International Journal for Housing Science and Its Applications
Publish August 6, 2025. Volume 46, Issue 3 Pages 5574-5584

https://doi.org/10.70517/ijhsa463481

Characterization of spatial and temporal evolution of rural tourism flows based on spatial autoregressive models

Tingting Jin[1,*]
[1] Boda College of Jilin Normal University, Siping, Jilin, 136000, China
Corresponding authors: (e-mail: 13634464978@139.com).

Abstract With the improvement of transportation network and the upgrading of consumer demand, rural tourism presents significant spatio-temporal evolution characteristics. In-depth analysis of rural tourism flow characteristics and revealing its spatio-temporal evolution law are of great significance to optimize the regional tourism layout, promote the implementation of rural revitalization strategy, and promote the high-quality development of rural tourism. Based on the rural tourism development data of 31 provinces in mainland China from 2015 to 2023, this study explores the spatio-temporal evolution characteristics of rural tourism flows by using spatial autocorrelation analysis, spatial autoregressive model and spatial trend surface analysis. The results of the study show that: the heat of rural tourism shows an "S"-shaped trajectory of change, with the most rapid growth between 2018 and 2020; the distribution within the year is characterized by "four peaks and four valleys", with the passenger flow in spring and autumn being larger than that in summer, and the lowest in winter, and the passenger rail transportation volume has a significant positive correlation with the GDP of rural tourism (the coefficient is 2.758, $p<0.001$). In terms of spatial distribution, the source of rural tourism was "high in the south and low in the north, high in the east and low in the west", and the Yangtze River Delta, Pearl River Delta, Henan, Anhui and other provinces were the main sources of tourists, and the spatial lag coefficients were 0.232 and 0.252, respectively ($p<0.001$), which indicated that there was an obvious spatial clustering effect. The robustness test further verifies the robustness of the regression results. It is found that the synergistic development of railroad network and aviation network positively affects rural tourism flows, and with the improvement of transportation conditions, the spatial agglomeration effect of rural tourism source market gradually weakens, and the scope of the source market keeps expanding.

Index Terms Rural tourism, spatio-temporal evolution, spatial autoregressive model, source distribution, heat analysis, tourism flow

I. Introduction

With the in-depth development of tourism, rural tourism to meet the different needs and preferences of tourists came into being. Rural tourism is the mutual adaptation, interaction and result of urban demand and rural, especially peri-urban agricultural integrated resources, opening up a new way to increase income in traditional agriculture [1]. In addition, the development of rural tourism not only helps to develop the rural economy, improve the rural environment, but also plays an important role in the construction of rural civilization and rural governance, has become an important way to implement rural revitalization, solve the "three rural" problem, rural tourism in rural development, urban-rural coordination of the positive role of rural tourism has been formed a consensus [2]-[5]. Rural tourism has a great attraction to tourism consumers, and the rural tourism industry has been growing in a "blowout" manner, and the development of rural tourism has presented a broad market prospect.

The spatial and temporal selection process of rural tourists determines the flow and direction of rural tourism flows, which directly affects the formation of hot and cold patterns of rural tourism flows [6]. Understanding the temporal and spatial distribution characteristics of rural tourism flows is an important theoretical guidance for strengthening the construction of rural tourism destinations and enhancing the image of rural tourism destinations. The prospect of rural tourism is promising, especially with the proposal of rural revitalization strategy, rural tourism has been a booming trend, a large number of rural tourism scenic spots of different scales, grades and types have emerged, rural tourism scenic spots have been transformed from point distribution to face distribution, the spatial layout is no longer only dependent on the outskirts of the city or the surrounding of the famous scenic spots but is expanding to more areas suitable for development, and the development of rural tourism has gradually shifted from attaching importance to a single The development of rural tourism has gradually transitioned from emphasizing the development of a single scenic spot to the development stage of destination system integration [7]-[10]. How to fully tap the advantages of rural characteristics and resources, optimize the spatial layout of rural tourism, promote regional cooperation, and guide the efficient flow of rural tourism flow is a realistic problem that local governments and rural tourism business operators and managers need to think about urgently. On the one hand, for the local government, due to the natural substrate, economic development, resource endowment, etc. of each region is different, the role and status in the rural tourism development network is very different, how to comprehensively and objectively examine the regional economic development, resource endowment and other realistic conditions, in the basis of their own development of reasonable positioning, the scientific development of rural tourism development planning, the rural tourism destination building and regional cooperation has an important role [11]-[15]. On the other hand, for rural tourism enterprises, how to develop rural tourism products under the premise of understanding the needs of rural tourists, avoiding homogenized competition, and creating rural tourism brands with local characteristics and reflecting local culture is of great practical significance for promoting the quality and efficiency of rural tourism and enhancing the attractiveness and competitiveness of rural tourism [16]-[19]. Therefore, in the context of the booming development of rural tourism, it has become an important scientific proposition to study the law of spatio-temporal flow of rural tourists, to understand the influencing factors of the spatio-temporal choice of rural tourists, and to master the mechanism of the formation of the hot and cold pattern of rural tourism flow.

