ATLANTIS
PRESS

The Measurement and Promotion Strategy of Provincial Tourism Efficiency in China Based on Three-stage DEA and Malmquist Index

Yan Zhang[1,*] Yeqin Fu[2]

[1] *College of Science and Technology University of Sanya, Sanya, Hainan 572022, China*
[2] *School of Tourism, Hainan Normal University, Haikou, Hainan 571158, China*
[*]*Corresponding author. Email: 28237148@qq.com*

ABSTRACT

This paper uses three-stage DEA (Data Envelopment Analysis) and Malmquist index model to measure and analyze the tourism economic efficiency of 31 provincial regions in China from 2014 to 2016. The results show that: after removing the interference of environment and random factors, the average comprehensive efficiency of tourism in China has increased by 24.5%, which indicates that the tourism efficiency of most provincial regions is underestimated; by analyzing the dynamic trend, it is found that the average increase of technical efficiency is 2.7%, the average increase of technical change is 12.3%, and the average increase of MI index is 15.4%, which indicates that the overall tourism efficiency of China is on the rise The main factor that restricts the efficiency of tourism industry in China is scale efficiency. Finally, the provincial regions are divided into several types, and suggestions are put forward to improve the efficiency of regional tourism from the aspects of improving the efficiency of tourism management and expanding the scale of tourism investment.

Keywords: Tourism efficiency, Environment variable, Three-stage DEA model, Malmquist index.

1. INTRODUCTION

Tourism, a general term for enterprises and institutions that provide services for tourists or various activities related to tourists, refers to a modern service industry with interactivity and influence based on natural geographical and cultural resources. China's rapid economic and social development elevated tourism as a pillar industry: in 2019, the number of Chinese tourists exceeded 6 billion and the total tourism revenue exceeded 6 trillion yuan. Provinces that generally focus on tourism make use of their natural, historical and cultural resources to promote commerce, transportation and star-rated hotels, forming an extensive growth dominated by scale expansion. In view of the rapid economic growth and favorable

tourism policies, the provinces have adopted various measures to enhance their tourism competitiveness. However, although tourism, which relies on capital and labor input, can effectively drive regional economic development in the short term, it will lead to overexploitation of natural resources and excess of human resources in the long term. In addition, despite its expanding scale, tourism still suffers from unfavorable development level and uneven growth among regions. The 13th Five-Year Plan advocates the transformation, upgrading, quality and efficiency improvement of tourism industry. Therefore, comprehensive and effective measurement of tourism efficiency and exploration of strategies to improve tourism efficiency contribute much to facilitate the transformation of its development mode.

The main methods adopted at present to evaluate tourism efficiency include SFA (Stochastic Frontier Approach) and DEA models. The former is employed in many papers including *Regional Evaluation of China's Tourism Industry Efficiency and Its Deconstruction Analysis: An Enqrirical Study Based on SFA* written by Hu Yaguang [1], and *Estimation of*

China's Tourism Efficiency and Analysis of Regional Differences Based on SFA Model by Meng Tao [2]. SFA, a parameter method for evaluating efficiency, features pre-set regression equation. It determines whether the regression equation is accurate by testing the significance of unknown parameters, and determines the optimal front surface by testing parameters. However, such method can only have one output index and is not applicable to efficiency evaluation with multiple outputs. Non-parametric DEA theory, another important efficiency evaluation method, is adopted in such as: Gu Jiang's *Appraisal and Model Foundation of the Efficiency in Tourism Production in China* [3], and Liang Mingzhu's *An Evaluation and Analysis of Tourism Efficiency in Different Cities and Regions of Guangdong Province* [4]. DEA is characterized by decision making units that can evaluate multi-input and multi-output indicators without estimating and testing parameters. However, the weight is too flexible and the influence of environment and random factors on efficiency is not taken into account.

In summary, tourism efficiency should include both internal management efficiency and external environmental efficiency. The traditional one-stage DEA cannot avoid the error influence of random factors, and the DEA efficiency value may be underestimated or overestimated, which is different from the traditional one. In contrast, the three-stage DEA method [5] proposed by Fried et al. can effectively eliminate the influence of external environmental factors on the input, so as to make the estimated tourism efficiency more real. The main issues to be solved in this paper include: 1) Constructing a provincial-level regional tourism efficiency evaluation index system on the context of the "13th Five-Year Plan" to fully reflect the input and output of China's tourism industry and the main influencing factors; 2) Eliminating the impact of management inefficiency and random errors on the evaluation of the tourism efficiency of various provincial administrative regions using three-stage DEA method; 3) Revealing the present situation of tourism development in different parts of China and analyzing the influence of external environmental factors on tourism efficiency by determining the tourism efficiency of 31 provincial regions in China from 2014 to 2016, in a bid to provide insight for improving the quality and efficiency of tourism industry and its transformation and upgrading.

2. VARIABLE SELECTION

2.1 Input-output Variables

Labor, capital, and land are considered as the main driving force for economic growth in economics. Compared with the primary and secondary industries, tourism that belongs to the tertiary industry is less dependent on land, which affects the development of tourism. Since land is not the main factor affecting tourism, this paper chooses input factors from the perspective of labor and capital. Based on the input-output theory, an evaluation index system for tourism efficiency in various provincial regions in the context of the "13th Five-Year Plan" is constructed from the angle of economy and human utilization ("Table 1").

The input indicators mainly cover the number of tourism employees and the original price of tourism fixed assets. The former is adopted to indicate the input of labor factors. Travel agencies, tourist attractions and star-rated hotels serve as the three major forces driving tourism development. Material capital is calculated using the perpetual inventory method. However, since the data of investment and depreciation of tourism enterprises in various regions are not available, the sum of the original price of fixed assets of travel agencies, tourist attractions and star-rated hotels is taken as the original price of fixed assets of tourism. In terms of output, the total income from tourism is selected as economic output.

2.2 Environment Variables

Environmental variables refer to the factors that are not directly affected by tourism but are essential for tourism, covering not only microeconomic environment and infrastructure, but also social environment and cultural environment.

As for microeconomic environment, the more developed the regional economy, the higher the degree of opening to the outside world, the greater the degree of tourism investment and development, and the greater the income from tourism. In this paper, residents' income is taken as an environmental variable to measure the external economic environment required by tourism consumption and residents' demand for tourism.

With regard to infrastructure, the transportation network density is taken as the variable. The denser the transportation network, the higher the travel accessibility rate. This indirectly stimulates the development of tourist attractions and star-rated

hotels, which prompts residents' desire for travel, and drives tourism consumption.

As regards social environment, the cultural quality of residents is adopted to reflect the scale of tourism research, planning, tourism colleges and the quality of population in a region.

In terms of cultural environment, the places of interest in each region influence local tourism. The

number of scenic spots, national forest parks, famous historical and cultural cities, key cultural relic protection units and A-level scenic spots can be pooled. Scoring criteria should also be developed and weighted separately, so that the weighted average method can be used to calculate the total value to understand residents' travel preferences.

