Data-Driven Spatial Econometric Analysis Model for Regional Tourism Development

Xuemei Zhao and Guofeng Yu
Anhui Business and Technology College, Hefei 231131, China

Correspondence should be addressed to Xuemei Zhao; 11261201@qq.com

Received 18 November 2020; Accepted 28 April 2021; Published 8 May 2021

1. Introduction

As a multifunctional and comprehensive industry, tourism is an economic issue, a livelihood issue, and an issue of social harmony. In recent years, Anhui province has been committed to becoming a major tourism province, and the tourism economy has become a new growth point for Anhui's economic development. In 2017, Anhui province received 5.49 million inbound tourists, a year-on-year increase of 13.07%; it received 626 million domestic tourists, a year-on-year increase of 19.88%, reaping a total tourism revenue of 619.7 billion yuan, a year-on-year increase of 25.64%. According to preliminary estimates, the added value of the tourism industry accounted for 6.6% and 16% of the province’s gross domestic product (GDP) and service industry, respectively, with contribution rates of 10.4% and 13.3%, respectively. Although the tourism economy in Anhui province has been developing rapidly, regional tourism development is not coordinated, and the pattern that some regions are developed whereas others are underdeveloped has not changed. Therefore, how to promote the coordinated development of regional tourism in Anhui province and improve the level of regional tourism development in Anhui is currently an important issue in the development of tourism.

By analyzing the evolution characteristics of the spatial pattern of tourism development in Anhui province and the main factors of spatial-temporal pattern evolution of tourism development, indicating that regional per capita tourism income in Anhui province has obvious positive spatial autocorrelation and relatively obvious local spatial cluster characteristics, it is helpful to summarize the evolution law, construct the spatial pattern of regional tourism development, and promote the development of regional tourism in Anhui province.
investigated the spatial distribution patterns of per capita tourism income in China’s 31 provinces and their evolution and found that China’s regional tourism economy exhibits a strong spatial dependence and a long-term polarization pattern of high-value agglomerations and low-value agglomerations at the local level. Wu [3] conducted a spatial correlation analysis on Chinese provincial tourism economic growth and demonstrated that the provincial tourism economic growth has a spatial dependence, in addition to estimating the contributions of capital and labour to tourism economic growth and the spatial spillover effect in the process of tourism economic growth using a spatial panel econometric model. Wang [4] conducted a spatial analysis of the linkage and difference in the tourism industry development in the Greater Pearl River Delta region and demonstrated that there is a negative spatial correlation in the overall tourism economy of the Delta region. Based on these past studies, in this study, we analyzed the spatial characteristics and influencing factors of the tourism industry of Anhui province by combining the current tourism situation in Anhui province to find appropriate countermeasures suitable for the development of tourism in Anhui Province and promote the healthy development of the tourism industry in the province.

About the tourism impact on regional economy and society, foreign scholars mainly focus on the applied aspects of the analysis. Yang and Fik [5] examined two types of spatial effects in regional tourism growth: spatial spillover and spatial heterogeneity. A spatial growth regression framework is used to model the growth in regional tourism and identify the economic and spatial factors that explain the variability in tourism growth across 342 prefectural-level cities in China from 2002 to 2010. Jackson and Murphy [6] investigated the applicability of cluster theory in supporting the movement from comparative advantage to competitive advantage for four regional towns located on the Murray River in Australia. Chaabouni [7] investigated tourism efficiency and its determinants using a two-stage double bootstrap approach for a global panel of 31 Chinese provinces over the period 2008–2013. Bias-corrected data envelopment analysis (DEA) efficiency scores were first calculated by employing the smoothed homogeneous bootstrapped procedure. They were then regressed on a set of explanatory variables using the double-truncated regression approach. For other background and applications in mathematical model aspects, see [8–22].