As an important part of China's tourism industry system, rural tourism plays a key role in promoting rural economic development, optimizing industrial structure, and expanding farmers' income-generating channels. With the improvement of people's living standards and the transformation of leisure concepts, the demand for rural tourism market continues to grow, showing the development trend of diversification, personalization and high quality. As the spatial manifestation of tourism activities, the spatial and temporal evolution of rural tourism flows not only reflects the development and utilization of regional tourism resources, but also reflects the laws of inter-regional economic ties and population flows. An in-depth study of the spatial and temporal evolution of rural tourism flows is of great significance for understanding the structural changes in the tourism market and optimizing the allocation of tourism resources. In recent years, scholars at home and abroad have carried out extensive research on rural tourism flows, mainly focusing on flow scale measurement, spatial structure characteristics, influencing factors and other aspects. However, most of the existing studies are limited to static analysis and pay insufficient attention to the dynamic evolution characteristics of rural tourism flows; at the same time, it is difficult for traditional research methods to comprehensively reveal the complex spatial and temporal characteristics of rural tourism flows, especially in terms of spatial dependence and heterogeneity, there are obvious deficiencies. With the acceleration of high-speed rail network construction and the improvement of civil aviation system, the influence of transportation conditions on rural tourism flows has become more and more prominent, but the existing studies have not explored the interaction between transportation network and rural tourism flows in sufficient depth. In addition, due to the limitation of data acquisition, there is a relative lack of systematic research on the spatial and temporal evolution characteristics of rural tourism flows under long time series. Therefore, it is of great theoretical and practical value to construct a scientific analytical framework to comprehensively reveal the spatio-temporal evolution of rural tourism flows and its internal mechanism, in order to promote the high-quality development of rural tourism.

Based on the above background, this study firstly constructs a spatial autoregressive model of rural tourism flows, selects endogenous variables such as railroad passenger traffic, airport throughput and rural tourism GDP, and combines them with exogenous variables such as household population and industrial structure to comprehensively analyze the influencing factors of rural tourism flows. Secondly, spatial autocorrelation analysis is applied to identify the spatial agglomeration characteristics of rural tourism flows and their evolution trends. Once again, based on the spatial variable coefficient model and the spatial trend surface analysis method, the spatial directionality of rural tourism source distribution and its changing law are revealed. Finally, the reliability and stability of the research results are verified by combining a variety of robustness tests. Through the above research, this paper constructs a comprehensive analytical framework of the spatial and temporal evolution characteristics of rural tourism flow, reveals the spatial and temporal evolution law of rural tourism flow and its internal mechanism, and provides scientific basis for optimizing the spatial layout of regional tourism, implementing differentiated tourism development strategies, and promoting the high-quality development of rural tourism.

II. Sources of data for the study

II. A.2.1 Selection of influencing factors

In this paper, railroad passenger traffic, airport throughput and rural tourism GDP are taken as endogenous variables, in which railroad passenger traffic and airport throughput are used to measure the developed degree of railroad network and civil aviation system in each study area, and gross domestic product (GDP) is chosen to measure the level of economic development of the city. In order to strengthen the robustness and credibility of the results of the regression fitting analysis of the spatio-temporal evolution of rural tourism flows, a series of control variables are selected to weaken the exogenous influence of the explanatory variables, including the household population, the proportion of the value-added of the tertiary industry, and so on, by systematically combing through the existing literature, and by borrowing from the relevant theories of tourism economics and tourism geography. The endogenous and exogenous variables are shown in Table 1.

Table 1: Variables

-	Specific variables
Endogenous variables	Railway passenger volume (y_1)
	Airport Throughput (y_3)
	GDP(y_2)
Exogenous variables	Household population (pop)
	Population growth rate ($popr$)
	Science and Technology Expenditure (se)
	Education expenditure (ee)
	The proportion of added value of the secondary industry (sr)
	The proportion of added value of tertiary industry (tr)
	The proportion of employees in the secondary industry (psr)
	Proportion of employees in the tertiary industry (ptr)

II. B.2.2 Data sources and description

Data for each variable are from the dataset of rural tourism development in 31 provinces in mainland China, 2015-2023. The indicator data of exogenous data come from China Statistical Yearbook, China Statistical Yearbook of Regional Economy, China Environmental Statistical Yearbook, China Transportation Statistical Yearbook, and Zhongjing.com; the indicator data of foreign direct investment come from Wind database, statistical yearbooks and statistical bulletins of provinces, and are converted to RMB as the settlement unit according to the average exchange rate of the year, and the data of this indicator of TAR for the years 2016-2019 are missing, and are made up by the moving average method; data of A-class tourist attractions of each province come from the data set of rural tourism development of 31 provinces in mainland China from 2015 to 2023. The data of this indicator is missing in 2019, and the moving average method is used to make up for it; the data of A-class tourist attractions in each province are from China Tourism Statistical Yearbook (2016-2023) and China Cultural Relics and Tourism Statistical Yearbook (2022-2023). Among them, the 2018 National Tourism Statistical Yearbook only has the total number of tourist attractions in each province, and there is no statistics by grade.

III. Research Methodology

In this paper, spatial and temporal evolution characteristics of rural tourism flow research and analysis, the use of spatial autocorrelation analysis, spatial constant coefficient model, spatial variable coefficient model and spatial trend surface analysis method and other related methods and theories.

