Journal of Applied Sciences

ISSN 1812-5654

Journal of Applied Sciences 13 (17): 3387-3393, 2013
ISSN 1812-5654 / DOI: 10.3923/jas.2013.3387.3393

Forecast and Inspection of Inbound Tourism Market in China

Wei Min
School of Management, Xiamen University, Xiamen, 361005, China

Abstract: Experts and scholars from home and abroad have made multi-angle exploration and research on the tourism market forecasting, attained a lot of research achievement and also laid a solid foundation for the completion of this article. In China, as an important channel of foreign exchange earnings in trade in services, tourism industry has an important duty of expanding foreign exchange earnings. Derived from the expenditure of foreign tourists, international tourism revenue is a part of national income from other countries which flowing into the destination country or region and is used to compensate for the value of tourism products (services). The development of China's tourism industry and tourism market can not only attract domestic and foreign idle funds flowing to the tourism industry, but also to attract a large number of tourists that come from the country (territory) outside, thereby increasing our foreign exchange earnings. So, international tourism revenue becomes an important indicator to measure tourism especially inbound tourism. This study applied the regression model to analyzing and forecasting the rules and trends of tourism development.

Key words: Forecast, inspection, inbound tourism market

INTRODUCTION

The composition of the tourism industry including the travel agency industry, the lodging industry which as represented by the hotels, transportation, food and beverage industry, tour entertainment, sales industry of tourism goods and souvenir, tourism management agencies at all levels as well as industry organizations, that are, "food, shelter, transportation, travel, entertainment, shopping", the six elements of tourism. In order to objectively analyze and forecast the situation of China's tourism market, this study intends to explore on inbound tourism revenue. As tourism revenue is an important indicator to measure the degree of a city's tourism development, Alleyne (2006) believed it is conducive to the formulation of the government's macroeconomic policy and operating decisions of tourism enterprises through the forecast of tourism revenue. By source, tourism revenues can be divided into two parts, one part is from domestic tourism and the other is from inbound tourism foreignexchange earnings. Tourism foreign exchange earnings can easily be affected by some changes from the tourist destination and tourist source country, like the changes of exchange rate or other macroeconomic policies.

In China, as an important channel of foreign exchange earnings in trade in services, tourism industry has an important duty of expanding foreign exchange earnings. The development of China's tourism industry and tourism market can not only attract domestic and foreign idle funds flowing to the tourism industry, but also to attract a large number of tourists that come from the country (territory) outside, thereby increasing our foreign exchange earnings. So, international tourism revenue becomes an important indicator to measure tourism especially inbound tourism. International tourism revenue is foreign currency revenues that obtained by operating overseas tourists travel to the domestic country or region. Derived from the expenditure of foreign tourists, international tourism revenue is a part of national income from other countries which flowing into the destination country or region and is used to compensate for the value of tourism products (services). Expressed as the increase of the total social value of the tourist destination country or region, it is equivalent to the tourist destination countries (regions) output goods to other countries, is a special form of Foreign trade.

REVIEWS

Experts and scholars from home and abroad have made multi-angle exploration and research on the tourism market forecasting, attained a lot of research achievement and also laid a solid foundation for the completion of this article. Tourism market forecasting began in the 1960s and developed rapidly in the 1980s. Since the 1990s, it entered a new stage of development. Until now, the tourism market forecasting is still a hot issue in the academic circles at home and abroad. Before the 1990s, tourism market forecasting mostly used traditional qualitative or quantitative research methods. After 1990s, with the development of artificial intelligence theory and its wide application in various industries, scholars introduced these methods gradually into the prediction and evaluation of tourism market research. The accuracy and operability of tourism market forecasting also have improved.