Table 1. Evaluation index system for tourism efficiency in China

First-level index	Second-level index	Third-level index	Reference source
Investment index	Number of tourism employees	Number of employees in star-rated hotels	Literature[7]
		Number of employees in travel agencies	Literature [8]
		Number of employees in A-level scenic spots	the research
	Original price of tourism fixed assets	Original price of fixed assets of star-rated hotels	the research
		Original price of fixed assets of travel agency	the research
		Original price of fixed assets in tourist attractions	the research
Output index	Total tourism revenue	Total operating income of tourism enterprises	Literature [8]
Environment variable	Residents' income	Per capita disposable income share	Literature[10]
	Traffic Conditions	Traffic network density	the research
	Residents' cultural quality	Proportion of population with higher education	the research
	Travel preferences	The proportion of tourists received in the country	Literature [11]

3. RESEARCH DESIGN AND DATA SOURCES

3.1 The Basic Steps of the Three-stage DEA Model

3.1.1 The First Stage: DEA-CCR Model

DEA model is an efficiency measurement method proposed by Charnes et al. in the 1970s. It is a linear programming model obtained using relevant theories of operational research under the assumption that the return to scale remains unchanged. In 1984, Banker et al. proposed an efficiency measurement model with variable returns to scale, the BCC model, which decomposes the overall efficiency in the CCR model into pure technical efficiency and scale efficiency. The DEA model is subdivided into input and output. Input-oriented refers to minimizing the required input variables and maximizing the output by controlling the weight coefficient of input variables under the given conditions of output. Since tourism is

characterized by unchangeable output and adjustable input, this paper selects an input-oriented DEA-CCR model to measure the economic efficiency of tourism in 31 provincial regions of China.

3.1.2 The Second Stage: SFA Model

The first stage fails to avoid the influence of random errors and external states, and the calculated efficiency value was inaccurate. In the second stage, the stochastic frontier analysis model was utilized to establish a regression equation with multiple external state variables as explanatory variables and input slack variables as explained variables. The input variables in the first stage were adjusted to make different decision making units have the same external environment and luck, so as to eliminate the influence of external environment and random error. A SFA regression equation is established for each input slack variable:

$$S_{ni} = f^n(Z_i; \beta^n) + V_{ni} + U_{ni} , \quad i = 1, 2, \cdots, I ; \quad n = 1, 2, \cdots, N$$

Where, S_{ni} is the relaxation variable of the i decision unit on the n input, N is the number of input index, I is the number of decision-making units; Z_i is the external environment vector; β^n is

unknown parameter. $V_{ni} + U_{ni}$ is a comprehensive error term, in which V_{ni} reflects the statistical noise, and $U_{ni} \geq 0$ is management inefficiency. The maximum likelihood method is used to estimate the

unknown parameters. The initial input is adjusted

according to the following formula:

$$X_{ni}^* = X_{ni} + [\max\{Z_i\hat{\beta}\} - z_i\hat{\beta}] + [\max\{\hat{V}_{ni}\} - \hat{V}_{ni}]$$

Among them, X_{ni} is the initial investment in the

first stage, and X_{ni}^* is the adjusted investment in the second stage. The first brackets indicate that all provincial regions are adjusted to have the same external state, and the second brackets indicate that

statistical noise (random error) is removed from all provincial regions. The condition of random error V_{ni} is estimated as:

$$\hat{E}[V_{ni}|V_{ni} + U_{ni}] = S_{ni} - Z_i\hat{\beta}^n - \hat{E}[U_{ni}|V_{ni} + U_{ni}]$$

To estimate V_{ni}, you must first calculate U_{ni}. This article uses the formula given by Luo Dengyue [6] to estimate U_{ni}:

$$E[U_{ni}|V_{ni} + U_{ni}] = \frac{\sigma\lambda}{1 + \lambda^2}\left[\frac{\varphi(\frac{\varepsilon_i\lambda}{\sigma})}{\Phi(\frac{\varepsilon_i\lambda}{\sigma})} + \frac{\varepsilon_i\lambda}{\sigma}\right]$$

$\sigma^2 = \sigma_{un}^2 + \sigma_{vn}^2$, where σ_{un}^2 is the variance

of U_{ni}, σ_{vn}^2 the variance of V_{ni}, ε_i is the

$$\lambda = \frac{\sigma_u}{\sigma_v}$$

comprehensive error term, σ_v, Φ, and φ are distribution function and the density function of the standard normal distribution.

The third stage: the traditional DEA model The

input variable X_{ni}^* of the second stage is employed

to replace the initial input variable X_{ni}, and the

CCR model is adopted to calculate the tourism efficiency of each provincial region. The efficiency at this time eliminates the influence of external conditions and statistical noise, and more accurately reflects the efficiency of the region's tourism industry.

3.2 Malmquist Index

Malmquist index model is a quantitative index to analyze productivity changes in two different periods [12]. This index is used to investigate the change of total factor productivity in each provincial administrative region. The TFP index from T period to T +1 period is defined as:

$$M(x^t, y^t, x^{t+1}, y^{t+1}) = \left[\frac{D^t(x^{t+1}, y^{t+1})}{D^t(x^t, y^t)} \times \frac{D^{t+1}(x^{t+1}, y^{t+1})}{D^{t+1}(x^t, y^t)}\right]^{\frac{1}{2}}$$

M is greater than 1, indicating that TFP is increasing from period t to period t+1. Malmquist index can be decomposed into technical efficiency change index (TEC) and technical progress change index (TC). A TEC greater than 1 indicates an improvement in relative technical efficiency and that a certain region is closer to the production frontier, and TC greater than 1 indicates progress in production technology. TEC can be further

decomposed into scale efficiency change (SEC) and pure technical efficiency change (PTEC), specifically expressed as: M=TEC×TC=PTEC×SEC ×TC. If PTEC and SEC are greater than 1, it means that it has a positive effect on the M index.

The three-stage DEA model can only analyze cross-sectional data and cannot measure changes in operating efficiency. Through the adjusted input variables in the second stage, the Malmquist index is

used to calculate the actual value of TFP after excluding environmental factors.

3.3 Data Source and Pre-processing

31 provincial regions in China are taken as objects. The number of employees in tourism-related industries and the original price of tourism fixed assets are taken as input vectors, and the total income of tourism is taken as output vectors. The input and output data are all from China Tourism Statistical Yearbook and provincial statistical yearbook from 2014 to 2016. The number of tourism employees refers to the total number of employees in star-rated hotels, travel agencies and A-level scenic spots, and the original price of fixed assets in tourism refers to the total of the original price of fixed assets in star-rated hotels, travel agencies and scenic spots. In order to facilitate the application of the SFA model in the second stage, the unit of tourism practitioners is ten thousand, and that of the original price of tourism fixed assets one hundred million yuan. The environmental variables data can be found in *China Statistical Yearbook*. The per capita disposable

income of each province and region is divided by the sum of the per capita income of the whole country to represent the proportion of the per capita disposable income. The coefficient of variation method is adopted to weigh the highway mileage of each grade and then divided by the area of each province to get the traffic network density of each provincial region. The proportion of the population with higher education was represented by the proportion of people with college education in the total population, and the number of tourists received by each provincial region was divided by the number of tourists in the whole country to measure the travel preference.