III. A. 3.1 Spatial autocorrelation analysis

According to the theory of spatial measurement and spatial economy, we first need to carry out the spatial autocorrelation test, and if there is autocorrelation, then establish the spatial econometric model to carry out the test and analysis of spatial measurement [20]. Spatial autocorrelation refers to the correlation between the observations of a variable due to the spatial proximity of the observation points, that is to say, spatial autocorrelation analysis is to test whether the variables are correlated in the study area, and to statistically measure the degree of correlation, which is generally expressed by the Moran's I index.The Moran's I index of the Moran's I index is calculated by the formula:

$$Moran's\ I = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}(Y_i - \bar{Y})(Y_j - \bar{Y})}{S^2 \sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}} \tag{1}$$

In the formula $S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)$, $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$, denotes i region's observation, n is the total number of regions, and W_{ij} is the spatial weight matrix, this paper adopts the distance criterion to define the mutual adjacency of spatial objects.Moran's I lies between $[-1,1]$, when it is greater than 0, it indicates that there exists a spatial positive correlation in the inbound tourism, and the closer the value is to +1, the stronger the positive correlation is; when it is smaller than 0, it shows that negative spatial correlation exists, the closer the value is to -1, the stronger the negative autocorrelation is; equal to 0 indicates that there is no correlation between inbound tourism in each region.

III. B. 3.2 Spatial constant coefficient model

Spatial constant coefficient models include spatial lag model (SLM), spatial error model (SEM). The spatial lag model mainly discusses whether there is a diffusion of each variable in a certain region, that is, whether there is a spillover effect, and its formula is as follows:

$$Y = \rho W_y Y + X\beta + \varepsilon \tag{2}$$

where Y is the dependent variable, X is the observation matrix of the independent variable, and ρ is the number of spatial autoregressive relationships, which takes values between -1 and 1, indicating the degree and direction of influence between neighboring regions. W_y is the spatial lagged dependent variable and ε is the vector of independent random error terms. β is the independent variable parameter, reflecting the effect of the independent variable X on the dependent variable Y.

The spatial dependence in the spatial error model, on the other hand, exists in the error perturbation term, which measures the extent to which the error shocks in neighboring regions about the dependent variable affect the observations in the region, and is given by the following equation [21]:

$$Y = X\beta + \varepsilon \tag{3}$$

$$\varepsilon = \lambda W\varepsilon + \mu \tag{4}$$

where Y, X, β and (4) have the same meaning, ε is the vector of random error terms; λ is the spatial error autocorrelation coefficient of the vector of cross-sectional dependent variables of $n\times 1$, which measures the direction and degree of influence of neighboring regions Y on the region Y, and takes the value of $[-1:1]$; μ is the vector of normally distributed random errors; $W\varepsilon$ is the spatial lag perturbation term.

III. C. 3.3 Spatial coefficient of variation model

Geographically weighted regression (GWR) recognizes that there is complexity, difference, and variability in data within space, and argues that regression models need to take such characteristics into account in order to make the results of the study more realistic [22]. In general, global regression models are:

$$Y_i = a_0 + \sum_k a_k x_{ik} + \varepsilon_i \tag{5}$$

However, the global regression model only shows the global average effect, while the local parameters are not fully represented. Taking the local parameters into account can more clearly reflect the different effects of the independent variables on the dependent variable in different regions. Therefore, the local area model is:

$$Y_i = a_0(U_i,V_i) + \sum_k a_k(U_i,V_i)x_{ik} + \varepsilon_i \tag{6}$$

where, (U_i,V_i) is the spatial coordinate of the i th sample point; $a_k(U_i,V_i)$ is the value of the continuous function $a_k(U,V)$ at the i th point. If $a_k(U,V)$ remains constant in space, the GWR model becomes global.

III. D. 3.4 Spatial trend surface analysis

III. D. 1) 3.4.1 Fundamentals

Spatial Trend Surface Analysis, a mathematical method that uses mathematical surfaces to model the trend of a target parameter in space. According to the relevant principle of the least squares method, the multivariate nonlinear function is transformed into a multidimensional linear fitting equation.

Let the actual observed value of a set of target parameters that has been obtained be $W_i(x_i, y_i)(i=1,2,\cdots,n)$, and the calculated value of the spatial trend surface fitting equation be $\hat{W}_i(x_i, y_i)$, then there are:

$$W_i(x_i, y_i) = \hat{W}_i(x_i, y_i) + \varepsilon_i \tag{7}$$

where, x_i, y_i are the geographic coordinates of the measurement points; ε_i is the residual, i.e., the difference between the computed value of the fitting equation and the measured value.

Theoretically, any function can be approximated by a polynomial in an appropriate range; by reasonably adjusting the number of polynomials, the fitted equations (polynomials) ultimately obtained can be made to converge infinitely to the actual function. The key to spatial trend surface analysis lies in determining the values of coefficients in the fitted equations (polynomials) with different numbers of times, so that the residual sum of squares $\sum \varepsilon_i^2$ tends to be minimized according to multiple sets of actual observations (x_i, y_i, W_i), as follows [23].