Among changes of the tourism market, many factors are difficult to predict, such as emergencies (including abrupt climate changes of tourism destinations, disease and political instability, etc.,). Thus, tourism market forecasting should have considerable flexibility. In the 1960s, some international scholars began to make the research of tourism market forecasting. In terms of forecasting methods, foreign research literature about the prediction of tourist market (demand) is abundant and has initially formed a set of systematic system. According to the research results at home and abroad and from the angle of tourism market measurement, the methods of tourism market forecasting can be divided into two categories, qualitative forecast methods (expert consultation, the survey of tourists' intentions, the experimental simulation and subjective probability) and quantitative forecast methods (time series forecasting, regression analysis and forecasting, Markov chain forecasting, gray system model forecasting, neural network forecasting, analytic hierarchy process and fuzzy mathematical model forecasting). In order to facilitate the analysis, this study summarizes these into three categories, time series forecasting, the regression model forecasting and the combination forecasting. Makridakis (1986) compared the results of different tourism market demand forecasting models and found that the researchers chose different models to predict, based on their own research experience, which led to different conclusions with the same data entry. After increasing the complexity of the prediction model, its accuracy doesn't increase correspondingly. So, simple prediction model is often not necessarily the shortcomings of forecasting, time series forecasting model in the tourism market forecasting are generally relatively simple and with low costs. Jackman and Greenidge (2010) used time series models to explain and predict the number of visitors from the key customer source markets to Barbados and its change trends and proposed a valuable perspective from the research of tourists behavioral factors. Goh and Law (2001) used the data of Hong Kong's main source countries to establish its model and make evaluations. They selected lots of time series models, including the moving average method (3 month moving average method, 12 month moving average method), exponential smoothing (simple exponential smoothing, Holt exponential smoothing, Winter exponential smoothing method), seasonal autoregressive integrated moving average method (SARIMA: Multiplicative seasonal ARIMA), autoregressive integrated moving average method (ARIMA), autoregressive moving average method with intervention and used Sale U-factor, average absolute percentage error (MAPE), Mean Absolute Deviation (MAD), mean square mistaken (MSE), Root Mean Square Error (RMSE), percentage of the root-mean-square errors(RMSPE) to test the accuracy of the forecasting (Aslan *et al.*, 2009).

In the numerous research literatures, the regression model is most widely used in the tourism market (demand) forecasting. It uses Ordinary Least Squares (OLS) to establish regression analysis model based on the sample values (real values) of the relevant variables, then use hypothesis testing to determine whether the set of regression analysis model is significant (Abbott *et al.*, 2012). Finally, it uses the regression model to analyze and forecast the rules and trends of tourism development. In addition, we should also recognize that, when we make the tourism market forecasting for a specific destination, we must take the characteristics of the various alternative forecasting methods and the characteristic factors of specific prediction problems into account and choose the best options.

ESTABLISHMENT OF THE MODEL

Samples sequence data preprocessing: Forecast is a quantitative method that relies on historical data to analyze the future trends (Coshall, 2005). Through the collection and collation of large amounts of data and then compare the running guidance indicators of travel demand and relevant variables, the laws governed the operation (Davidson and MacKinnon, 1998). Thus, it is very necessary to collect raw date. The raw data this study used is mainly come from the China Statistical Yearbook and the website of the National Bureau of Statistics of the People's Republic of China,and the data of 2011 is from the webs ite of the National Bureau of Statistics of the People's Republic of China. We collect data about China's inbound tourism market from 1992-2011 and the Table 1 lists the data of China's foreign exchange earnings on inbound tourism in 1992-2011. They can be regarded as one continuous random time sequence that formed over time. By using the method of random time series analysis, combined with Eviews 6.0 software and correction of abnormal values, we hope to get a better prediction results from the forecast of China's inbound tourism market. China's inbound tourism revenue in 1992-2011 is as shown in Table 1.

China's inbound tourism market trends can be visually seen from the time sequence diagram. So, we used Eviews6.0 software to draw this time sequence diagram, the result is shown below.

Table 1: China's inbound tourism revenue in 1992-2011 Unit: 100 million dollars

Years	Revenue	Growth rate (%)	Years	Revenue	Growth rate (%)
1992	39.50		2002	203.85	14.52
1993	46.80	18.48	2003	174.06	-14.61
1994	73.23	56.47	2004	257.39	47.87
1995	87.00	18.80	2005	292.96	13.82
1996	102.00	17.24	2006	339.49	15.88
1997	121.00	18.63	2007	419.00	23.42
1998	126.00	4.13	2008	408.43	-2.52
1999	141.00	11.90	2009	396.75	-2.86
2000	162.24	15.06	2010	458.00	15.44
2001	178.00	9.71	2011	484.64	5.81

