the abovementioned interrelated variables are transformed into a few independent or

TABLE 2: Revised tourist flow of cultural tourism in the Yangtze River Delta from 2008 to 2018.

Year/year	Passenger flow/10000 person times
2008	1090.09
2009	1110.62
2010	1128.60
2011	1175.27
2012	1241.94
2013	1268.52
2014	1364.39
2015	1392.41
2016	1379.66
2017	1485.59
2018	1590.33

unrelated variables, and then, these variables are integrated to establish a regression model.

Principal component analysis is performed on  $X_7, X_8, X_{15}, X_{16}, X_{17}, X_{18}, X_{20}, X_{22},$  and  $X_{23}$  variables. Using SPSS statistical software, the eigenvalues and eigenvectors are obtained. The cumulative variance contribution rate of the first two eigenvalues has reached 91.041%. It is generally believed that the effective information can be retained when the cumulative contribution rate of principal components reaches 85%. Therefore, this paper only needs to take the first two principal components to reflect most of the information of all indexes:

$$F_1 = 0.149X_7 + 0.180X_8 + 0.227X_{15} - 0.217X_{16} - 0.204X_{17} + 0.199X_{18} + 0.179X_{20} - 0.179X_{22} - 0.210X_{23}, \quad (13)$$

$$F_2 = 0.455X_7 - 0.359X_8 + 0.060X_{15} + 0.137X_{16} + 0.245X_{17} + 0.219X_{18} + 0.342X_{20} + 0.289X_{22} - 0.059X_{23}. \quad (14)$$

The standardized regression equation is obtained by multiple regression analysis between the evaluation values of  $F_1$  and  $F_2$  and the dependent variable  $Y$ .

$$Y = 0.973F_1 + 0.113F_2 (R^2 = 0.972). \quad (15)$$

By substituting formula (10) and (13) into (14), the standardized regression equations for the standardized variables  $X_7, X_8, X_{15}, X_{16}, X_{17}, X_{18}, X_{20}, X_{22},$  and  $X_{23}$  are obtained.

$$Y = 0.204X_7 + 0.138X_8 + 0.230X_{15} - 0.200X_{16} - 0.159X_{17} + 0.219X_{18} + 0.220X_{20} - 0.142X_{22} - 0.215X_{23}. \quad (16)$$

Using formula (7), the principal component regression model is obtained as follows:

$$y = 1596.733x_7 + 964.461x_8 + 5.013x_{15} - 576.898x_{16} + 529.266x_{17} + 0.164x_{18} + 7.999x_{20} - 392158.897x_{22} - 1478326.246x_{23} - 158927.732. \quad (17)$$

**3.4. Model Validation.** Through the test formula of formula (8) and (9), it can get  $R^2 = 0.984$ ;  $\bar{R}^2 = 0.859$ , showing that formula (17) has a high degree of fit and can make a reasonable forecast of passenger volume. By substituting the values of  $X_7, X_8, X_{15}, X_{16}, X_{17}, X_{18}, X_{20}, X_{22},$  and  $X_{23}$  from 2008 to 2018 into formula (17), the predicted passenger volume of each year can be obtained. The comparison between the predicted value and the actual value (Table 2) is shown in Figure 1.

It can also be seen from Figure 1 that the predicted value of the principal component regression model has a good fit with the actual value, with the highest error of 1.23%, the lowest error of 0.01%, and the average error of 10 years is only 0.41%, which can basically meet the needs of the prediction of cultural tourism passenger flow in the Yangtze River Delta.

**3.5. Model Performance Analysis.** In order to analyze the application performance of the model, this paper uses the method to predict the cultural tourism passenger flow in the Yangtze River Delta (Shanghai as the leader, Nanjing, Suzhou, Wuxi, Changzhou, Zhenjiang, Yangzhou, and Taizhou in Jiangsu Province and Hangzhou, Ningbo, Huzhou, Jiaying, Shaoxing, Zhoushan, and Taizhou in Zhejiang Province, a total of 15 cities as both wings) in 2020. MAPE (average absolute percentage error, which can directly reflect the pros and cons of the prediction effect), MAE (average absolute error, the smaller the value is, the smaller the error is), RMSE (root mean square error, the smaller the value is, the smaller the error is), and EC (equalization coefficient, the higher the value is, the higher the fitting degree is) are taken as evaluation indexes to verify the performance of the model. The results are shown in Table 3. The calculation formula of each evaluation index is as follows:

$$\text{MAPE} = \frac{1}{V} \sum_t \left| \frac{C_p(t) - C_r(t)}{C_r(t)} \right| \times 100\%, \quad (18)$$

$$\text{MAE} = \frac{1}{V} \sum_t |C_p(t) - C_r(t)|, \quad (19)$$

$$\text{RMSE} = \sqrt{\frac{\sum_t (C_p(t) - C_r(t))^2}{V}}, \quad (20)$$

$$\text{EC} = 1 - \frac{\sqrt{\sum_t (C_p(t) - C_r(t))^2}}{\sqrt{\sum_t (C_p(t))^2} + \sqrt{\sum_t (C_r(t))^2}}. \quad (21)$$