Let $x_1 = x$, $x_2 = y$, $x_3 = x^2$, $x_4 = xy$, $x_5 = y^2$, and ...... then:

$$\hat{W} = a_0 + a_1 x_1 + a_2 x_2 + \cdots + a_p x_p \tag{8}$$

$$\sum \varepsilon_i^2 = \sum_{i=1}^{n} \left[W_i - \hat{W}_i \right]^2 = \sum_{i=1}^{n} \left[W_i - \left(a_0 + a_1 x_{1i} + a_2 x_{2i} + \cdots + a_p x_{pi} \right) \right]^2 \tag{9}$$

Find the partial derivatives of $\sum \varepsilon_i^2$ with respect to $a_0, a_1, \cdots, a_p$, respectively, and make them equal to 0 to obtain a system of multiple linear equations:

$$\begin{cases} na_0 + a_1 \sum_{i=1}^{n} x_{1i} + \cdots + a_p \sum_{i=1}^{n} x_{pi} = \sum_{i=1}^{n} W_i \\ a_0 \sum_{i=1}^{n} x_{1i} + a_1 \sum_{i=1}^{n} x_{1i} x_{1i} + \cdots + a_p \sum_{i=1}^{n} x_{pi} x_{1i} = \sum_{i=1}^{n} x_{1i} W_i \\ \cdots\cdots \\ a_0 \sum_{i=1}^{n} x_{pi} + a_1 \sum_{i=1}^{n} x_{1i} x_{pi} + \cdots + a_p \sum_{i=1}^{n} x_{pi} x_{pi} = \sum_{i=1}^{n} x_{pi} W_i \end{cases} \tag{10}$$

The coefficients of polynomials such as $a_0, a_1, \cdots, a_p$ can be found by using the system of linear equations in Eq. (10) containing $p+1$ equations and $p+1$ unknowns to finalize the expression of the equation for fitting the spatial trend surface.

III. D. 2) 3.4.2 Fitness test for fitting equations to spatial trend surfaces

In spatial trend surface analysis, for the identified analysis samples (i.e., one or several sets of measured data of target parameters), the selection of the number of times of the fitting equations directly affects the final degree of fitting and convergence to the actual spatial variation characteristics of the target parameters. Optimizing the most reasonable number of polynomials through the following fitness test process is an indispensable part of spatial trend surface analysis, and it is also a key issue related to whether the obtained fitting equations and their spatial expression results can be applied in actual research and practice.

1) The R^2 test for the fitness of the trend surface

The fitness coefficient R^2 is an important indicator to measure and reflect the goodness of fit of the trend surface to the measured data (actual surface). $R^2 = \frac{SS_R}{SS_T} = 1 - \frac{SS_D}{SS_T}$, the larger R^2 is, the better the fit of the trend surface is.

where $SS_D = \sum_{i=1}^{n} \left(W_i - \hat{W}_i \right)^2$ is the residual sum of squares; $SS_R = \sum_{i=1}^{n} \left(\hat{W}_i - \bar{W} \right)^2$ is the regression sum of squares; $\bar{W}$ is the mean of W_i:

$$SS_T = \sum_{i=1}^{n}\left(W_i - \hat{W}\right)^2 + \sum_{i=1}^{n}\left(\hat{W}_i - \overline{W}\right)^2 = SS_D + SS_R \tag{11}$$

2) F -test for the fitness of the trend surface

Is used to analyze the significance of the fit of the trend surface to the measured data (actual surface):

$$F = \frac{SS_R / p}{SS_D / n - p - 1} \tag{12}$$

At the significance level α, if the calculated F value is greater than the critical value $F_{(\alpha)}$, the trend surface equation fit is considered to be significant; otherwise, it is not significant.

3) Successive tests of trend surface fitness

Find the difference between the regression sum of squares of the higher polynomial fit equation and the regression sum of squares of the lower polynomial fit equation, divided by the degrees of freedom of the regression sum of squares, to obtain the mean squared deviation of the regression due to the increasing number of polynomials. Dividing this mean squared deviation by the remaining mean squared deviation of the higher order polynomial yields the F value of the fitness comparison test for the above two order trend surface equations. If the resulting F value is significant, the higher order polynomial makes a new contribution to the improvement in fit; if the F value is not significant, the higher order polynomial makes no new contribution to the improvement in fit.

IV. Characterization of the spatial and temporal evolution of rural tourism flows

The spatio-temporal evolution characteristics of rural tourism flows are an important characterization of the quality of rural tourism development and regional differences. In this chapter, we will identify the hot and cold patterns of rural tourism flows, explore the spatial directional characteristics of the distribution of sources in rural tourism, and study the spatio-temporal evolution characteristics of the hot and cold patterns of rural tourism flows.

IV. A. 4.1 Analysis of the temporal evolution of the heat of rural tourism

(1) Annual change characteristics

The heat value of each scenic spot in the countryside is summarized according to the year and month, and the annual map of the heat value of rural tourism is drawn, as shown in Figure 1. Rural tourism heat value is "S" type change trajectory, rural tourism heat year by year growth trend is very obvious. 2014-2017 period, the development of rural tourism is slower, rural tourism heat value is lower, although it has been growing, but the rate of growth is slow; Rural tourism heat value in 2018 increased rapidly in 2018-2020 period has maintained a high growth rate, but it is not easy for rural tourism heat value to grow. 2020 has maintained a high growth trend; 2021 began to slow down, rural tourism heat into a stable growth stage.