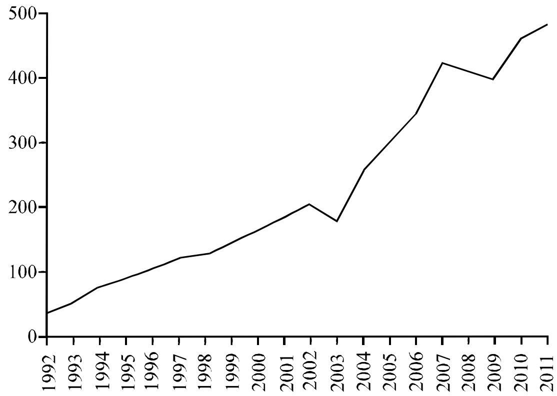


Fig. 1: Trend graph of China's inbound tourism revenue in 1992-2011, Unit: 100 million dollars

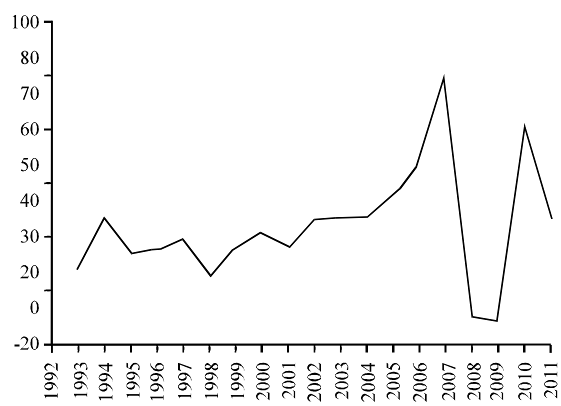


Fig. 2: Trend graph of time series after the first-order differential

From Table 1 and Fig. 1, we can draw such a result that China's inbound tourism market developed quickly and inbound tourism revenue showed a steady growth trend though there were some fluctuations in several years.

From Table 1, a big fluctuation in 2003 that China's inbound tourism revenue dropped sharply due to the impact of SARS. In order to eliminate the adverse effects which come from the 2003 data volatility to forecast, we smooth the raw data of 2003. In actual operation, we use the arithmetic average of the two years before and after 2003 instead of the raw data in 2003, that is adjusting the raw data in 2003 to $ 23.062 billion.

From Fig. 1-5, it is obviously to see that $\{y_t\}$ has a clear upward trend which said $\{y_t\}$ a non-stationary time series. To obtain the stationary time series, we use the differencing method. But we should choose different differencing processes according to different time series that own its characteristics. From the trend graph of $\{y_t\}$, we can see that the sequence contains a curve trend. Thus, usually, we can use low-level (second-order or third-order) differential to eliminate the curve trend and adjust $\{y_t\}$ into a stationary time series.

After the first-order differential of , we get sequence and the graph of is shown below (Fig. 2).

We did parametric test and runs test on $\{r_t\}$ and drew conclusion that this sequence a non-stationary time series. So, we continue the second-order differential of $\{y_t\}$ and got another new sequence $\{x_t\}$ whose graph is shown by Fig. 3.

Through direct observation from Fig. 3, $\{x_t\}$ looks almost relatively stable. Next, we did a unit root test on $\{x_t\}$ and test results was shown in Fig. 4.

The result of unit root test proves that $\{x_t\}$ a stationary time series. Next,we can do time series analysis of stationary time series $\{x_t\}$.

Model identification and model order: As mentioned before, we can analyze and identify the type of the model through investigating the Autocorrelation coefficients (ACF) and the partial correlation coefficient (PACF) of the sample. There are three types of stationary time series and we listed each type's statistical characteristics of both autocorrelation function and partial autocorrelation function.

Through Eviews6.0, we calculated the sample autocorrelation function and sample partial autocorrelation function of the sequence $\{x_t\}$. The result is shown by Fig. 5.