2. Research Methods and Data Sources

2.1. Descriptive Analysis. First of all, descriptive analysis is conducted from the time dimension, mainly using some commonly used indicators and coefficients to measure regional development differences. In this paper, the coefficient of variation is used for analysis. The coefficient of variation is a normalized measure of the degree of dispersion of probability distribution, which is defined as the ratio of standard deviation to average. The coefficient of variation eliminates the influence of data level and measurement unit, measures the relative discrete degree of data, and can be used to compare the discrete degree of different groups of data. The larger the coefficient of variation is, the greater the difference degree of variable value is. The smaller the coefficient of variation is, the smaller the difference degree of variable value is. In practical application, the coefficient of variation is generally used to compare the discrete degree of different groups of data:

$$V_\sigma = \frac{\sigma}{\bar{x}} \times 100\%,$$

where $V_\sigma$ is the coefficient of variation, $\sigma$ is the standard deviation, and $\bar{x}$ is the mean.

2.2. Spatial Autocorrelation. From the dimension of space, we analyze and test the spatial correlation from the global spatial autocorrelation analysis. Whether the data have a spatial dependence must be examined when using spatial econometric methods. If a spatial dependence is absent, then the standard econometric method is adopted. Otherwise, the spatial econometric method is used.

Autocorrelation of spatial sequences is rather complicated, and spatial autocorrelation refers to the fact that regions with similar locations have similar variable values. If high-value aggregation and low-value aggregation are present, it indicates the presence of “positive space autocorrelation,” whereas if high values and low values are distributed completely randomly, it indicates the absence of any spatial autocorrelation.

Given the complexity of spatial autocorrelation, a series of methods for measuring spatial autocorrelation have been proposed in the past, such as Morans I, Garyy’s C, and Getis, the most common of which is the Morans I [8]. The global Morans I is often used to analyze the correlation index of the whole region, and the local Morans I is used to analyze the correlation index of each regional unit in the region. The global Morans I is calculated as follows:

$$I = \frac{W_{11}(Y_1 - \bar{Y})(Y_1 - \bar{Y}) + \cdots + W_{in}(Y_n - \bar{Y})(Y_n - \bar{Y}) + \cdots + W_{nn}(Y_n - \bar{Y})(Y_n - \bar{Y})}{s^2(W_{11} + \cdots + W_{in} + \cdots + W_{nn})},$$

in $s^2 = (Y_1 - \bar{Y})^2 + \cdots + (Y_n - \bar{Y})^2/n$, where

$$y = X\beta + \epsilon,$$

$$\epsilon = \lambda w_y + \mu,$$

is the sample variance and $W_{ij}$ is the spatial weight matrix $(i, j)$ (for measuring the distance between the measurement regions $i$ and $j$), which is generally constructed based on the
distance and adjacency relationship of geographic features. In this study, the binary matrix \( W_{ij} \) is constructed based on the adjacency relationship. The calculation result of Moran’s \( I \) is between -1 and 1. When it is greater than 0, it indicates the presence of a positive spatial correlation. When it is less than 0, it indicates the presence of a negative spatial correlation. When it is close to 0, it indicates that the adjacent regions are independent of each other, and thus no spatial autocorrelation among them is present.

The global Moran’s \( I \) only examines the spatial agglomeration of the entire region and cannot reveal which regions are high-value agglomeration regions or low-value agglomeration regions in the entire region. To elucidate the spatial agglomeration near a certain region, it is necessary to further examine local spatial autocorrelation, which can be performed through the local Moran’s \( I \), with the following calculation formula:

\[
I = \frac{(Y_i - \bar{Y})(W_{i1}(Y_1 - \bar{Y})^2 + \cdots + W_{in}(Y_n - \bar{Y})^2)}{s^2},
\]

in \( s^2 = (Y_1 - \bar{Y})^2 + \cdots + (Y_n - \bar{Y})^2/n \), where \( Y_j \) is the observation value of Region \( j \) and \( W \) is the spatial weight matrix. Like the global Moran’s \( I \), in the case of the local Moran’s \( I \), if Moran’s \( I > 0 \), it indicates the adjacent regions have similar feature values, assuming the “high value-high” or “low value-low” value distributions; if Moran’s \( I < 0 \), it indicates the adjacent regions have non-similar feature values, assuming the “high value-low” or the “high value-low” distributions, which indicates high spatial heterogeneity among the regions.