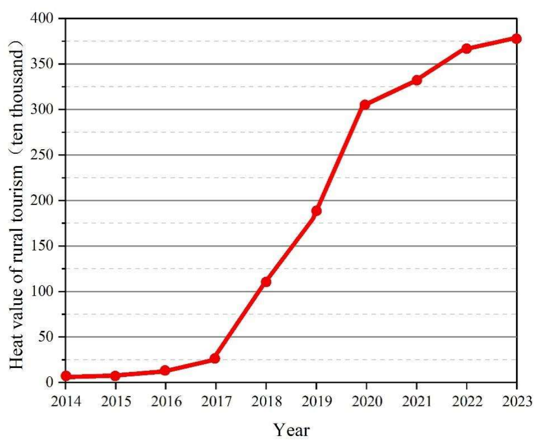


Figure 1: The annual change of rural tourism heat

(2) Seasonal and intra-week change rule

The annual change of rural tourism heat value is shown in Figure 2. From the point of view of the yearly change of rural tourism heat, the seasonal difference of rural tourism heat is significant, overall "four peaks and four valleys" type of change characteristics, holidays and weekends effect is prominent. April, May, August and October is the peak season of rural tourism, and respectively in the Ching Ming Festival, May Day, summer vacation, the 11th

National Day and other holidays, the peak of the flow of passengers. In April and May, the climate is suitable, and it is the time of spring blossom, plus the increase of leisure time during the holidays such as Ching Ming Festival and May Day Labor Day, which plays a great role in promoting rural tourism.

However, it can be seen from the figure that the intensity of rural tourism flows is greater in April than in May, which is different from the time distribution characteristics of general tourism flows. August is the summer vacation, the majority of teachers and students travel by the time constraints are smaller, the overall passenger flow is larger, parent-child travel at this time can not be underestimated as the main source of passengers. At the end of August to the beginning of September ushered in the school season, the heat of rural tourism appeared "diving" type decline, after the temperature gradually cooler, and coincided with the Mid-Autumn Festival holiday, the heat of rural tourism began to pick up, to the National Day period once again appeared to the peak of passenger flow. October autumn high and refreshing, is the golden period of autumn picking, coupled with the National Day Golden Week holiday time is longer, the temperature is suitable, the flow of passengers showed a "blowout" growth, the heat of rural tourism is unusually high. January, June, November and December are the off-season for rural tourism. Except for the New Year's Day and Chinese New Year holidays, which contribute to the growth of rural tourism heat, all other periods of time have smaller passenger flow and lower rural tourism heat. It can be seen that the heat of rural tourism has more obvious seasonal fluctuation characteristics, with larger flows in spring and fall, intermediate flows in summer and smaller flows in winter, and the overall distribution of spring flows > fall flows > summer flows > winter flows.

From the point of view of the weekly distribution of rural tourism heat, the highest frequency and the most stable "sawtooth" cycle in the figure corresponds to the weekly double holidays, forming a periodic cycle of the peak forest structure. It shows that the double holiday is an important time for rural tourism, and the heat of rural tourism shows an obvious weekend effect. The reason for this is that, on the one hand, with the year-on-year increase in the number of private cars and the gradual improvement of transportation conditions, coupled with the convenience provided by the big data era for the acquisition of tourism information and real-time query, the heat of rural tourism, which is mainly based on short- and medium-distance trips, has risen sharply. On the other hand, the spatial and temporal distance and consumption level of rural tourism scenic spots are more moderate, and the huge weekend recreation market is frequently directed to the suburbs (counties) under the effect of the law of close travel, which gradually becomes the ideal choice for citizens' weekend self-guided tours.

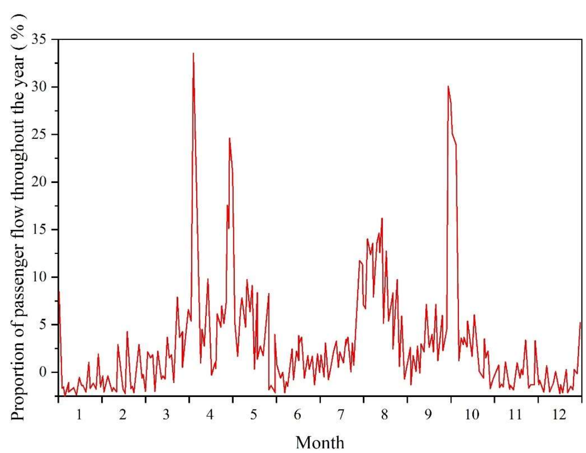
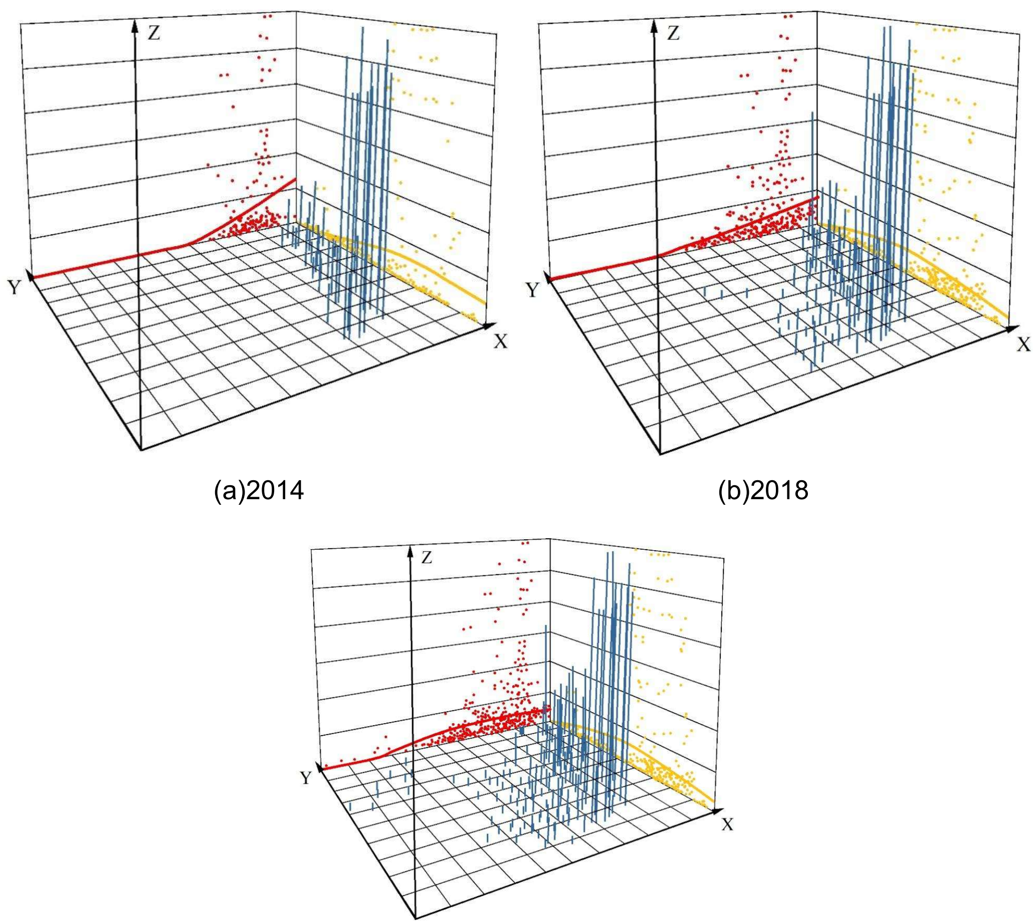