In order to verify whether the input-output data of the first stage are relevant, Stata14.0 software is employed to calculate the Pearson correlation coefficient of the average of the three indexes ("Table 2"). "Table 2" reveals that the Pearson correlation coefficients of the input-output mean values of each provincial region are positive and passed the test, meeting the model requirements.

Table 2. Pearson correlation coefficient of the mean value of input-output index

	Number of tourism employees	Original price of tourism fixed assets	Tourism income
Number of tourism employees	1	0.7677**	0.8345**
Original price of tourism fixed assets	0.7677**	1	0.7485**
Tourism income	0.8345 **	0.7485 **	1

a   Note: ** indicates a significant correlation at the 0.01 level (two-sided).

4.   EMPIRICAL ANALYSIS

4.1 DEA Analysis of the First Stage

MaxDEA Ultra is employed to calculate the tourism economic efficiency of 31 provincial administrative regions in mainland China from 2014 to 2016 ("Table 3"). Factoring out the influence of external environment and random factors, the overall efficiency of tourism in 2014 was effective in only two provinces of China, with an overall average efficiency of 0.456 and the lowest of 0.144 (Ningxia).

More than half of the provincial regions were in the state of diminishing returns to scale. In 2015 and 2016, the overall efficiency of provincial regions improved slightly, but not higher than 0.5, and no sea change occurs in the number of provinces with comprehensive technical efficiency and return to scale. DEA from the first stage fails to truly reflect the real situation of tourism economic efficiency in different parts of China. The influencing factors in the second stage are required to be separated given external and random factors.

Table 3. Tourism efficiency of the first stage

Year	Area	Overall efficiency	Technical efficiency	Scale efficiency	Return to scale	Slack Variable 1	Slack Variable 2
2014	Anhui	0.488	0.539	0.905	Drs	4.47	85.1
2014	Beijing	0.329	0.412	0.8	Drs	10.87	593.9
2014	Fujian	0.333	0.349	0.955	Drs	6.77	151.5
2014	Gansu	0.237	0.425	0.558	Irs	3.16	64.9
2014	Guangdong	0.352	0.831	0.424	Drs	17.92	382.9
2014	Guangxi	0.5	0.504	0.991	Drs	3.27	67.2

Year	Area	Overall efficiency	Technical efficiency	Scale efficiency	Return to scale	Slack Variable 1	Slack Variable 2
2014	Guizhou	1	1	1	-	0	0
2014	Hainan	0.177	0.283	0.626	Irs	2.93	110.8
2014	Hebei	0.3	0.304	0.987	Drs	7.41	202.9
2014	Henan	0.796	1	0.796	Drs	1.39	30.6
2014	Heilongjiang	0.316	0.519	0.61	Irs	3.16	50.9
2014	Hubei	0.481	0.576	0.834	Drs	5.03	125
2014	Hunan	0.333	0.361	0.923	Drs	14.76	134.6
2014	Jilin	0.794	0.832	0.955	Irs	0.59	16.2
2014	Jiangsu	0.51	1	0.51	Drs	9.62	251.9
2014	Jiangxi	0.578	0.598	0.966	Irs	3.59	42.6
2014	Liaoning	0.529	1	0.529	Drs	10.37	103.9
2014	Inner Mongolia	0.575	0.605	0.951	Irs	1.65	53.7
2014	Ningxia	0.144	0.707	0.203	Irs	1.06	38.9
2014	Qinghai	0.191	1	0.191	Irs	1.09	21.5
2014	Shandong	0.332	0.673	0.493	Drs	15.18	291.7
2014	Shanxi	0.648	0.693	0.935	Drs	1.92	51
2014	Shaanxi	0.4	0.402	0.996	Drs	4.69	105.4
2014	Shanghai	0.462	0.53	0.873	Drs	4.79	264.4
2014	Sichuan	0.59	0.838	0.704	Drs	4.22	111.9
2014	Tianjin	1	1	1	-	0	0
2014	Tibet	0.399	1	0.399	Irs	0.38	21
2014	Xinjiang	0.161	0.282	0.572	Irs	4.2	109.8
2014	Yunnan	0.328	0.339	0.966	Drs	6.78	166
2014	Zhejiang	0.446	0.772	0.577	Drs	9.75	329.4
2014	Chongqing	0.416	0.429	0.968	Irs	3.5	98.8
2014	Mean	0.456	0.639	0.748		5.31	131.56
2015	Anhui	0.541	0.587	0.922	Drs	3.89	84.7
2015	Beijing	0.29	0.368	0.788	Drs	11.33	618.7
2015	Fujian	0.362	0.363	0.997	Drs	6.28	147.4
2015	Gansu	0.263	0.449	0.587	Irs	3.25	63
2015	Guangdong	0.326	1	0.326	Drs	18.78	650.4
2015	Guangxi	0.556	0.571	0.973	Irs	4.52	53.4
2015	Guizhou	1	1	1	-	0	0
2015	Hainan	0.143	0.266	0.536	Irs	3.44	121.7
2015	Hebei	0.33	0.352	0.939	Drs	7.49	204.4
2015	Henan	0.843	1	0.843	Drs	1.03	26.2
2015	Heilongjiang	0.399	0.621	0.644	Irs	2.81	42.1
2015	Hubei	0.453	0.516	0.879	Drs	5.71	145.1
Year	Area	Overall efficiency	Technical efficiency	Scale efficiency	Return to scale	Slack Variable 1	Slack Variable 2
2015	Hunan	0.364	0.404	0.902	Drs	13.5	133.1
2015	Jilin	0.837	0.893	0.937	Irs	0.49	12.7
2015	Jiangsu	0.528	1	0.528	Drs	8.06	268.1
2015	Jiangxi	0.711	0.76	0.935	Drs	2.56	30.4
2015	Liaoning	0.364	0.377	0.965	Drs	7.86	144.6
2015	Inner Mongolia	0.475	0.503	0.945	Irs	2.53	80.6
2015	Ningxia	0.141	0.735	0.191	Irs	0.99	56.7
2015	Qinghai	0.191	0.82	0.233	Irs	1.17	28.7

