The Social Sciences 11 (19): 4617-4621, 2016
ISSN: 1818-5800

Forecasting the Chinese Tourist Arrivals to Thailand the Time Series Approach

[1]Xue Gong, [1]Songsak Sriboonchitta and [2]Siwarat Kuson
[1]Faculty of Economics, Chiang Mai University, Chiang Mai, Thailand
[2]Faculty of Economics, Maejo University, Chiang Mai, Thailand

Absrtact: The ARIMA Model is good for tourism demand forecasting when the uncertainty is low. However, when several uncertainty events happened, such as Chinese holidays, political turmoil and structural changes in our study, the model reacts very weakly. After comparing the out-of-sample forecast performances of ARIMA and Seasonal ARIMA (SARIMA) Models, we suggest that the SARIMA Model produce a more stable forecast especially when the structural change occurs and high uncertainty appears. We recommend the policy makers and relevant travel decision section to use SARIMA method to conduct the tourist forecasting.

Key words: Chinese tourist arrivals, Thailand, forecasting, ARIMA method, SARIMA method

INTRODUCTION

Tourism is an important sector to Thailand. Year by year there are mass tourists flowing into Thailand due to its special culture, nice atmosphere and kind residents (Li and Zhang, 1997). Before, the major tourist origin countries were Australia, Japan, Korea, Singapore and Malaysia, etc. (Song *et al*., 2003) but recent situations have changed. With China's GDP per capita rising to 4,433 USD in 2009, people have more money to enjoy their lives. The number of outbound Chinese tourists has increased from 452.510 in 1997-1,122,219 in 2009. Moreover with the recent hit movie Lost in Thailand. Chinese tourists have been booming into Thailand since 2012. Thailand has become one of the hottest outbound tourist destinations in China. The Tourism Authority of Thailand (TAT) announced that more than 1.5 million Chinese would visit by 2015. In the first week of April 2014 which is the Thai New Year, Chinese tourists brought 0.9 billion Thai Baht profit.

China has a large population. Furthermore with the effects of Word of Mouth (WOM), there are continually Chinese tourists coming to visit Thailand. Many Chinese even visited Thailand several times. The phenomenon is recent and can potentially last long into the future. It is quite important to generate accurate forecasts of future trends in tourist flows from China.

Tourism demand is the foundation of all tourism planning; especially when we know the tourism demand from a specific major country with its own special characteristics such as China, the problem becomes more important. The tourism marketers are interested in tourists demand for the several reasons. Firstly, the companies, such as the airlines, hotels and tour agencies, are interested in the demand for their products by tourists. Since the demand is a key to determine the business profitability, it will be important to the destination marketers. Secondly, tourism forecasting is important because of tourism investment, especially investments in destination infrastructures such as airports and highways. Therefore, accurate forecasts of the demand in the tourism sector of the economy will largely help the tourism sectors.

For tourism arrivals, there exist the structural changes due to many reasons. For example the source countries' economic booming such as China and India (Lee and Chang, 2008), visits to sites where recent movies and dramas have been filmed, such as Korean soap opera or The Lord of the Rings filmed in New Zealand (Connell, 2012; Rittichainuwat and Rattanaphinanchai, 2015) and events oriented to tourist arrivals, such as the Olympic Games (Lee and Taylor, 2005; Fourie and Santana-Gallego, 2011). To efficiently model and forecast tourist demand with structural changes, we find that there is limited literature to deal with these pheromones although it always happens, in our observations, there is only one piece of literature (Song *et al*., 2011). They use the structural time series models.

The forecasting method we selected is ARIMA and SARIMA Models. The major contribution to the literature in this study is that this is the first study to compare the performance of ARIMA and SARIMA Models during the period of structural changes and high uncertainty in the Chinese tourism demand analysis.

Corresponding Author: Xue Gong, Faculty of Economics, Chiang Mai University, Chiang Mai, Thailand

Literature review: Tourism forecasting is important to tourism planning and tourism policy (Song *et al*., 2003; Song and Witt, 2006). Many national travel institutions such as World Travel and Tourism Council (WTTC) together with Oxford Econometrics (OE) give out their expert opinions and also tourists demand forecasts. The Thailand Tourism Authority (TTA) and China National Tourism Admiration (CNTA), for example, have also been providing forecasts of tourist arrivals for their respective governments for both long-term and short-term planning and policymaking purposes.