Figure 2: Seasonal and weekly variation of rural tourism heat

IV. B. 4.2 Analysis of the spatial evolution of the distribution of rural tourism sources

In this section, the trend surface analysis tool of ArcGIS software platform will be used to analyze the spatial pointing characteristics of rural tourism guest sources. Based on the geographic location of each source of rural tourism and the intensity of passenger flow for the overall trend surface fitting, to get the trend surface analysis map of the spatial pointing of rural tourism sources in different time periods, specifically as shown in Figure 3. Figures (a) to (c) correspond to 2014, 2018, and 2023, respectively, in which the X and Y axes identify the geographic location of the source, the Z axis indicates the intensity of the passenger flow in each source, and the red line represents the east-west direction, while the yellow line indicates the north-south direction. The trend surface fitting results show that during the research period, the spatial trend lines of rural tourism source areas exhibited a clear spatial orientation characteristic, generally maintaining a spatial layout pattern of "high in the south and low in the north, high in the east and low in the west". The north-south trend line showed an upward trend in the southern region, while the east-west trend line exhibited a significant rise in the easternmost areas, indicating that regions such as the Yangtze River Delta, the Pearl River Delta, and provinces like Henan and Anhui are the main sources for rural tourism flows. At the same time, it is noted that there are differences in the trend line changes of rural tourism flow intensity in different directions, with the trend surface in the north-south direction tending to be flat, while the trend surface in the east-west direction is steeper, which implies that the divergent characteristics of rural tourism flow intensity in the east-west direction are more prominent. However, in the past ten years, this spatial differentiation feature has slowed down to different degrees in both the north-south and east-west directions, indicating that the spatial agglomeration effect of the rural tourism source market has gradually weakened. With the acceleration of tourism information flow and the speeding up of high-speed railways, subways and other transportation, the spatial and temporal distance between sources and destinations is shortened, which enhances the attractiveness of rural tourism sources, and the scope of the source market is constantly expanding, with a weakening of the agglomeration effect, which is basically consistent with the characteristics of the agglomeration of the sources as discussed in the previous section.



(a)2014 (b)2018

(c)2023

Figure 3: The overall trend of passenger flow intensity in source areas

V. Regression analysis of spatio-temporal evolution characteristics of rural tourism flows

In this chapter, we will combine the principles of spatial autocorrelation analysis, spatial constant coefficient model, and spatial variable coefficient model to further regression analysis of the spatio-temporal evolution characteristics of rural tourism flows.

V. A. 5.1 Analysis of calibration results

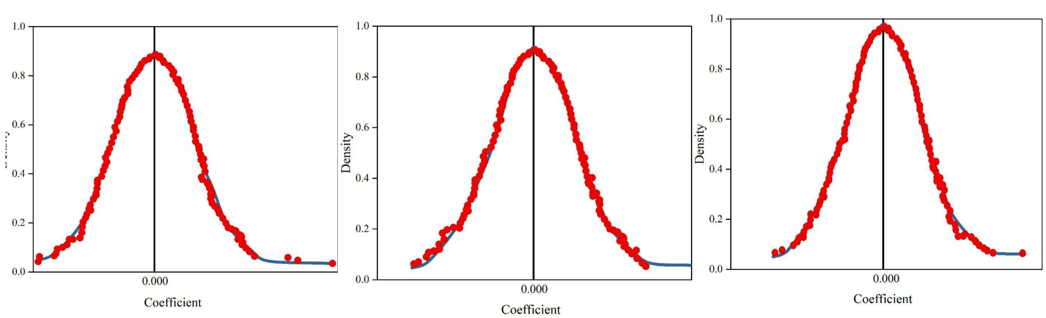
Due to the difference in the magnitude of variables, the unit of rural tourism GDP is taken as hundreds of billions of yuan. The regression results of the spatial and temporal evolution characteristics of rural tourism flows are shown in Table 2. The R² of each regression equation is greater than 0.400, indicating that the regression analysis results have good explanatory ability. The p-value of each endogenous explanatory variable is less than 0.010, which preliminarily verifies the validity of regression causality. The spatial lag coefficients of railroad passenger traffic and airport throughput are positive and significant, indicating that with the continuous development and improvement of air-railway network, the cooperative relationship between railroad network and aviation network is strengthened, presenting the phenomenon of high-value agglomeration.