In the abovementioned formula,  $C_p(t)$ ,  $C_r(t)$ , and  $V$  are the predicted output value, the measured value of cultural tourism passenger flow of each city in 2020, and the number of predicted samples, respectively. The prediction accuracy of the proposed model and the comparison model (the prediction model based on the gravity model and the prediction model based on GM (1, 1)) are compared fairly to the maximum extent by using the abovementioned indexes. Prediction results of this model are shown in Table 3.

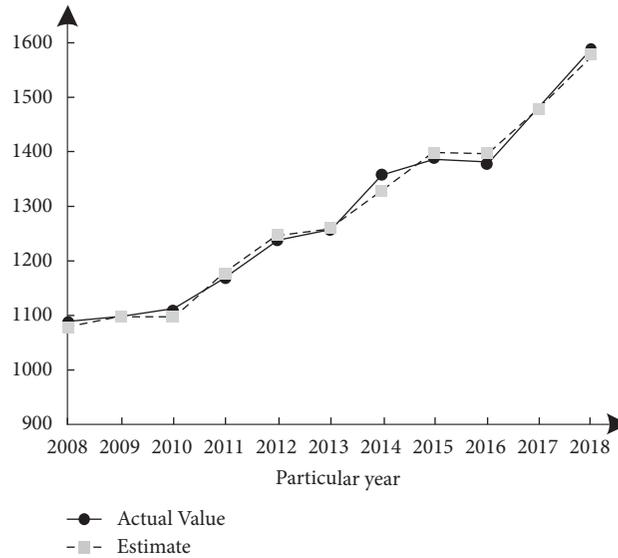


FIGURE 1: Comparison between the predicted value and actual value.

TABLE 3: Prediction results of this model.

Cultural tourism cities in the Yangtze River Delta	MAPE	MAE	RMSE	EC
Shanghai	0.57	1.26	2.05	0.97
Nanjing	0.60	1.14	2.05	0.98
Suzhou	0.62	1.30	2.01	0.98
Wuxi	0.35	1.06	1.65	0.98
Changzhou	0.51	1.10	1.79	0.98
Zhenjiang	0.50	1.05	1.81	0.99
Yangzhou	0.40	1.02	1.69	0.98
Taizhou	0.37	1.05	1.77	0.98
Hanzhou	0.69	1.14	1.96	0.97
Ningbo	0.43	1.12	1.72	0.98
Huzhou	0.49	1.11	1.82	0.99
Jiaxing	0.39	1.09	1.93	0.97
Shaoxing	0.44	1.13	2.00	0.98
Zhoushan	0.51	1.06	1.91	0.98
Taizhou	0.56	1.11	2.02	0.98
Average value	0.495	1.116	1.879	0.979

The average values of each index obtained by the algorithm in Table 3 are compared with those obtained by the two comparative prediction models, and the results are shown in Table 4.

After analyzing the prediction performance of the three prediction models in Table 4, it is found that the evaluation indexes of the proposed model are significantly better than those of the two comparison models, and the MAPE is reduced by more than 50 compared with the comparison model, which indicates that this model has higher prediction accuracy compared with the comparison model.

The running time (the larger the value is, the higher the complexity of the method is) is used as the evaluation index of prediction performance. The running time of this model and the two comparison models in the evaluation process is compared, and the results are shown in Table 5.

Analysis of Table 5 shows that the running time of the proposed model and the GM (1, 1) model is significantly

better than that of the gravity model. The average running time of the proposed model is 0.71 s, and the average running time of the GM (1, 1) model is 0.74 s. There is no significant difference between the two models. Combined with the data in Table 4, it can be concluded that the proposed model has a significant performance advantage compared with the comparison model.

**3.6. Application Test.** This paper compares the amount of foreign exchange earned by cultural tourism cities in the Yangtze River Delta after using this model to predict the passenger flow with the amount of foreign exchange earned before using this model (the average of the previous three years is the standard value). The results are shown in Figure 2, in which the numbers 1–15 represent the cultural tourism cities in the Yangtze River Delta.