It is widely known that tourism forecasts may be generated by either a quantitative approach (such as the econometric method) or qualitative approaches (such as expert observation or expert opinion). There are many studies focusing on the quantitative forecasting method (econometric approaches), such as cointegration analysis (Kulendran, 1996; Lim and McAleer, 2001), Vector Autoregression (VAR) Model (Witt and Witt, 1992; Song *et al*., 2006) and recent use of time series based models such as (Shareef and McAleer, 2005; Chan *et al*., 2005). Since, we only want to know the demand of Chinese tourists to Thailand, we will adopt the univariate time series method in our study such as the Autoregressive Integrated Moving Average (ARMA) model and the seasonal ARMA (SARMA) Model. The similar studies can be found in Song *et al*. (2000), Song and Li (2008) and Chu (2009).

The previous studies account for the impact of one-off events, tourists' taste changes and structural changes on the demand for tourism by dummy variables (Lim, 1997) or some modified structural change econometric methods. However, in our study, we focus on forecasting. The more simple methods could be more suitable for our forecasting purposes.

The econometric method we selected here is the seasonal ARIMA (SARIMA), this model has already been adopted in different literature such as traffic man-agement (Williams and Hoel, 2003), production value (Tseng and Tzeng, 2002) and tourism demand (Chang *et al*., 2009; Brida and Garrido, 2013). For our study, we selected this model since it both can capture level, long-term trend and seasonality all at once. We compare the performances of SARIMA and ARIMA Models in Chinese tourist arrivals to Thailand.

MATERIALS AND METHODS

The ARIMA based models, proposed by Box and Jenkins (1970) dominate the tourism demand literature. Depending on the frequency of the time series, either simple ARIMA or seasonal ARIMA (i.e., SARIMA) models could be used with the latter gaining an increasing popularity over the last few years as seasonality is such a dominant feature of the tourism industry that decision makers are very much interested in seasonal variation in tourism demand. With regard to the forecasting performance of the ARIMA and SARIMA Models, empirical studies present contradictory evidence. For example, Cho (2003) showed that the ARIMA Model outperforms two other time series models in all cases. Goh and Law (2002) suggested that the SARIMA Models outperform eight other time series methods while the non-seasonal (simple) ARIMA Model's performance was above the average of all forecasting models considered. However, Smeral and Weber (2005) found that the ARIMA or SARIMA Model cannot even outperform the Nave 1 (no-change) model. We will try and compare the ARIMA-based model by the tourism demand analysis in the following.

Autoregressive integrated moving average (ARIMA) method: In the ARIMA Model, a series is transformed to a condition of covariance stationary and then it is identified, estimated, diagnosed and at last forecast (Carey and Rob, 2002). The ARIMA Model can be represented by ARIMA (p, d, q), parameter p represent the order of (p) autoregrssion. (d) is the differencing time and (q) is for the moving average. The forecasts can be represented as:

$$F_{t+1} = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_p \varepsilon_{t-p} \quad (1)$$

According to Box and Jenkins (1970), the sample for the ARIMA is at least 50 observations. "An ARIMA Model can be viewed as a "filter" that tries to separate the signal from the noise and the signal is then extrapolated into the future to obtain forecasts" (Box and Jenkins, 1970).

Seasonal ARIMA (SARIMA) method: If the data exhibits seasonality, a multiplicative SARIMA model can be very useful. The seasonal components of the ARIMA Model are expressed by ARIMAs (P, D, Q) where capitalized letters denote the seasonal components of the model and s shows the order of periodicity or seasonality. Equivalently, SARIMA is an extension of ARIMA where seasonality in the data is accommodated using seasonal difference. The full formulation of a multiplicative seasonal ARIMA Model use the general form ARIMA (p, d, q)s(P, D, Q). ARIMA (1, 0, 1)(0, 1, 1) (Lim, 1997) with constant This model incorporates a constant term representing the long-term trend. Hence, this model