Table 2: Regression result

Variable	Coefficient	Z	P	95.000% confidence interval	
Endogenous variable: y_2					
y_1	2.758	7.36	0.000	2.038	3.484
Constant term	6.414	0.388	0.724	-2.667	3.946
Exogenous variables: pop , $popr$, psr , ptr					
R²	0.788				
RMSE	0.486				
Endogenous variable: y_3					
y_2	0.246	4.15	0.000	0.136	0.375
Constant term	-1.815	-1.53	0.131	-4.192	0.546
Exogenous variables: se , sr , tr					
R²	0.851				
RMSE	0.135				
Endogenous variable : y_1					
y_3	1.047	3.53	0.000	0.481	1.679
Constant term	0.316	3.79	0.000	0.155	0.472
Exogenous variable: ee					
R²	0.583				
RMSE	0.321				
$\rho 1$	0.232	66.35	0.000	0.216	0.23
$\rho 2$	0.252	18.57	0.000	0.235	0.278

V. B. 5.2 Robustness test

In order to verify the robustness of the regression results, two methods of robustness testing were used. First, a placebo test is conducted, which is based on the principle of eliminating “false significance” caused by the interference of other policies or random factors. Based on the Monte Carlo method, the coefficients of y_1, y_2 and y_3 are randomly disturbed 500 times. The changes of the coefficients after disruption are specifically shown in Fig. 4. Figures (a) to (c) correspond to the cases of disrupting y_1, y_2 and y_3, respectively. It can be seen that the coefficients of the disrupted variables obey a normal distribution with a mean of 0.000 on the whole, verifying the robustness of the results.



(a)Derange y_1 (b)Derange y_2 (c)Derange y_3

Figure 4: Placebo test

The corresponding elasticity analysis regression results were subsequently obtained using elasticity analysis, as shown in Table 3. The results show that the regression coefficients remain significant after taking the logarithm of the endogenous variables, verifying the robustness of the regression results.

Table 3: Regression result

Variable	Coefficient	Z	P	95.000% confidence interval	
Endogenous variable: y_2					
$\ln y_1$	1.283	11.32	0.000	1.086	1.527
Constant term	2.546	8.367	0.75	-13.78	18.963
Exogenous variables: pop , $popr$, psr , ptr					
R^2	0.942				
RMSE	0.241				
Endogenous variable: y_3					
$\ln y_2$	0.824	8.12	0.000	0.622	1.025
Constant term	-7.175	-2.59	0.01	-12.612	-1.736
Exogenous variables: se , sr , tr					
R^2	0.627				
RMSE	0.562				
Endogenous variable : y_1					
$\ln y_3$	1.062	6.35	0.000	0.715	1.365
Constant term	1.241	0.355	0.001	0.536	1.98
Exogenous variable: ee					
R^2	0.445				
RMSE	0.611				
$\rho 1$	0.242	90.54	0	0.237	0.248
$\rho 2$	0.278	13.8	0	0.239	0.318

VI. Conclusion

This study has systematically analyzed the spatial and temporal evolution characteristics of rural tourism flows in China from 2015 to 2023 by using spatial econometric methods, and has drawn the following main conclusions:

First, rural tourism heat shows significant stage and cycle characteristics in time. The annual change shows an "S" trajectory, especially after 2018, with rapid growth; the seasonal distribution shows a "four peaks and four valleys" pattern, with obvious holiday effects, and four tourism peaks formed in Qingming Festival, May Day, summer vacation and National Day; the intra-week distribution shows an obvious "jagged" cycle, with prominent weekend effects.

Secondly, the spatial distribution of rural tourism sources shows the trend of "high in the south and low in the north, high in the east and low in the west", with the Yangtze River Delta, the Pearl River Delta and provinces such as Henan and Anhui as important sources. With the passage of time, this spatial difference has weakened, indicating that the spatial agglomeration effect of the rural tourism source market is gradually weakening and the market scope is expanding.

Third, the regression analysis shows that there is a significant positive correlation between railroad passenger traffic and rural tourism GDP, with an elasticity coefficient of 1.283 ($p<0.001$), indicating that the improvement of transportation conditions has a significant role in promoting the development of rural tourism. The spatial lag coefficients are both positive and significant (0.232 and 0.252, respectively, $p<0.001$), confirming that there is a significant spatial spillover effect of rural tourism mobility.

Fourth, the results of placebo test and elasticity analysis verified the robustness of the research findings. The regression coefficients of the endogenous variables remained significant after taking logarithms, and the coefficients of the disrupted variables obeyed a normal distribution with a mean value of 0.000, indicating that the research results were not disturbed by random factors.

The above findings provide theoretical support for the formulation of regional rural tourism development strategies, and suggest that future rural tourism development should focus on improving the transportation network, optimizing the seasonal layout, and expanding the scope of the source market to promote the high-quality development of rural tourism.

References

[1] He, Y., Wang, J., Gao, X., Wang, Y., & Choi, B. R. (2021). Rural tourism: Does it matter for sustainable farmers' income?. Sustainability, 13(18), 10440.