According to the analysis of Figure 2, after using the proposed model to predict the passenger flow, the amount of

TABLE 4: Comparison results of prediction performance of different prediction models.

Prediction model	MAPE	MAE	RMSE	EC
Model of this paper	0.495	1.116	1.879	0.979
Prediction model based on the gravity model	1.326	2.735	2.789	0.962
Prediction model based on the GM (1, 1) model	0.866	2.003	2.104	0.941

TABLE 5: Running time of different prediction models.

Cultural tourism cities in the Yangtze River Delta	Model of this paper	Prediction model based on the gravity model	Prediction model based on the GM (1,1) model
Shanghai	0.71	1.05	0.58
Nanjing	0.74	1.16	0.92
Suzhou	0.77	1.13	0.84
Wuxi	0.70	1.08	0.73
Changzhou	0.66	1.15	0.81
Zhenjiang	0.69	1.02	0.69
Yangzhou	0.72	0.99	0.75
Taizhou	0.73	1.15	0.68
Hanzhou	0.69	1.03	0.75
Ningbo	0.71	1.01	0.72
Huzhou	0.70	0.97	0.67
Jiaxing	0.68	1.08	0.73
Shaoxing	0.68	1.04	0.64
Zhoushan	0.71	1.03	0.79
Taizhou	0.72	1.10	0.80

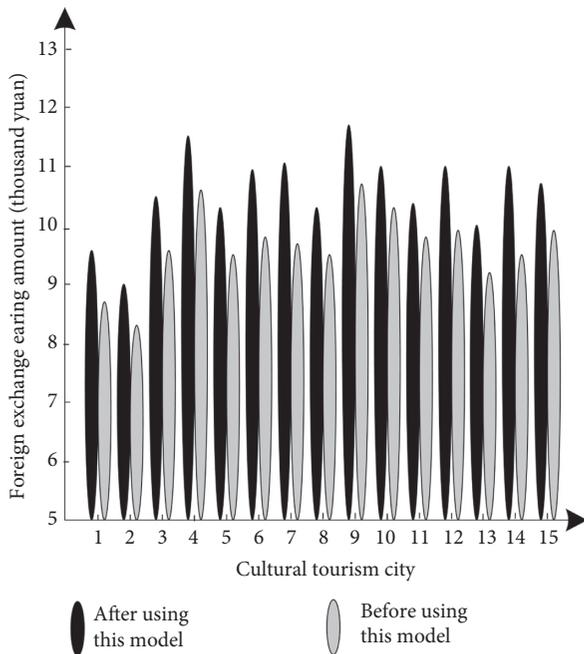


FIGURE 2: System sales comparison.

foreign exchange earnings of each city shows different increases, and the increase rate is maintained at more than

12%, which shows that the economic benefits of cultural tourism in the Yangtze River Delta can be significantly improved by using this model.

#### 4. Conclusions

Cultural tourism has gained much attention in the last decade and has promoted the preservation of a variety of tangible and intangible assets of culture. The applications of cultural assets for development of tourism has generated various debates such as the matter of whether the intangible values of cultural assets including those of their education, aesthetics, and history can properly be carried out in order to attract tourist. In order to accurately predict the cultural tourism passenger flow in the Yangtze River Delta and improve its economic benefits, this paper designs the prediction model of cultural tourism passenger flow in the Yangtze River Delta based on regression analysis. This paper considers many factors that affect the cultural tourism passenger flow and designs a prediction model of cultural tourism passenger flow in the Yangtze River Delta based on regression analysis. Through the experimental analysis, this model can accurately predict the cultural tourism passenger flow in the Yangtze River Delta in the process of practical application, and significantly improve the cultural tourism amount of foreign exchange in the Yangtze River Delta.

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was supported by University Humanities and Social Sciences Research Project of Anhui Province “Research on the high quality development mode of culture and tourism in the process of Yangtze River Delta integration,” Project No. SK2020A0431..