Year	Area	Overall efficiency	Technical efficiency	Scale efficiency	Return to scale	Slack Variable 1	Slack Variable 2
2015	Shandong	0.328	0.718	0.457	Drs	16.38	312.8
2015	Shanxi	0.783	0.856	0.914	Drs	0.99	30.1
2015	Shaanxi	0.376	0.386	0.974	Irs	6.05	109.3
2015	Shanghai	0.38	0.425	0.894	Drs	5.53	295
2015	Sichuan	0.561	0.905	0.62	Drs	5.29	138.6
2015	Tianjin	1	1	1	-	0	0
2015	Tibet	0.335	1	0.335	Irs	0.56	19.4
2015	Xinjiang	0.235	0.337	0.696	Irs	3.76	88.4
2015	Yunnan	0.326	0.328	0.996	Drs	7.89	166.4
2015	Zhejiang	0.424	0.722	0.588	Drs	9.72	366.9
2015	Chongqing	0.388	0.414	0.938	Irs	3.73	108
2015	Mean	0.46	0.635	0.757		5.34	146.66
2016	Anhui	0.548	0.549	0.998	Drs	3.67	86.6
2016	Beijing	0.333	0.337	0.99	Drs	9.32	620.4
2016	Fujian	0.377	0.389	0.968	Irs	6.08	139
2016	Gansu	0.248	0.46	0.538	Irs	3.47	84.5
2016	Guangdong	0.409	1	0.409	Drs	14	597.5
2016	Guangxi	0.652	0.701	0.93	Irs	2.46	43
2016	Guizhou	1	1	1	-	0	0
2016	Hainan	0.143	0.269	0.532	Irs	3.74	183.5
2016	Hebei	0.4	0.412	0.97	Irs	6.77	134.1
2016	Henan	0.516	0.648	0.796	Drs	5.06	123.2
2016	Heilongjiang	0.394	0.713	0.553	Irs	3.01	47.4
2016	Hubei	0.485	0.486	0.996	Drs	4.85	123.7
2016	Hunan	0.391	0.401	0.975	Irs	13.34	140.6
2016	Jilin	0.914	0.98	0.933	Irs	0.25	7
2016	Jiangsu	0.522	1	0.522	Drs	8.7	276.7
2016	Jiangxi	0.9	0.902	0.997	Irs	0.98	10.7
2016	Liaoning	0.403	0.409	0.984	Irs	5.88	132.6
2016	Inner Mongolia	0.635	0.686	0.925	Irs	1.46	42.2
2016	Ningxia	0.152	0.852	0.178	Irs	1.09	44.2
2016	Qinghai	0.181	1	0.181	Irs	1.31	33
2016	Shandong	0.326	0.751	0.434	Drs	16.68	301.6
2016	Shanxi	0.798	0.804	0.993	Drs	1	31.9
2016	Shaanxi	0.393	0.428	0.92	Irs	5.52	116.7
2016	Shanghai	0.372	0.374	0.995	Drs	6.06	407
2016	Sichuan	0.633	1	0.633	Drs	4.15	149.6
2016	Tianjin	1	1	1	-	0	0
Year	Area	Overall efficiency	Technical efficiency	Scale efficiency	Return to scale	Slack Variable 1	Slack Variable 2
2016	Tibet	0.329	1	0.329	Irs	0.63	30.7
2016	Xinjiang	0.227	0.372	0.61	Irs	4.45	113.3
2016	Yunnan	0.409	0.41	0.997	Drs	6.4	159
2016	Zhejiang	0.493	0.811	0.608	Drs	7.73	297.2
2016	Chongqing	0.427	0.469	0.91	Irs	3.31	93.5
2016	Mean	0.484	0.665	0.768		4.88	147.42

a   Note: irs means increasing returns to scale, drs means diminishing returns to scale,-means constant returns to scale, slack variable 1 is the number of tourism employees, slack variable 2 is the original price of fixed assets in the tourism industry.

4.2 Regression Analysis of SFA in the Second Stage

The slack variables of the two input variables in the first stage were taken as dependent variables, and residents' income, traffic conditions, residents' culture and travel preference were taken as explanatory variables. Frontier4.1 software was adopted to measure the results ("Table 4").

Table 4. SFA regression results of the number of tourism practitioners and the original price of tourism fixed assets in various provincial regions in China

	Year	Slack variable: The number of tourism employees	T-ratio	Input Variables of Original Price of Tourism Fixed Assets	T-ratio
Constant term	2016	-0.54955097E+01***	-0.55815806E+01	-0.22049968E+02***	-0.39779724E+01
	2015	-0.59268919E+01***	-0.73391566E+01	-0.32876471E+02***	-0.25523795E+02
	2014	-0.65726635E+01***	-0.71214641E+01	-0.20186240E+02***	-0.43459784E+02
Residents' income	2016	0.25070788E+01***	0.55082658E+01	0.97743703E+01***	0.42270439E+01
	2015	0.29440118E+01***	0.70561528E+01	0.12895832E+02***	0.92540028E+01
	2014	0.30878257E+01***	0.70365921E+01	0.10327542E+02***	0.57161681E+02
Traffic network density	2016	-0.15604111E+01*	-0.20522817E+01	-0.29866245E+01***	-0.11107410E+01
	2015	-0.28134232E+01***	-0.84366886E+01	-0.10262967E+02***	-0.41346311E+01
	2014	-0.15114602E+01**	-0.19877365E+01	-0.39904696E+01***	-0.34156396E+02
Residents' cultural quality	2016	-0.11749948E+01**	-0.13142269E+01	-0.39230647E+01***	-0.80340481E+00
	2015	-0.89840470E+00*	-0.11532673E+01	-0.35135801E+01***	-0.38221008E+01
	2014	-0.37169051E+01***	-0.40258485E+01	-0.97240935E+01***	-0.21931190E+02
Travel preferences	2016	0.66355191E+00*	0.30287377E+01	0.20642889E+01***	0.17697124E+01
	2015	0.10209010E+01**	0.52973023E+01	0.45323581E+01***	0.63885181E+01
	2014	0.12249749E+01**	0.76299911E+01	0.26313490E+01***	0.33124790E+02
σ^2	2016	0.25908484E+02***	0.26432837E+02	0.53875223E+03***	0.22799008E+01
	2015	0.32887769E+02***	0.33167134E+02	0.27492032E+03***	0.27155112E+03
	2014	0.28691147E+02***	0.28774727E+02	0.13705602E+03***	0.13692621E+03
γ	2016	0.99999999E+00***	0.13828095E+08	0.95851723E+00***	0.39396061E+02
	2015	0.99999999E+00***	0.27843942E+07	0.99999999E+00***	0.14535301E+03
	2014	0.99999999E+00***	0.64098062E+05	0.99999999E+00***	0.17601899E+08

a   Note: *, **, *** mean significant at the 10%, 5%, and 1% significance level respectively.

"Table 4" shows that all the four environmental variables have passed the significance test. The γ value is above 0.9, which proves the vital role of external environment and random error on the tourism economic efficiency in all provinces and regions in China. Regression coefficients less than zero indicate that the reduction of environmental variables will lead to the increase of slack variables, resulting in the waste of corresponding input variables. Regression coefficients greater than zero indicate that the increase of environmental variables will lead to the increase of input redundancy, thus undermining efficiency. The specific results are as follows. The regression coefficient of residents' income level to the number of tourism employees is positive, indicating that the higher the residents' income, the lower the local residents' willingness to engage in the service industry will be, prompting more migrants to engage in the low-end service industry, which cuts down the number of locals engaged in the tourism service industry. The regression coefficient of the traffic network density to the input slack variable is negative, indicating that better infrastructure and travel accessibility rate facilitate tourists, and increase tourism income, which is in accordance with common sense. The regression coefficient of residents' educational level is negative. This shows that the higher the quality of local residents, the deeper the cultural deposits, the stronger their travel desire, the more comprehensive and profound their understanding of cultural tourism, the higher the quality and ability of the corresponding tourism practitioners, the lower the redundancy of the number of practitioners, and the full use of the value of the fixed assets of the tourism industry.