[2] Zhang, L. (2022). THE INFLUENCE OF COORDINATED DEVELOPMENT OF RURAL TOURISM MANAGEMENT AND ECOLOGICAL CIVILIZATION CONSTRUCTION ON VILLAGERS'EMOTIONAL DISORDER. Psychiatria Danubina, 34(suppl 2), 579-579.

[3] Valderrama, E. L., & Polanco, J. A. (2024). Understanding how collaborative governance mediates rural tourism and sustainable territory development: a systematic literature review. Tourism Recreation Research, 49(4), 888-904.

[4] Yongguang, Z. O. U., Haoying, Z. H. O. N. G., & Yifu, C. U. I. (2024). Performance Evaluation and Enhancement Path of Rural Tourism Promoting Rural Revitalization. Economic geography, 44(8), 201-210.

[5] Zhangxin, Y. I. N., Liuyan, D. A. I., Chang, G. A. N., & VODA, M. (2024). COUPLING COORDINATION RELATIONSHIP BETWEEN TOURISM DEVELOPMENT AND URBAN-RURAL INTEGRATION. Geographia Technica, 19(2).

[6] Li, Y., Gong, G., Zhang, F., Gao, L., Xiao, Y., Yang, X., & Yu, P. (2022). Network structure features and influencing factors of tourism flow in rural areas: Evidence from China. Sustainability, 14(15), 9623.

[7] Qi, J., Lu, Y., Han, F., Ma, X., & Yang, Z. (2022). Spatial distribution characteristics of the rural tourism villages in the Qinghai-Tibetan Plateau and its influencing factors. International Journal of Environmental Research and Public Health, 19(15), 9330.

[8] Du, J., Zhao, B., & Feng, Y. (2024). Spatial distribution and influencing factors of rural tourism: A case study of Henan Province. Heliyon, 10(7).

[9] Yi, Y., Siow, M. L., Ibrahim, R., & Abdul Aziz, F. (2024). Understanding the Role of Tourist-Oriented Villages in Promoting Rural Tourism in China: Integrating Rural Landscapes and Tourist Services. Journal of Quality Assurance in Hospitality & Tourism, 1-41.

[10] Wu, X., & Chen, C. (2022). Spatial distribution and accessibility of high level scenic spots in Inner Mongolia. Sustainability, 14(12), 7329.

[11] Qin, X., Li, X., Chen, W., Tan, H., Luo, L., & Xu, X. (2022). Tourists' digital footprint: The spatial patterns and development models of rural tourism flows network in Guilin, China. Asia Pacific Journal of Tourism Research, 27(12), 1336-1354.

[12] Shen, S., Wang, H., Quan, Q., & Xu, J. (2019). Rurality and rural tourism development in China. Tourism Management Perspectives, 30, 98-106.

[13] Kataya, A. (2021). The impact of rural tourism on the development of regional communities. Journal of Eastern Europe Research in Business and Economics, 2021(10).

[14] SUN, J. X., ZHANG, L. Y., & LUO, Y. L. (2023). Development of rural tourism resources in China under the goal of common prosperity: Current situation, problems and the development path. Journal of natural resources, 38(2), 318-334.

[15] Liu, C., Dou, X., Li, J., & Cai, L. A. (2020). Analyzing government role in rural tourism development: An empirical investigation from China. Journal of Rural Studies, 79, 177-188.

[16] Li, T., Liu, J., Zhu, H., & Zhang, S. (2018). Business characteristics and efficiency of rural tourism enterprises: an empirical study from China. Asia Pacific Journal of Tourism Research, 23(6), 549-559.

[17] Li, Z. (2025). Rural Tourism and the Development of Characteristic Towns in China and Its Influence on Rural Areas. International Journal of Agricultural and Environmental Information Systems (IJAEIS), 16(1), 1-23.

[18] RĂCĂŞAN, B. S., & EGRESI, I. (2020). TOURISM, BRANDING AND TERRITORIAL IDENTITY IN THE RURAL SPACE. LOCAL AUTHORITIES'PERSPECTIVE. Territ. Identity Dev, 4, 18-38.

[19] Su, Y., Wang, Y., & Li, R. (2025). Mechanisms to Overcome the Homogenization of Rural Tourism Products and Improve the Competitiveness of Rural Tourist Destinations: A Case Study from China. Systems, 13(4), 287.

[20] Julie Le Gallo,Fernando A. López & Coro Chasco. (2020). Testing for spatial group-wise heteroskedasticity in spatial autocorrelation regression models: Lagrange multiplier scan tests. The Annals of Regional Science,64(2),287-312.

[21] Cai Jiaxin,Hu Weiwei,Yang Yuhui,Yan Hong & Chen Fangyao. (2024). Outlier detection in spatial error models using modified thresholding-based iterative procedure for outlier detection approach. BMC Medical Research Methodology,24(1),89-89.

[22] Hyunwoo Lee & Young Woong Park. (2025). Integrated subset selection and bandwidth estimation algorithm for geographically weighted regression. Pattern Recognition,165,111589-111589.

[23] Xianming Zhang,Faqiang Zhan,Chunyan Hao,Ying Duan Lei & Frank Wania. (2025). Prioritizing Chemical Features in Non-targeted Analysis through Spatial Trend Analysis: Application to the Identification of Organic Chemicals Subject to Mountain Cold-Trapping. Environmental science & technology.