Such conclusion can be drawn from Fig. 5 and autocorrelation coefficients are all fall within the confidence interval and are close to zero; partial autocorrelation coefficients are all fall within the confidence interval and are close to zero, too. So, we can conclude again that the original sequence is a stationary sequence. Usually, for a sample contains about 20 data, the maximum lag number of the model often take 4 or 5. If there are too many coefficients, it will improve the model's complexity and it may reduce the precision of the model. Thus, we set p between 1-5, set q between 1-3, taking the coefficient of determination and adjusted coefficient of

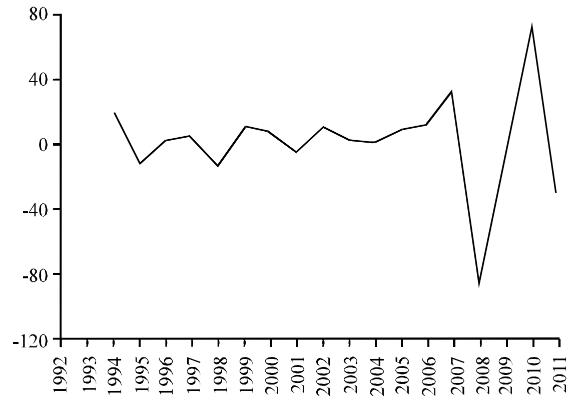
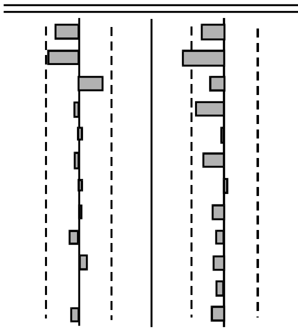


Fig. 3: Trend graph of time series after the second-order differential

		t-statistic	Prob*
Augmented dicky-fuller test statistic		-6.596007	0.0004
Test critical values:	1% level	-4.667883	
	5% level	3.733200	
	10% level	3.310249	

Fig. 4: Result of unit root test



Autocorrelation	Partial correlation		AC	PAC	Q-stat	Prob
		1	-0.327	-0.327	2.2597	0.133
		2	-0.433	-0.604	6.4770	0.139
		3	-0.314	-0.194	8.8488	0.031
		4	-0.051	-0.400	8.9168	0.063
		5	0.035	-0.038	8.9505	0.111
		6	-0.063	-0.291	9.0691	0.170
		7	0.029	0.029	9.0973	0.246
		8	0.031	-0.171	9.1313	0.331
		9	-0.122	-0.113	9.7231	0.373
		10	0.114	-0.155	10.309	0.414
		11	0.014	-0.086	10.318	0.502
		12	-0.099	-0.180	10.912	0.536

Fig. 5: Sample autocorrelation function and Sample partial autocorrelation function of $\{x_t\}$

Table 2: ACF and PACF theoretical model of ARMA (p,q)

Model	ACF	PACF
AR (p)	Tailing, attenuation tends to zero (geometric or oscillating type)	It will be truncated after a p-order
MA (q)	It will be truncated after a q-order	Tailing, attenuation tends to zero (geometric or oscillating type)
ARMA (p, q)	Tailing and attenuation tends to zero after a q-order geometric or oscillating type)	Tailing and attenuation tends to zero after a *p*-order geometric or oscillating type)

determination and AIC as indicators, we finally select the model ARMA (2,1) which is the relatively appr opriate model after repeated fitting.

Vaviable	Coefficient	Std. Error	t-statistic	Prob.
AR (1)	-0.318906	0.290504	-1.097766	0.2922
AR (2)	-0.747008	0.313302	2.384311	0.0330
MA (1)	-0.382390	0.388842	-0.983407	0.3434

R-squared	0.557480	Mean dependent var	0.804375
Adjusted R-squared	0.489400	S.D. dependent var	33.05783
S.E. of regression	23.62189	Akaike info criterion	9.329586
Sum squared resid	7253.920	Schwarz criterion	9.474446
Log likelihood	-71.63668	Hannan-Quinn criter.	9.337004
Durbinn-watson stat	1.863268		
Inverted AR roots	-0.16+0.851	-0.16-0.851	
Inverted MA roots	0.38		