Table 1: Summary statistics for tourist Arrivals in 1997-2015

Tests	(Whole period) 1997-2014	1997-2008 (Period I)	2009-2014 (Period II)
Mean	114300.18	62637.96	217624.7
Median	71709.5	62740.5	181822
Maximum	513441	113990	513441
Minimum	6490	6490	35370
Std. Dev.	110528.13	20517.07	141089.5
Skewness	2.04	0.01	0.53
Excess Kurtosis	3.29	0.02	6.08
Jarque-Bera	253.61(***)	0.03	0.05
Observations	216	144	72

assumes a more stable trend than the ARIMA (0, 1, 1) (0, 1, 1) (Lim, 1997) Model and that is the main difference between them. If we add the indicated MA (1) and SMA (1) terms to the preceding model, we obtain an ARIMA (1, 0, 1)×(0, 1, 1) model with constant, whose forecasting equation is:

$$\hat{Y}_t = \mu + Y_{t-12} + \phi_1(Y_{t-1} - Y_{t-13}) - \theta_1 e_{t-1} - \Theta_1 e_{t-12} + \theta_1 \Theta_1 e_{t-13}$$

RESULTS AND DISCUSSION

Data description: The empirical study in this paper is based on inbound tourism demand to Thailand from recent major international markets China. It is the biggest feeder market, accounting for more than 14% of total international tourist arrivals in 2014 and also the highest revenue generating market for Thailand. The data set examined in this study is monthly total tourist arrivals from China from January 1997 to December 2014 which are obtained from EcoWin. We plot the whole data set in Fig.1. The Figure shows that there is structural change in 2009. With China's GDP per capita rising to 4,433 USD in 2009, people have more money to enjoy their lives. Chinese tourists have been booming into Thailand since Dec 2012. Although between 2013 and 2014, the Thailand turmoil happened making some drops in the data series.

To illustrate structural change clearly, we plot the recent changes in Fig. 1 and summarize the different periods in Table 1. In the last two years of period II (January 2013-December 2014), the average monthly tourist is around six times of the first 2 years of period I (January 1997-December 2008). Another interesting thing is that although in the whole period the data exhibits the non-normal distribution (it can be checked from Jarque-Bera statistics), however in each period it is normal distribution. The evidences of structural change are rigorous.

The drawback of arima method when the high uncertainty exhibit: The ARIMA Model works well when the uncertainty is low (Carey and Rob, 2002; Lee *et al*., 2008; Chang *et al*., 2009). However, one significant problem of SARIMA Model for forecasting is relatively

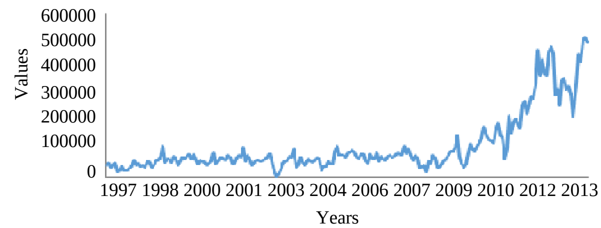

Fig. 1: The chinese tourist arrivals to thailand

inaccurate forecasts during high uncertainty periods, such as the Chinese holiday, Thai political turmoil and also the period of structural changes in our case. Some problems even repeat every year where a Chinese holiday comes.

To illustrate the severity of these inaccurate forecasting problems during the high uncertainty period, we investigated how ARIMA and SARIMA methods tackle it. When we presented one-period-ahead forecasts with a rolling window, there is a natural question raised. If we use the long horizon span data set, that is, as much as data we can get to forecast the post-structure change tourist arrivals will the result be good or should we use the short but close data span? Here we answer the question by using a different sample size. We presented one-period-ahead forecasts with a rolling window of size 60 (80) or 108. We used the in-sample data from 1997-2008 which is the half of our data, the out-of sample forecasts is from 2009-2014, totally 108 out-of-sample forecasts.

We wished to try out the first 108 observations, to make data stationary we log the data and take difference. Then, we used the ADF test to test if the data series is stationary or not. The test statistics was 9.33, therefore the null hypothesis was rejected. The data series is stationary. For the first 108 observations, we tried several different models and compare the AIC value, they are ARIMA (1, 0, 1), ARIMA (1, 1, 1) and six different SARIMA Models, ARIMA (1, 0, 1), ARIMA (1, 1, 1), SARIMA (2, 1, 2) (2, 1, 2) and SARIMA (1, 1, 2) (2, 1, 1), SARIMA (2, 1, 1)(2, 1, 1), SARIMA (2, 0, 1)(2, 1, 0), SARIMA (1, 0, 1)(1, 1, 0) (Lim, 1997), SARIMA (0, 1, 1) (0, 1, 1) and SARIMA (1, 0, 1) (0, 1, 1), respectively. We