## References

- [1] X. Wang, E. Yao, and S. Liu, “Urban rail transit passenger flow forecasting for large special event based on a/c data,” *Beijing Jiaotong Daxue Xuebao*, vol. 42, no. 1, pp. 87–93, 2018.
- [2] C. Lin, K. Wang, D. Wu, and B. Gong, “Passenger flow prediction based on land use around metro stations: a case study,” *Sustainability*, vol. 12, no. 17, p. 6844, 2020.
- [3] X. Liu, L. Li, X. Liu, and T. Zhang, “Analysis of passenger flow and its influences on hvac systems: an agent based simulation in a Chinese hub airport terminal,” *Building and Environment*, vol. 154, pp. 55–67, 2019.
- [4] K. Mbbchir, R. Statistics, C. G. Hons, and M. Alain Vuylstekec, “Prediction of patient length of stay on the intensive care unit following cardiac surgery: a logistic regression analysis based on the cardiac operative mortality risk calculator, euroscore,” *Journal of Cardiothoracic and Vascular Anesthesia*, vol. 32, no. 6, pp. 2676–2682, 2018.
- [5] Y. Liu, C. Lyu, X. Liu, and Z. Liu, “Automatic feature engineering for bus passenger flow prediction based on modular convolutional neural network,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 99, pp. 1–10, 2020.
- [6] W. Zhou, W. Wang, and D. Zhao, “Passenger flow forecasting in metro transfer station based on the combination of singular spectrum analysis and adaboost-weighted extreme learning machine,” *Sensors*, vol. 20, no. 12, p. 3555, 2020.
- [7] S. Y. Song, “Analysis and prediction of passenger flow differences in shanghai 3a scenic spots and above based on arima model,” *Statistics and Applications*, vol. 08, no. 3, pp. 537–552, 2019.
- [8] H. Chen, B. Wang, W. He, and J. Zheng, “Research on passenger flow early warning of urban rail transit station based on system dynamics,” *MATEC Web of Conferences*, vol. 308, no. 2, p. 01003, 2020.
- [9] C. Li, “Combined forecasting of civil aviation passenger volume based on arima-regression,” *International Journal of System Assurance Engineering and Management*, vol. 10, no. 5, pp. 945–952, 2019.
- [10] F. Wang, H. Y. Liu, C. F. Shao, and J. J. Zhang, “Prediction model of public transport vehicle allocation based on multiple regression analysis,” *Advances in Transportation Studies*, vol. 3, pp. 69–78, 2018.
- [11] Y. Liu, C. Liu, and Z. Zheng, “Traffic congestion and duration prediction model based on regression analysis and survival analysis,” *Open Journal of Business and Management*, vol. 08, no. 02, pp. 943–959, 2020.
- [12] L. Z. Jing, Q. S. Li, J. L. Xu, X. L. Jia, and Y. J. Han, “Average travel time prediction model in basic expressway sections based on v/c ratio and truck percentage,” *Chang’an Daxue Xuebao (Ziran Kexue Ban)/Journal of Chang’an University (Natural Science Edition)*, vol. 38, no. 5, pp. 106–113, 2018.
- [13] T. Wu, P. Zhang, J. Qin, D. Wu, and Y. Wan, “A flood-discharge-based spatio-temporal diffusion method for multi-target traffic hotness construction from trajectory data,” *IEEE Access*, vol. 99, pp. 1–2, 2020.
- [14] L. Liu, R.-C. Chen, and S. Zhu, “Impacts of weather on short-term metro passenger flow forecasting using a deep lstm neural network,” *Applied Sciences*, vol. 10, no. 8, p. 2962, 2020.
- [15] T.-Z. Chen, Y.-Y. Chen, and J.-H. Lai, “Estimating bus cross-sectional flow based on machine learning algorithm combined with wi-fi probe technology,” *Sensors*, vol. 21, no. 3, p. 844, 2021.
- [16] X. Chang, J. Wu, H. Sun, G. Wang, and X. Bao, “Understanding and predicting short-term passenger flow of station-free shared bike: a spatiotemporal deep learning approach,” *IEEE Intelligent Transportation Systems Magazine*, vol. 1, no. 99, pp. 1–3, 2021.
- [17] D. Li, L. Deng, and Z. Cai, “Statistical analysis of tourist flow in tourist spots based on big data platform and da-hkrvm algorithms,” *Personal and Ubiquitous Computing*, vol. 24, no. 1, pp. 87–101, 2020.
- [18] H. J. Heng and P. Ren, “Short-Term passenger flow forecasting for the check-in process in the airport based on time series,” *Computer Simulation*, vol. 37, no. 02, pp. 31–37, 2020.
- [19] W. Wang, Y. Wang, G. Correia, and Y. Chen, “A network-based model of passenger transfer flow between bus and metro: an application to the public transport system of Beijing,” *Journal of Advanced Transportation*, vol. 2020, no. 15, 4 pages, Article ID 6659931, 2020.
- [20] L. Qiu, H. Zhou, Z. Wang, S. Zhang, L. Zhang, and W. Lou, “High-speed elevator car air pressure compensation method based on coupling analysis of internal and external flow fields,” *Applied Sciences*, vol. 11, no. 4, p. 1700, 2021.