Fig. 6: Established graph of model

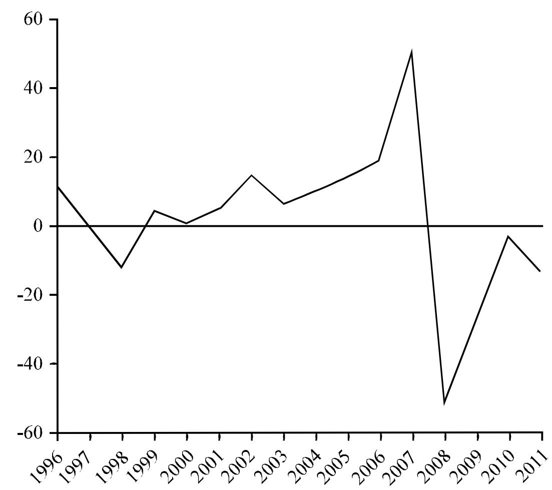


Fig. 7: Residual time series trend graph of model

Model parameter estimation: After a preliminary identification of the type of the model, the next job is to estimate the model parameters. This study use Eviews6.0 to estimate the parameters of the model.

Model ARMA (2, 1) is estimated using Eviews6.0 and we get the estimation result and the residual schematic diagram. As shown by Fig. 6 and 7.

Thereby, we can obtain the mathematical expression of the Model 1:

$$X_t = -0.3189X_{t-1}-0.7470X_{t-2}+a_t-0.3824a_{t-1} \quad (1)$$

Model test: In order to analyze the reliability of the model, the model must be tested. We identify and estimate the model ARMA (p, q) under such an assumed condition that random interference term $\{\alpha_t\}$ is a white noise sequence. So, if the estimated model is correct, the model's random interference term $\{\alpha_t\}$ will be a white noise sequence (Tideswell *et al*., 2001). While $\{\alpha_t\}$ is not a white noise sequence, it means something wrong with the identification and estimation of the model so that we need try again. In the test, the key point is identifying whether the residual sequence is an autocorrelation sequence or not. We can use statistic Q_{LB} mentioned above to do χ^2 test on residual series.

TEST OF INBOUND TOURISM REVENUE FORECAST

As previously mentioned, before doing predictive accuracy test and forecast covers inspection on the selected time series model ARMA (2, 1), we need to build a model of other types to compare the predictive values. Traditional regression analysis is selected in this study to build a model of inbound tourism market and predict (Li and Lo, 2004).

Traditional regression analysis method modeling: Originated in the 19th century Gaussian least squares method and formed the standard mathematical theory of statistical analysis in the early 20th century, regression analysis is one of the most widely used mathematical analysis methods in the statistical analysis (Ucar and Omay, 2009). Regression analysis is to look for the inherent regularity of variables, based on the historical observation data collection and through computer software technology, in order to establish a expression that representatives the interdependence between a variable with the other one or several variables (Smeral and Wuger, 2005). In the regression analysis forecast, the key points are the following issues: (1) Determine whether there is a mutual relationship between the several variables by observing the sample, if there is, it means that we can establish a regression equation, (2) Test the confidence level and confidence interval of the regression equation, (3) Predict the value of the dependent variable through the regression equation, as well as calculate the prediction accuracy (Li, 2000). Regression analysis mainly has two methods, one is a simple regression analysis and the other is multiple regression analysis (Gouveia *et al*., 2005). A simple regression analysis is to analyze the relationship between an independent variable and a dependent variable while multiple linear regressions are to analyze the relationship among several independent variables and a dependent variable. In order to facilitate analysis, this article omitted the microstructure factors of tourism market and select a simple regression model to do relative forecast.

By using Eviews6.0, we fit the inbound tourism revenue series $\{y_t\}$ and the time series $\{t\}$ (t = 1, 2, 3,...). According to $\{y_t\}$ and $\{t\}$, we can draw scatter plot of the relationship between them.

According to the Scatter plot above, we finally obtained optimal fit function (Model 2) after repeated fitting:

$$\hat{y} = 50.3968+1.7092t^2-0.0305t^3 \quad (2)$$

Fitting result is shown in Fig. 8.