2.3. Spatial Econometric Model. The use of traditional econometric models on explanatory variables that have passed the test for spatial autocorrelation will lead to unreasonable estimation results; in this case, it is necessary to add a spatial weight matrix to modify the model and construct a spatial econometric model for analysis. Spatial econometric models are mainly categorized into two types: one is the spatial lag model (SLM), which is mainly used to study the behaviour of neighbouring regions and the spatial impacts on the behaviour of other regions in the whole system, such as diffusion or spillover. The SLM is thus suitable for estimating the intensity of spatial interactions. It has the following calculation formula:

\[
y = \rho W_{y}X\beta + \epsilon,
\]

where \( y \) is the explained variable, \( X \) is exogenous explanatory variable matrix of \( n \times K \), \( \rho \) is the spatial regression coefficient, \( W \) is the space weight matrix of order \( n \times n \), \( W_{y} \) is the explained variable of spatial lag, and \( \epsilon \) is the vector of the random error term. Another is the spatial error model (SEM), in which the relationship between the regions is represented by the random interference term, with the following calculation formula:

\[
y = X\beta + \epsilon, \quad \epsilon = \lambda W_{y} + \mu,
\]

where \( \epsilon \) is the vector of the random error term, \( \lambda \) is the spatial error coefficient of the vector of the explained variable of \( n \times 1 \) cross section, and \( \mu \) is the random error vector of normal distribution.

3. Spatial Econometric Analysis of the Regional Tourism Economy

To mine the spatial heterogeneity and spatial dependence information regarding the regional tourism economy of Anhui province, we used the per capita tourism income as the raw data, on which descriptive analysis was performed. The global spatial correlation was analyzed using the global Moran’s \( I \), and the local spatial correlation patterns of the regional tourism economy were further determined using the local Moran’s \( I \) and scatter plots.

3.1. Descriptive Analysis. First, we performed a descriptive analysis of the regional tourism economy of Anhui province using the mean, standard deviation, and coefficient of variation. The results are reported in Table 1 and Figure 1. Overall, the gap between the per capita tourism revenues of regions in Anhui province has gradually narrowed. The People’s Government of Anhui province issued the “Implementation Opinions on Further Accelerating the Development of Tourism Industry,” which adopts rational layout plans and promotion of coordinated development of the regional tourism to vigorously develop the tourism economy. It has effectively orchestrated regional tourism development and enabled the regions with relatively underdeveloped tourism to develop rapidly, thereby gradually narrowing the regional development gap. However, the coefficient of variation was still greater than 1, indicating that the dispersion in per capita tourism revenues in various regions is still high and that a development imbalance persists among the regions.

3.2. Spatial Autocorrelation Analysis. According to the calculation formula of the spatial autocorrelation, the per capita tourism income of each region was used in the calculation, and the results are shown in Table 2. It indicates that the values of Moran’s \( I \) of per capita tourism income of different regions of Anhui province were all greater than 0. Their normal statistics passed the significance test at the significance level of 0.1, indicating that, in terms of the spatial distribution, the per capita tourism revenues of different regions in Anhui province exhibited a significant positive spatial autocorrelation, that is, the regions with high per capita tourism income were adjacent to each other, as were those with low per capita tourism income (the presence
of spatial agglomeration among the regions). Table 2 indicates that the values of Moran’s I of different regions in the province from 2011 to 2016 were all greater than 0.2 and have exhibited a rising trend, indicating that, during the period, the spatial dependence of tourism development in various regions has gradually increased, and the trend that regions with high per capita tourism income levels and those with low per capita tourism income levels cluster in their respective regions has also intensified. In 2017, the values of Moran’s I slightly decreased, indicating that the agglomeration of per capita tourism avenues of different regions in Anhui province in 2017 was slightly weakened. At the same time, the coefficient of variation of per capita tourism revenues of different regions in 2017 was significantly greater than that in 2016, indicating that the difference in per capita tourism income in different regions increased.