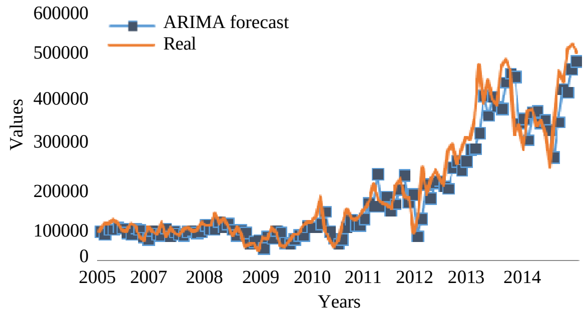

Fig. 2: The results of ARIMA Model forecast with rolling window 108

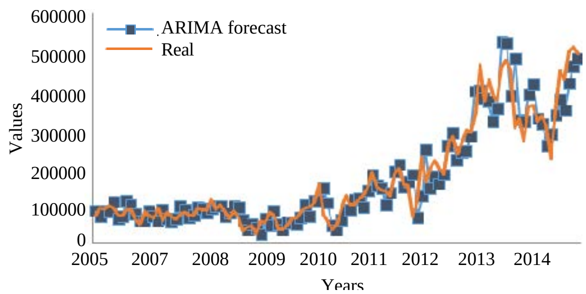

Fig. 3: The results of SARIMA Model Forecast with Rolling Window 80

select the model SARIMA (1, 0, 2) (2, 1, 2) according to each parameter is significant and relatively small AIC value (Table 2).

However, the best in-sample model may not give best forecasts results. As a result, we both use the ARIMA and SARIMA methods to do forecasting, the results show that ARIMA models with different sample horizon has difficult to tackle the uncertainty. The results can be checked in the following Fig. 2 and Table 2.

It can be seen that prior to the structural changes (2009), both methods can predict well. However, when the forecast period goes to the high uncertainty period, the ARIMA methods have less accuracy than SARIMA. The forecast performance can be checked again in Table 3 and Fig. 2 and Fig. 3.

Table 3 shows that the predictions accuracy when there are high uncertainties with different sample sizes. The SARIMA models perform much better. This result is consistent with Smeral and Weber (2000), Cho (2003) and Chu (2009). In the Chinese holiday, we include the three longest official holidays; they are the Spring Festival (January and February), Labor Day (May) and National Day (October). The RMSPE values of SARIMA are half of the ARIMA, no matter structural changes or in the holidays.

Table 2: The performance of ARIMA and SARIMA Models deal with uncertainty measure by RMSPE

Models	Parameters	SE	t-statistics	Sig.
Model 1				
Constant	10.91	0.074	147.432	***
ar1	0.505	0.131	3.854	***
ma1	0.215	0.154	1.396	
Model 2				
Parameter	SE	t-statistics	Sig.	
ar1	0.557	0.136	4.102	***
ma1	-0.568	0.155	-3.671	***
ma2	-0.332	0.12	-2.767	***
sar1	-0.909	0.14	-6.5	***
sar2	-0.725	0.098	-7.397	***
sma1	0.029	0.206	0.14	
sma2	-0.041	0.194	-0.209	
Model 3				
ar1	0.5568	0.1356	4.106	***
ma1	-0.5623	0.1519	-3.702	***
ma2	-0.3367	0.1167	-2.885	***
sar1	-0.8887	0.079	-11.249	***
sar2	-0.731	0.0749	-9.76	***

Log likelihood = -29.67, AIC = 67.34; Log likelihood = -25.86, AIC = 67.72; Log likelihood = -25.88, AIC = 63.76

Table 3: The Performance of ARIMA and SARIMA Models deal with uncertainty measure by RMSPE

Variables	2005-2008	2009-2014	Chinese holiday	Thai coup
ARIMA (108)	0.0527	0.0902	0.0872	0.0847
SARIMA (108)	0.0367	0.0776	0.0745	0.0332
ARIMA (60)	0.0652	0.0944	0.0890	0.0986
SARIMA (80)	0.0469	0.0870	0.0790	0.0317

CONCLUSION

The ARIMA Model is good for tourism demand forecasting when the uncertainty is low. However, when several uncertainty events happened, such as holiday and political turmoil in our study, the model reacts weakly. Based on this, this study compares the forecast performances of ARIMA Models and Seasonal ARIMA Models during the period of high uncertainty. When considering the seasonality in tourist demand, the Seasonal ARIMA Model forecast much better in Chinese tourist arrivals to Thailand. We suggest the policy makers and relevant travel decision section to use SARIMA method to conduct the tourist forecasting. The future work should focus on providing a more reliable method to solve the structural changes and high uncertainty in tourism demand problem.