In the regression analysis, the coefficient of determination (R^2) can generally be used to measure the fit degree of the prediction model, but in practice, it is often difficult to obtain accurate distribution function of R^2 because R^2 is easily influenced by factors such as the number of independent variables and samples (Huang and Min, 2002). Therefore, in order to reach a more comprehensive understanding of the coefficient of determination, some scholars studied the sampling nature of R^2. From Fig. 9, both the coefficient of determination (R^2) of this model and adjusted coefficient

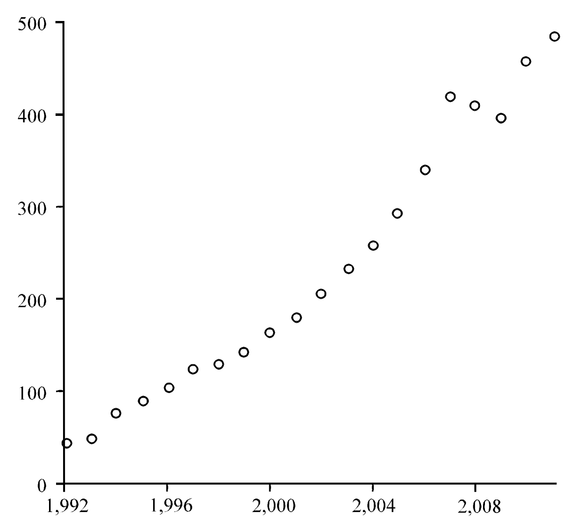


Fig. 8: Scatter plot of the relationship between $\{y_t\}$ and $\{t\}$

Variable	Coefficient	Std. Error	t.Statistic	Prob.
C	50.39680	8,733260	5.770674	0.0000
T^2	1.709211	0.205290	8.325820	0.0000
T^3	-0.030510	0.010482	-2.910572	0.0097

R-squared	0.984999	Mean dependent var	228.3950
Adjusted R-squared	0.983234	S.D. dependent var	144.8593
S.E. of regression	18.75674	Akaike info criterion	8.838464
Sum squared resid	5980.860	Schwarz criterion	8.987824
Log likelihood	-85.38464	Hannan.quinn criter.	8.867621
F-statistic	558.1326	Durbin-watson stat.	1.312509
Prob. (F-statistic)	0.000000		

Fig. 9: Fitting results from the traditional regression analysis method

of determination reach to over 98%, so that the model pass the adaptive test with a good fitting effect. The Results, which from the t-test and Significant test of the parameters of the model, show that the model is reasonable and can be used to further forecast of inbound tourism market.

Predictive accuracy test: To facilitate the analysis, in this study, we used the average absolute relative error Eq.:

$$MAPE = \sum \frac{|(y_i - \hat{y}_i / y_i|)}{n} \quad (3)$$

This equation above can be used to do accuracy test on Model 1 and 2. And the test result is that the average absolute relative error of model ARMA (2,1) is 6.03% while the traditional regression model's is 8.99%. It can be seen that, when doing forecast on inbound tourism market, the predictive accuracy of time series model is higher than traditional regression model's.

Forecast covers test: Chu (1998) set a way to assess the forecasting performance is forecast covers test. Conceptually, a different forecast model has a different connotation and what the forecast covers test want to determent is whether the predictive value of the model contains all relevant information in the competitive model. The method of forecast covers test is relative to Combination of Forecasts. Koc and Altinay (2007) believed combination of forecasts is that, when predict a target, use two or more different prediction methods. It is a combination of several forecasting methods as well as a integration of several qualitative thinking. But for the forecast of tourism market, it is often a combination of qualitative method and quantitative method. The main purpose of the combination is to take advantages of various methods and to improve the prediction accuracy as far as possible . For example, in the period with big macroeconomic fluctuations, it is difficult to create a single prediction model that is good fitting to the frequent fluctuations.

Set y_t as series that observed actually, set (f_{1t}, f_{2t}) as the two sets of corresponding predicted values, one set of the values is from Model 1 and the another set is from Model 2 and the Model 1 is the selected model while the Model 2 is the competitive one. After calculating the value-weighted average of the two sets of predicted values, we can obtain the combined predictive value (f_{ct}) as following:

$$f_{ct} = (1-\lambda)\, f_{1t}+\lambda f_{2t} \quad (4)$$

In the equation above, λ is the weight of the two kinds of prediction model. In general, the predicted value having a smaller error variance should be given a higher weight. Therefore, if $\lambda \to 0$, then it indicates that the predictive value f_{1t} is much accurate and has smaller error variance; but if $\lambda \to 1$, it indicates that the predictive value f_{2t} is much accurate. If the weight is neither 0 nor 1, results predicted by combination forecast will be better than those predicted by the original forecast.