We selected two time sections of 2015 and 2017 and generated scatter plots of the local Moran’s I values of per capita tourism revenues of different regions of Anhui province; the study shows that most regions clustered in the third quadrant, indicating that the low-value agglomeration type dominates and thus the existence of a positive spatial correlation. The “core regions” of high value clustered in the first quadrant included Huangshan and Chizhou Cities, and it is worth noting that Wuhu City caught up and entered the first quadrant, making the south Anhui tourism cluster centred in Huangshan and Chizhou slightly move towards central Anhui (Wuhu), which is largely related to the commercial development, cultural exchange, and the driving and developing of the Fangte Happy World in Wuhu. Xucheng, Tongling, Anqing, and Maanshan Cities were in the third quadrant (high value-low value), whereas Hefei, the provincial capital, has been always in the fourth quadrant (low value-high value), mainly because, among the regions, Hefei is a comprehensive development area, with inadequate dominant tourism resources, and its tourism revenues are derived mainly from businesses, exhibitions, and conferences.

The above results show that the regional tourism levels in Anhui province were characterized by spatial heterogeneity, profound local spatial clustering characteristics, and local spatial autocorrelation.

3.3. Spatial Econometric Model. Due to the presence of significant spatial dependence of regional per capita tourism revenues in Anhui province, the spatial distribution of the tourism economy in various regions has changed from 2011 to 2017. Thus, traditional econometric regression models are not appropriate for examining the spatial correlations between regions and not able to comprehensively reveal the effects of the factors, making it impossible to obtain complete and rational conclusions. To further analyze factors that have a significant impact on the per capita tourism revenue, we adopted a spatial econometric regression model to analyze the causes of spatial differences to better interpret the spatial correlations and impacts.

Tourism development is affected by many factors; theoretically, it is affected by macroeconomic factors, resource factors, population factors, policy systems, and other factors. To further analyze Anhui’s regional tourism development, we chose the indicator of per capita tourism income as the explained variable of the model to directly reflect the operation status and effectiveness of the regional tourism economy and the development healthiness of the tourism market; we used the per capita GDP (X 1), the number of star-rated hotels (X 2), the fixed assets investment (X 3), and the number of employees in the tertiary industry (X 4) as explanatory variables to perform the spatial regression analysis.
Based on spatial econometric-related theories, we used the Stata software package to generate the spatial lag model and the spatial error model. The lambda value of the autoregressive coefficient of the error term passed the significance test, whereas rho, the spatial autoregressive coefficient of the spatial lag model, did not, indicating that the spatial error model is able to fully extract and express spatial correlation information and thus superior to the spatial lag model. Therefore, we adopted the spatial error model to examine the influencing factors of the regional per capita tourism income, and the results are reported in Table 3.

Among the influencing factors of per capita tourism income, the per capita GDP, the total number of star-rated hotels, the fixed assets investment, and the number of employees in the tertiary industry all passed the significance test at the significance level of 1%, indicating that they have played a significant role in tourism development.

### 4. Results

The results are as follows:

1. The regression coefficient of per capita GDP was positive, indicating that the impact of the local economic development level on tourism development is positive as a promoting effect, and the regression coefficient is small, indicating that local economic development level is one of the main factors affecting tourism development.

2. The regression coefficient of the number of star-rated hotels was positive and the highest, indicating that the number of star-rated hotels in each region has a positive effect on tourism development and is the most important factor affecting tourism development.

3. Both the fixed assets investment and the number of employees in the tertiary industry passed the significance test, indicating that both are factors that affect tourism development, but their regression coefficients were negative, which is contradictory to the previous notion that increasing fixed assets investment and the number of employees of the tertiary industry can promote tourism income. This result is mainly due to the limitation of the data used here, in which the amount of fixed assets investment in the development of tourism industry and the number of employees in the tourism industry were lacking, which affected the model estimation results. Once the abovementioned data become available, the model can be further improved.

### 5. Conclusions

Based on the above analyses, the global spatial autocorrelation (dependency) of regional per capita tourism income levels in Anhui province was strong, exhibiting spatial agglomerations. The regions with "high value-high value" agglomeration mainly included Huangshan, Chizhou, and Wuhu, and those with "low value-low value" agglomeration mainly included Chuzhou, Lu’an, Fuyang, Suzhou, Huainan, and Bengbu, whereas Hefei, the capital city, has always been in the region with "low value-high value" agglomeration. Based on the spatial characteristics of tourism development in Anhui province, we provide the following recommendations:

1. The key to developing the regional tourism industry [7] and increasing tourism revenue is to vigorously develop the local economy. After the local economy is developed, people’s disposable income will be increased, which will stimulate tourism demand and promote the development of the tourism industry. Anhui province’s tourism investment has been continuously growing, and the implementation of the "335" tourism construction action plan has achieved an accumulated investment of 804.5 billion yuan during the "Twelfth Five-Year Plan" period, which was 5.3 times of the total investment during the "Eleventh Five-Year Plan" period, accounting for 8.9% of the province’s total accumulated fixed asset investment. Further adjusting the proportion structure of fixed assets investment in the tourism industry, formulating reasonable investment plans, and forming a more sophisticated investment system can continuously improve the investment efficiency. At the same time, the relevant departments should use funds dedicated to the key points of tourism industry development to maximize capital effectiveness, such that the waste of funds can be minimized.

2. The differences among different regions of Anhui province were profound. Due to the limitation of geographical location, the agglomeration effect of southern Anhui has had very little driving and spatial spillover effects on central and northern Anhui. Southern Anhui should emphasize its agglomeration effect and spatial spillover effect to drive the development of its neighbouring regions [8], such that tourism can be developed more rapidly in regions with inexplicit resources.

We should encourage each region to engage in regional cooperation and promote tourism for all, actively implement various tourism marketing strategies to increase the publicity of tourist attractions [9], strengthen the focus of tourism promotion on regions neighbouring known scenic spots (e.g., Huangshan and Jiuhuashan), and increase the investment in the tourism infrastructure and supporting facilities, especially the transportation construction, thereby making the access to tourist attractions unimpeded and consequently achieving coordinated development of tourism.

---

**Table 3: Results of the spatial error model.**

<table>
<thead>
<tr>
<th>Model parameter</th>
<th>Coefficient</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>X 1</td>
<td>0.2003439</td>
<td>0.000</td>
</tr>
<tr>
<td>X 2</td>
<td>859.4962</td>
<td>0.000</td>
</tr>
<tr>
<td>X 3</td>
<td>−5.690162</td>
<td>0.000</td>
</tr>
<tr>
<td>X 4</td>
<td>−59.6762</td>
<td>0.000</td>
</tr>
<tr>
<td>C</td>
<td>3,370.822</td>
<td>0.013</td>
</tr>
</tbody>
</table>
(3) Because the spatial location factor is an important factor affecting the development of the regional tourism industry in Anhui province, the regional tourism industry development in the province is affected not only by the economic development level of the region and the number of star-rated hotels but also by the development level of tourism in the neighbouring regions, and the levels of development of different regions are interdependent to some extent. Improvement of the tourism level or improvement of economic development in a certain region has positive effects on the tourism development in neighbouring regions. Therefore, to develop regional tourism in Anhui province and to increase regional tourism revenue, we should simultaneously grasp the development statuses of the region and neighbouring regions and make full use of all favourable factors. We should also strengthen the exchange of experience in the development of the tourism industry in neighbouring regions, especially information sharing regarding infrastructure and tourism talents. In areas with rich explicit tourism resources, we should utilize the resources and geographical advantages to drive the tourism development in adjacent areas, thereby forming a large-scale tourism industry cluster.

(4) To address the spatial dependence problem of regional tourism development in Anhui province, tourism management departments should coordinate the development of regional tourism in the province, and the government should actively develop effective and adaptive tourism development plans and construct “smart Anhui tourism,” using cutting-edge information and big data methods to achieve integrated tourism across the province. Moreover, the tourism industry in different regions should be integrated to conduct unified guidance and management regarding tourism product development, infrastructure construction, and resource and environment protection to continuously deepen the coordinated development of regional tourism while ensuring sustainable tourism development in Anhui province.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by 2020 Anhui Quality Engineering Project: Accounting and Statistics Teaching Team (2020jxtd031), 2019 Anhui Business and Technology College Teaching Innovation Team Project (2019xjixtd01), 2018 Anhui Natural Science Research Project (KJ2018A084), 2017 Anhui Natural Science Research Project (KJ2017B004), 2017 Anhui Vocational and Adult Education Association Project (azjxh17109), and 2016 Anhui Provincial Tourism Bureau Key Research Project (2016FACJ2985-8).

References