4.3 DEA Analysis After the Adjustment of Input Variables in the Third Stage

The adjustment of the initial input variable in the second stage is re-applied to the DEA model to show the value of tourism economic efficiency and return to scale of all provincial administrative regions in China ("Table 5").

Table 5. The efficiency of tourism economy in the third stage (after adjusting the input variables)

Year	Area	TE	PTE	SE	RTS
2014	Anhui	0.672	0.786	0.855	Irs
2014	Beijing	0.522	0.6	0.869	Irs

Year	Area	TE	PTE	SE	RTS
2014	Fujian	0.527	0.779	0.677	Irs
2014	Gansu	0.22	0.728	0.303	Irs
2014	Guangdong	0.726	0.831	0.873	Drs
2014	Guangxi	0.65	0.929	0.7	Irs
2014	Guizhou	1	1	1	-
2014	Hainan	0.122	0.797	0.153	Irs
2014	Hebei	0.425	0.678	0.627	Drs
2014	Henan	1	1	1	-
2014	Heilongjiang	0.337	0.797	0.423	Irs
2014	Hubei	0.712	0.836	0.852	Irs
2014	Hunan	0.464	0.467	0.992	Irs
2014	Jilin	0.598	0.944	0.634	Irs
2014	Jiangsu	1	1	1	-
2014	Jiangxi	0.731	0.831	0.88	Irs
2014	Liaoning	0.787	1	0.787	Drs
2014	Inner Mongolia	0.498	1	0.498	Irs
2014	Ningxia	0.059	0.841	0.07	Irs
2014	Qinghai	0.107	1	0.107	Irs
2014	Shandong	0.66	0.696	0.948	Drs
2014	Shanxi	0.67	0.924	0.725	Irs
2014	Shaanxi	0.508	0.772	0.658	Irs
2014	Shanghai	0.739	1	0.739	Irs
2014	Sichuan	0.982	1	0.982	Irs
2014	Tianjin	0.873	0.959	0.91	Irs
2014	Tibet	0.1	1	0.1	Irs
2014	Xinjiang	0.148	0.677	0.219	Irs
2014	Yunnan	0.463	0.698	0.664	Irs
2014	Zhejiang	0.974	1	0.974	Irs
2014	Chongqing	0.439	0.815	0.539	Irs
2014	Mean	0.571	0.851	0.67	
2015	Anhui	0.685	0.747	0.916	Irs
2015	Beijing	0.644	0.796	0.808	Irs
2015	Fujian	0.582	0.793	0.734	Irs
2015	Gansu	0.251	0.758	0.331	Irs
2015	Guangdong	0.644	1	0.644	Drs
2015	Guangxi	0.697	0.747	0.933	Irs
Year	Area	TE	PTE	SE	RTS
2015	Guizhou	1	1	1	-
2015	Hainan	0.119	0.699	0.171	Irs
2015	Hebei	0.482	0.629	0.766	Irs
2015	Henan	1	1	1	-
2015	Heilongjiang	0.406	0.882	0.461	Irs
2015	Hubei	0.653	0.735	0.889	Irs
2015	Hunan	0.487	0.51	0.954	Drs
2015	Jilin	0.689	1	0.689	Irs
2015	Jiangsu	1	1	1	-
2015	Jiangxi	0.837	0.862	0.97	Drs
2015	Liaoning	0.626	0.746	0.839	Irs

2015	Inner Mongolia	0.586	1	0.586	Irs
2015	Ningxia	0.048	0.857	0.056	Irs
2015	Qinghai	0.102	1	0.102	Irs
2015	Shandong	0.592	0.73	0.811	Drs
2015	Shanxi	0.783	0.963	0.813	Irs
2015	Shaanxi	0.534	0.671	0.795	Irs
2015	Shanghai	0.684	1	0.684	Irs
2015	Sichuan	0.99	1	0.99	Irs
2015	Tianjin	0.783	1	0.783	Irs
2015	Tibet	0.118	1	0.118	Irs
2015	Xinjiang	0.25	0.791	0.315	Irs
2015	Yunnan	0.487	0.627	0.776	Irs
2015	Zhejiang	0.983	1	0.983	Irs
2015	Chongqing	0.425	0.709	0.6	Irs
2015	Mean	0.586	0.847	0.694	
2016	Anhui	0.69	0.722	0.955	Irs
2016	Beijing	0.634	0.729	0.87	Irs
2016	Fujian	0.612	0.732	0.836	Irs
2016	Gansu	0.218	0.72	0.303	Irs
2016	Guangdong	0.723	1	0.723	Drs
2016	Guangxi	0.741	0.832	0.89	Irs
2016	Guizhou	1	1	1	-
2016	Hainan	0.116	0.72	0.162	Irs
2016	Hebei	0.574	0.608	0.945	Irs
2016	Henan	0.715	0.72	0.993	Drs
2016	Heilongjiang	0.376	0.823	0.457	Irs
2016	Hubei	0.664	0.695	0.956	Irs
2016	Hunan	0.486	0.505	0.964	Irs
2016	Jilin	0.645	0.962	0.671	Irs
2016	Jiangsu	0.964	1	0.964	Drs
2016	Jiangxi	0.936	0.94	0.996	Irs
2016	Liaoning	0.63	0.717	0.88	Irs
2016	Inner Mongolia	0.616	0.962	0.64	Irs
2016	Ningxia	0.061	0.895	0.068	Irs
2016	Qinghai	0.106	1	0.106	Irs
2016	Shandong	0.54	0.771	0.7	Drs
Year	Area	TE	PTE	SE	RTS
2016	Shanxi	0.782	0.884	0.884	Irs
2016	Shaanxi	0.537	0.641	0.837	Irs
2016	Shanghai	0.708	0.949	0.746	Irs
2016	Sichuan	0.995	1	0.995	Drs
2016	Tianjin	0.738	1	0.738	Irs
2016	Tibet	0.103	0.999	0.103	Irs
2016	Xinjiang	0.241	0.705	0.342	Irs
2016	Yunnan	0.597	0.638	0.936	Irs
2016	Zhejiang	1	1	1	-
2016	Chongqing	0.451	0.72	0.626	Irs
2016	Mean	0.587	0.825	0.719	

The comparison between "Table 3" and "Table 5" shows that after excluding external environmental factors and random error, tourism economic efficiency and return to scale have changed to varying degrees: (1) Effective comprehensive efficiency can only be found in Guizhou and Tianjin before the adjustment. After the adjustment, the number of effective areas did not significantly increase, but the regions changed. Except for Guizhou, Henan, Jiangsu and Zhejiang all achieved effective comprehensive efficiency in different periods, with the mean value of comprehensive efficiency rising from 0.467 before adjustment to 0.581 after adjustment. (2) As for the pure technical efficiency, the number of effective areas increased significantly, so did the number of effective areas, from 7 before adjustment to 11 after adjustment, and the mean value increased from 0.646 before adjustment to 0.841 after adjustment. (3) Scale efficiency decreases after adjustment. The average scale efficiency before adjustment is 0.758, and the average after adjustment is 0.694. Despite the small change in quantity, the regions where scale efficiency is effective have changed. Besides Guizhou, there are Henan, Jiangsu and Zhejiang. But Tianjin was ruled out. The scale efficiency decreases overall after the adjustment, indicating that the input scale of each region becomes more rational after the factor of management inefficiency is removed. The practice of blind expansion was abolished and management was increasingly valued. (4) The return to scale in most regions is increasing after adjustment, proving their strong development potential. (5) The mean value of the comprehensive efficiency of tourism in all regions ("Figure 1") shows that only Jilin and Tianjin experienced an obvious decline during 2014-2016, while the rest of the provincial administrative regions grow to varying degrees, among which Sichuan, Zhejiang, Guangdong, Henan and Jiangsu outperform. This proves the significant difference before and after removing the influencing factors and the accuracy of the three-stage DEA on the comprehensive efficiency of tourism.