REFERENCES

Box, G.E.P. and G.M. Jenkins, 1970. Time Series Analysis: Forecasting and Control Holden-Day, San Francisco, CA., USA.

Brida, J.G. and N. Garrido, 2011. Tourism forecasting using SARIMA models in Chilean regions. Int. J. Leisure Tourism Marketing, 2: 176-190.

Carey, G. and L. Rob, 2002. Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. Tourism Manage., 23: 499-510.

Chan, F., C. Lim and M. McAleer, 2005. Modelling multivariate international tourism demand and volatility. Tourism Manage., 26: 459-471.

Chang, C.L., S. Sriboonchitta and A. Wiboonpongse, 2009. Modelling and forecasting tourism from East Asia to Thailand under temporal and spatial aggregation. Math. Comput. Simul., 79: 1730-1744.

Cho, V., 2003. A comparison of three different approaches to tourist arrival forecasting. Tourism Manage., 24: 323-330.

Chu, F.L., 2009. Forecasting tourism demand with ARMA-based methods. Tourism Manage., 30: 740-751.

Connell, J., 2012. Film tourism-Evolution progress and prospects. Tourism Manage., 33: 1007-1029.

Fourie, J. and M. Santana-Gallego, 2011. The impact of mega-sport events on tourist arrivals. Tourism Manage., 32: 1364-1370.

Kulendran, N., 1996. Modelling quarterly tourist flows to Australia using cointegration analysis. Tourism Econ., 2: 203-222.

Lee, C.C. and C.P. Chang, 2008. Tourism development and economic growth: A closer look at panels. Tourism Manage., 29: 180-192.

Lee, C.K. and T. Taylor, 2005. Critical reflections on the economic impact assessment of a mega-event: The case of 2002 FIFA World Cup. Tourism Manage., 26: 595-603.

Lee, C.K., H.J. Song and J.W. Mjelde, 2008. The forecasting of International Expo tourism using quantitative and qualitative techniques. Tourism Manage., 29: 1084-1098.

Li, L. and W. Zhang, 1997. Thailand: The Dynamic Growth of Thai Tourism. In: Tourism and Economic Development in Asia and Australasia, Go, F.M. and C.L. Jenkins (Eds.). Pinter, London, pp: 286-303.

Lim, C. and M. McAleer, 2001. Cointegration analysis of quarterly tourism demand by Hong Kong and Singapore for Australia. Applied Econ., 33: 1599-1619.

Lim, C., 1997. Review of international tourism demand models. Ann. Tourism Res., 24: 835-849.

Rittichainuwat, B. and S. Rattanaphinanchai, 2015. Applying a mixed method of quantitative and qualitative design in explaining the travel motivation of film tourists in visiting a film-shooting destination. Tourism Manage., 46: 136-147.

Shareef, R. and M. McAleer, 2005. Modelling international tourism demand and volatility in small island tourism economies. Int. J. Tourism Res., 7: 313-333.

Smeral, E. and A. Weber, 2000. Forecasting international tourism trends to 2010. Ann. Tourism Res., 27: 982-1006.

Song, H. and G. Li, 2008. Tourism demand modelling and forecasting: A review of recent research. Tourism Manage., 29: 203-220.

Song, H. and S.F. Witt, 2006. Forecasting international tourist flows to Macau. Tourism Manage., 27: 214-224.

Song, H., G. Li, S.F. Witt and G. Athanasopoulos, 2011. Forecasting tourist arrivals using time-varying parameter structural time series models. Int. J. Forecasting, 27: 855-859.

Song, H., P. Romilly and X. Liu, 2000. An empirical study of outbound tourism demand in the UK. Appl. Econ., 32: 611-624.

Song, H., S.F. Witt and G. Li, 2003. Modelling and forecasting the demand for Thai tourism. Tourism Econ., 9: 363-387.

Tseng, F.M. and G.H. Tzeng, 2002. A fuzzy seasonal ARIMA model for forecasting. Fuzzy Sets Syst., 126: 667-676.

Williams, B.M. and L.A. Hoel, 2003. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results. J. Transp. Eng., 129: 664-672.

Witt, S.F. and C.A. Witt, 1992. Modeling and Forecasting Demand in Tourism. Academic Press, Cambridge, Massachusetts, ISBN:9780127607405, Pages: 195.