Therefore, by estimating the weight R, we can judge whether it is necessary to form a combination forecast.

To illustrate how to estimate λ, we set $e_{it} = y_t - f_{it}$, I = 1,2 as the error of the two sets of predicted values and set ε_t as the error of the value predicted by combination forecast, therefore:

$$y_t - f_{ct} = [(1-\lambda)f_{1t}+\lambda f_{2t}] = e_{1t}+\lambda\,(f_{1t}-f_{2t}) = \varepsilon_t \quad (5)$$

After transposition:

$$e_{1t} = \lambda\,(f_{1t}-f_{2t})+\varepsilon_t \quad (6)$$

CONCLUSION

Therefore, on the basis of the analysis of the time series model and the regression model, this study learns from the inbound tourism market forecasting experience of experts in this field. Considering minimizing the forecasting error for inbound tourism market in China, the combination forecasting methods can be applied to predict the relevant contents of inbound tourism market. This study is compared with the traditional method of regression analysis, increases tests for the forecasting accuracy and coverage of the time series model and uses combination forecasting methods according to the test results, in order to improve the forecasting accuracy and operability. So, this study attempts to provide a new thought for the construction and evaluation of China's tourism market forecasting model and on this basis, to provide practical policy recommendations for the scientific development of China's tourism industry.

REFERENCES

Abbott, A., G. De Vita and L. Altinayc, 2012. Revisiting the convergence hypothesis for tourism markets: Evidence from Turkey using the pairwise approach. Tourism Manage., 33: 537-544.

Alleyne, D., 2006. Can seasonal unit root testing improve the forecasting accuracy of tourist arrivals? Tourism Econ., 12: 45-64.

Aslan, A., M. Kaplan and F. Kula, 2009. International tourism demand for Turkey: A dynamic panel data approach. Res. J. Int. Stud., 9: 65-73.

Chu, F.L., 1998. Forecasting tourist arrivals: Nonlinear sine wave or ARIMA. J. Travel Res., 36: 79-84.

Coshall, J.T., 2005. A selection strategy for modelling UK tourism flows by air to European destinations. Tourism Econ., 11: 141-158.

Davidson, R. and J.G. MacKinnon, 1998. Graphical methods for investigating the size and power of hypothesis tests. Manchester School, 66: 1-26.

Goh, C. and R. Law, 2011. The methodological progress of tourism demand forecasting: A review of related literature. J. Travel Tourism Market., 28: 296-317.

Gouveia, P.M.D.C.B. and P.M.M. Rodrigues, 2005. Dating and synchronizing tourism growth cycles. Tourism Econ., 11: 501-515.

Huang, J.H. and J.C.H. Min, 2002. Earthquake devastation and recovery in tourism: The Taiwan case. Tourism Manage., 23: 145-154.

Jackman, M. and K. Greenidge, 2010. Modelling and forecasting tourist flows to barbados using structural time series models. Tourism Hosp. Res., 10: 1-13.

Koc, E. and G. Altinay, 2007. An analysis of seasonality in monthly per person tourist spending in Turkish inbound tourism from a market segmentation perspective. Tourism Manage., 28: 227-237.

Li, Y., 2000. Geographical consciousness and tourism experience. Ann. Tourism Res., 27: 863-883.

Li, Y.P. and R.L.B. Lo, 2004. Applicability of the market appeal-robusticity matrix: A case study of heritage tourism. Tourism Manage., 25: 789-800.

Makridakis, S., 1986. The art and science of forecasting An assessment and future directions. Int. J. Forecasting, 2: 15-39.

Smeral, E. and M. Wuger, 2005. Does complexity matter? Methods for improving forecasting accuracy in tourism: The case of Austria. J. Travel Res., 44: 100-110.

Tideswell, C., T. Mules and B. Faulkner, 2001. An integrative approach to tourism forecasting: A glance in the rearview mirror. J. Travel Res., 40: 162-171.

Ucar, N. and T. Omay, 2009. Testing for unit root in nonlinear heterogeneous panels. Econ. Lett., 104: 5-8.