The comparison of efficiency of administrative regions in each province from 2014 to 2016 shows changes to varying degrees as for efficiency before and after adjustment. (1) Obvious changes can be found as for comprehensive efficiency of tourism in each region after the adjustment. From 2014 to 2016, more than 65% of the regional comprehensive efficiency values were higher than those before the adjustment. Among them, Henan and Jiangsu provinces in 2014, 2015 and Zhejiang province in 2016 were at the optimal frontier after adjustment, that is, they reached the optimal input-output ratio, indicating estimation errors of traditional one-stage DEA. (2) Decomposition efficiency changes obviously. From 2014 to 2016, the pure technical efficiency of all regions exceeded the pre-adjustment efficiency. In terms of scale efficiency, only Guangdong, Jiangsu, Sichuan and Zhejiang showed an upward trend, indicating that the pure technical efficiency of all provincial regions in China was low and the scale efficiency was overestimated before the adjustment. (3) There are many changes in the state of returns to scale, especially in nearly half of the provincial regions during 2014-2015. The return to scale changes from decreasing to increasing, indicating that the return to scale of most provinces has not been optimized after adjustment.
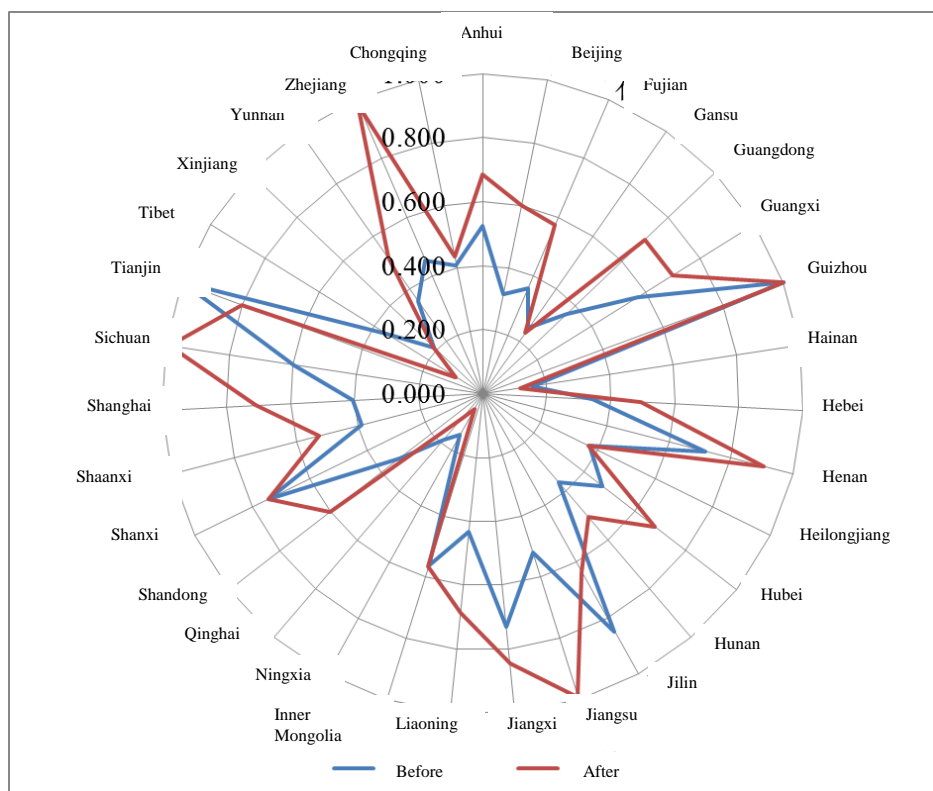
Figure 1 Average of comprehensive efficiency of tourism economy in each province and region during 2014-2016.

The principle of DEA model shows that the comprehensive efficiency can be decomposed into the product of scale efficiency and pure technical efficiency. In order to present the relationship between the decomposed efficiency and the comprehensive efficiency, the adjusted average value of each year's efficiency is made into a scatter plot. "Figure 2" shows that the pure technical efficiency scatter mostly concentrates above 0.6, indicating that

the pure technical efficiency plays a leading role in the comprehensive efficiency, while the scale efficiency scatter mostly concentrates on a straight line, indicating that the change trend of the scale efficiency is consistent with the change trend of the comprehensive efficiency. The tourism efficiency in China relies more on pure technical efficiency and advanced management technology to promote the overall development of tourism economic efficiency.
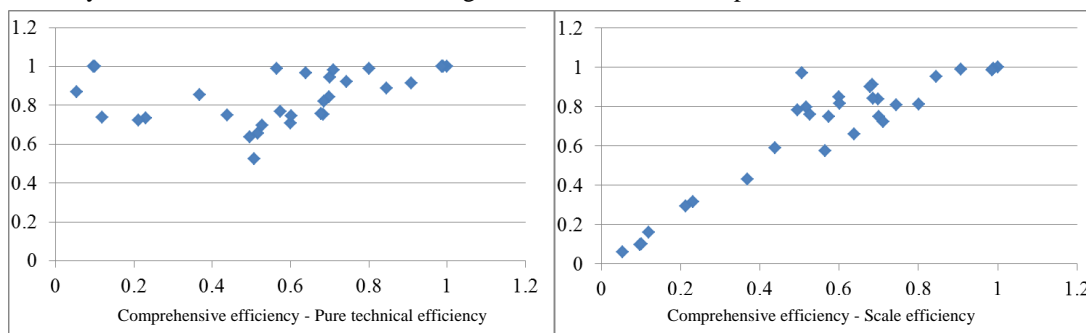

Figure 2 Effect of pure technical efficiency and scale efficiency on comprehensive efficiency.

4.4 Distribution of Tourism Economic Efficiency Types

Taking into account the mean value of pure technical efficiency and scale efficiency of each

province from 2014 to 2016 after removing the influencing factors, the 31 provincial regions in China are divided into four types ("Figure 3"). The first type is "best", that is, pure technical efficiency and scale efficiency are high, including areas at the

forefront. The second is "excellent", with scale efficiency higher than 0.6 and pure technical efficiency lower than 0.9, that is, areas that need to improve tourism operation and management. The third is "good", where the pure technical efficiency is higher than 0.9 and the scale efficiency is lower than 0.6, that is, places that need to increase the total number of tourism and expand the scale of tourism investment. The fourth is "potential type", with the pure technical efficiency lower than 0.9 and the scale efficiency lower than 0.6. Despite its low efficiency value, it still has great potential in tourism management and organization capacity and tourism scale.

"Figure 3" shows that 10 provincial-level regions of Sichuan, Zhejiang, Guizhou, Jiangsu, Jiangxi,

Shanghai, Tianjin, Guangdong, Jilin and Inner Mongolia belong to the first type. They are the regions with the best tourism development, leading other regions in China in terms of tourism investment scale, management level and tourism income. Fourteen provincial-level regions including Hunan, Hebei, Yunnan, Shaanxi, Hubei, Henan, Anhui, Beijing, Liaoning, Fujian, Shandong, Chongqing, Guangxi and Shanxi belong to the second category and need to improve their pure technical efficiency to improve the efficiency of tourism. Tibet and Qinghai belong to the third type and should improve the scale efficiency of tourism. Five provincial-level regions, Xinjiang, Heilongjiang, Gansu, Hainan and Ningxia, belong to the fourth type, and they have the greatest potential for the improvement of tourism management ability and tourism investment.
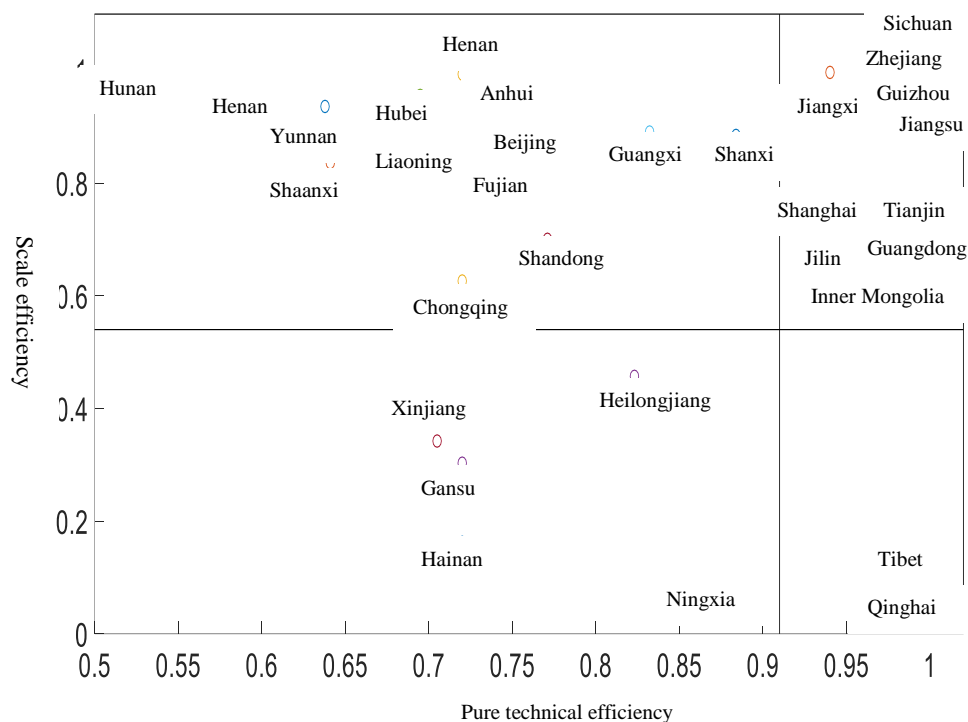

Figure 3 A regional type map of China divided by the mean of pure technical efficiency and scale efficiency during 2014-2015.

4.5 Analysis of Malmquist Calculation Results

In order to accurately evaluate the change of operation efficiency of tourism industry in various provincial regions over time, this paper applies the Malmquist index model to study the dynamic change trend, and draws the following four conclusions. This paper lists the average Malmquist index and its

decomposition in each provincial region from 2014 to 2016 ("Table 6").

- As for the national average, the technical efficiency change increased by 2.7%, and technical change increased by 12.3%. Such increase prompts an increase of 15.4% in national MI index, which was greater than 1, indicating a growing trend in the overall tourism operation efficiency during 2014-2015.

- The MI index of Henan and Qinghai is less than 1, lagging behind other provinces in China. Among them, Henan's technical efficiency shows a declining trend, indicating its low input-output efficiency level. The large-scale investment in tourism makes up Henan's flaws in tourism technical efficiency. Although the overall tourism volume of Henan ranks among the top in China, it relies on extensive growth, resulting in excessive waste of tourism resources. Therefore, Henan should cultivate advanced tourism management and accelerate the construction of high-quality tourism team. The reason why the MI index of Qinghai Province is too low lies in the insufficient change of scale efficiency. Qinghai Province is located in the western part of China, which emphasizes the exploitation of mineral resources over the investment scale of tourism. Therefore, it can further improve the operation efficiency of tourism industry by expanding tourism investment and improving the ability of resource allocation.

- Xinjiang, Beijing and Hebei all performed well, and the MI index ranked the top three among 31 provincial-level regions, indicating that the tourism industry in these three regions has a good development trend. They enjoy unique advantages in tourism resource allocation, so they should maintain the level of tourism management and development orientation.

- Shanghai outperforms in technical changes than other regions, which indicates that Shanghai, a previous shaker in tourism, takes tourism as the foundation of the city, which is closely related to a large number of high-level tourism talents in Shanghai. The tourism operation efficiency of Liaoning is the worst, and the pure technical efficiency is only 0.853, which is at the lowest level. It has deviated from the optimal tourism operation scale. Therefore, tourism investment should be appropriately reduced and tourism output level should be maintained to bring its tourism operation scale back to the right track.

Table 6. Average Malmquist index and decomposition of 31 provincial regions from 2014 to 2016

Area	MI Index	Technical efficiency change (EC)	Technical changes (TC)	Pure Technical Efficiency Change (PEC)	Scale efficiency change (SEC)
Anhui	1.13	1.015	1.114	0.959	1.059
Beijing	1.327	1.11	1.2	1.121	1.003
Fujian	1.258	1.079	1.167	0.971	1.114
Gansu	1.117	1.013	1.098	0.997	1.014
Guangdong	1.176	1.01	1.162	1.101	0.936
Guangxi	1.163	1.071	1.086	0.974	1.132
Guizhou	1.085	1	1.085	1	1
Hainan	1.081	0.977	1.107	0.954	1.031
Hebei	1.379	1.169	1.178	0.948	1.232
Henan	0.962	0.861	1.122	0.873	0.984
Heilongjiang	1.101	1.033	1.056	1.053	0.977
Hubei	1.136	0.971	1.166	0.913	1.063
Hunan	1.126	1.103	1.029	1.125	0.979
Jilin	1.154	1.031	1.117	1.009	1.019
Jiangsu	1.129	0.986	1.146	1	0.986
Jiangxi	1.2	1.146	1.049	1.082	1.061
Liaoning	1.042	0.898	1.157	0.853	1.052
Inner Mongolia	1.285	1.109	1.162	0.981	1.13
Ningxia	1.037	0.972	1.074	1.035	0.937
Qinghai	0.962	0.91	1.056	1	0.91
Shandong	1.065	0.913	1.165	1.076	0.849
Shanxi	1.208	1.083	1.118	0.981	1.103
Shaanxi	1.143	1.032	1.109	0.92	1.125
Shanghai	1.178	0.981	1.2	0.975	1.008
Sichuan	1.157	1.008	1.148	1	1.008
Tianjin	1.007	0.914	1.102	1.016	0.9
Area	MI Index	Technical efficiency change (EC)	Technical changes (TC)	Pure Technical Efficiency Change (PEC)	Scale efficiency change (SEC)
Tibet	1.029	0.962	1.061	1	0.962
Xinjiang	1.466	1.313	1.119	1.03	1.255
Yunnan	1.311	1.144	1.144	0.958	1.192
Zhejiang	1.211	1.013	1.195	1	1.013
Chongqing	1.154	1.017	1.134	0.943	1.081
Average	1.154	1.027	1.123	0.995	1.036

5. CONCLUSION

After factoring out the interference of external uncertain environment and random error, the deviation of comprehensive efficiency, scale efficiency and pure technical efficiency is corrected. Tourism economic efficiency has a significant change after adjustment, and the following conclusions are drawn. (1) Tourism efficiency and decomposition proportion change, comprehensive efficiency and pure technical efficiency increase obviously, while scale efficiency decreases as a whole. Pure technical efficiency dominates the value of comprehensive efficiency, and scale efficiency and comprehensive efficiency tend to be the same. (2) The comparison of returns to scale before and after adjustment shows that more than half of provincial regions' returns to scale have changed. The return to scale changes from diminishing to increasing, indicating that the return to scale is underestimated because of external uncertainty and random error. (3) Zhejiang, Jiangsu, Jiangxi, Sichuan and Guizhou are the excellent representatives of China's tourism industry. Although Sichuan and Guizhou are located in the remote western part of China, they are among the top thanks to rational tourism investment and effective operation and management. The tourism efficiency of Guangxi, Shanxi, Shanghai, Tianjin and Guangdong has always been at the top level in China. Jilin, Inner Mongolia and Shandong are at the moderate level, and the tourism efficiency of Ningxia, Hainan and Gansu provinces and regions is low. (4) The pure technical efficiency of all provinces in China is relatively high, and the change speed of scale efficiency and comprehensive efficiency is roughly the same. The pure technical efficiency of Qinghai and Tibet in the western region is very high, even higher than that of the eastern coastal provinces which are highly open, but the scale efficiency is low. Hunan province in central China has high scale efficiency but low pure technical efficiency. Hunan should reduce the negative effect of management inefficiency. The scale efficiency of the central and eastern regions is very high, but the pure technical efficiency of some provinces is not high. The eastern coastal provinces should strive to improve their own tourism management efficiency and realize the transformation from extensive growth to intensive growth. In short, further improving the efficiency of tourism requires sharp rise in efficiency of scale.

Based on the above analysis and empirical results, the following suggestions are proposed: (1) To promote the development of tourism at an appropriate scale. The low scale efficiency restricts the development of China's tourism industry. Therefore, the fixed assets of the tourism industry and the input of tourism practitioners in the central and western regions should be expanded, including travel agencies, star-rated hotels and other tourism companies, so as to increase the operation scale of the tourism industry. (2) Strive to change the external environment of tourism in each region. Economic capacity for development should be cultivated, and efforts should be made to raise the income level of residents and provide economic support for the development of tourism. All regions should also improve infrastructure, speed up urban construction, restore the ecological environment, and make good use of their own characteristic tourism resources. On the basis of not damaging the environment, new scenic spots should be developed, and the dividends of national tourism policies should be utilized to expand and strengthen the tourism industry. (3) Improve the efficiency of tourism operation. The development of science and technology also contributes a lot to the improvement of the efficiency of the tourism industry. Experience in engineering, hotels, urban planning and tourism marketing should be introduced to effectively improve the management ability and level of the tourism industry and improve the satisfaction of tourists.

AUTHORS' CONTRIBUTIONS

Yan Zhang was responsible for experimental design and wrote the manuscript, and Yeqin Fu analysed data and contributed to revising.

REFERENCES

[1] Hu Yaguang. Regional Evaluation of China's Tourism industry Efficiency and Its Deconstruction Analysis: An Enqrirical Study Based on SFA [J]. Contemporary Finance & Economics, 2019(10):107-119. (in Chinese)

[2] Meng Tao. Estimation of China's Tourism Efficiency and Analysis of Regional Differences Based on SFA Model [J]. Commercial Time, 2018(28):117-118. (in Chinese)

[3] Gu Jiang, Hu Jing. Appraisal and Model Foundation of the Efficiency in Tourism Production in China [J]. Journal of Tongji University (Social Science Section), 2008(4):93-98. (in Chinese)

[4] Liang Mingzhu, Yi Tingting. An Evaluation and Analysis of Tourism Efficiency in Different Cities and Regions of Guangdong Province [J].

Economic Geography, 2012(10):158-164. (in Chinese)

[5] Fried HO, Lovell CA K, Schmidt S S, Yanisawarng S. Accounting for Environmental Effects and Statistical Noise in Data Envelopment Analysis [J]. Journal of Productivity Analysis, 2002, 17(2): 157-174.

[6] Luo Dengyue. A Note on Estimating Managerial Inefficiency of Three-Stage DEA Model [J]. Statistical Research, 2012, 29(4): 104-107. (in Chinese)

[7] Jin Chunyu, Cheng Hao, Song Guangrui. Research on the Evaluation of China's Efficiency of Regional Tourism Industry Based on Three-stage DEA Model [J]. Tourism Tribune, 2012, 27(11): 56-65. (in Chinese)

[8] Wang Huiying. A Research on China's Tourism Efficiency: An Managerial and Environmental Perspective [J]. Tourism Science, 2014, 28(5): 31-40. (in Chinese)

[9] Wang Family, Cui Fengyu. Empirical Study on Measuring and Promoting Operation Efficiency of Regional Tourism in China [J]. Contemporary Economic Management, 2011, 33(6): 53-60. (in Chinese)

[10] Fang Yelin, Wang Fang. The Measurement and Optimization of Provincial Tourism Efficiency Considering Environment Variables in Mainland China: From the Perspective of Regional Economy and Traffic Location [J]. Tourism and Hospitality Prospect, 2019, 3(2): 40-53. (in Chinese)

[11] Cao Yanxue, Ma Lan. Evaluation on the Efficiency of Tourism Poverty Alleviation in China's Ethnic Regions Based on the Three-stage DEA [J]. East China Economic Management, 2017, 31(9): 91-97. (in Chinese)

[12] MALMQUIST S. Index numbers and indifference curves [J]. Trabajos de Estatistica, 1953, 4(2): 209-242.