Predicting tourism demand by A.R.I.M.A. models

Biljana Petrevska

Routledge
Taylor & Francis Group

Predicting tourism demand by A.R.I.M.A. models

Biljana Petrevska

Faculty of Tourism and Business Logistics, Goce Delcev University, Stip, Macedonia (the former Yugoslav Republic of)

ABSTRACT

The paper provides a short-run estimation of international tourism demand focusing on the case of F.Y.R. Macedonia. For this purpose, the Box–Jenkins methodology is applied and several alternative specifications are tested in the modelling of original time series and international tourist arrivals recorded in the period 1956–2013. Upon the outcomes of standard indicators for accuracy testing, the research identifies the model of A.R.I.M.A.(1,1,1) as most suitable for forecasting. According to the research findings, a 13.9% increase in international tourist arrivals is expected by 2018. The forecasted values of the chosen model can assist in mitigating any potential negative impacts, as well as in the preparation of a tourism development plan for the country.

1. Introduction

Tourism is one of the world's largest industries and thus represents a major area of interest, not just because of its size in terms of the enormous number of travellers, passengers, visitors and tourists, or the size of their consumption, but also because of its enormous impact on national economies and people's lives. Tourism has proven to be a surprisingly strong and resilient sector of economic activity and a fundamental contributor to economic recovery by generating billions of dollars in exports and creating millions of jobs.

Yet, this dynamic industry faces many challenges that affect its development. In order to cope with them, the planners and policy-makers apply the process of forecasting as the only way to furnish information, which permits them to reach decisions before the occurrence of certain events. In order to create a comprehensive tourism development plan as a base for formulating tourism policy, reliable estimates of future demand must be obtained. So, the main aim of introducing forecasting processes in tourism is to envisage the success of the destination by ensuring that visitors are hosted in a way that maximises the benefits to stakeholders with minimal negative effects, costs and impacts (Edgell, DelMastro Allen, Smith, & Swanson, 2008; Goeldner & Ritchie, 2006; Mason, 2015; Wilkinson, 1997).

CONTACT Biljana Petrevska ✉ biljana.petrevska@ugd.edu.mk

The main aim of the current research is to identify a model that best describes and forecasts international tourism demand in F.Y.R. Macedonia. To this end, the time series of international tourist arrivals is considered as a significant indicator of tourism activity and tourism development of the country. The data sample covers the period between 1956 and 2013. The research is structured as follows: following the introduction, Section 2 provides a brief overview of the literature addressing the main topic of interest. The description of the methodology applied is presented in Section 3. Section 4 covers the analysis, main results and discussion. The forecasts and accuracy of the selected specifications are given in Section 5, while the conclusions and some future challenges are presented in the final section.

The study contributes to increasing the limited number of research articles in this area by enriching the poorly developed empirical academic work in F.Y.R. Macedonia. Some exceptions are noted in the work of Petrevska (Petrevska, 2012, 2013), who argues for the need to apply forecasting methods in preparing the country for coping with the unexpected challenges surrounding international tourist arrivals.

2. Literature review

Forecasting tourism demand has attracted much interest in academia. A variety of econometric models have been applied. In this regard, according to the comprehensive study of Song and Li (2008), the forecasting methodology is very diverse since researchers employ both time series and econometric approaches in estimating tourism demand. Consequently, Wong, Song, Witt, and Wu (2007) and Song, Witt, Wong, and Wu (2008) advocate for the introduction of combined forecasts directed at obtaining more accurate models. On the other hand, combining forecasts is more beneficial for long-run estimations.

Although the basic indicator describing tourism demand has gradually been modified, tourist arrivals is still the most applicable indicator. It is noticeable that many authors break this variable down in several in-depth variables, such as holiday tourist arrivals, business tourist arrivals, tourist arrivals for visiting friends and relatives (Kulendran & Wong, 2005; Turner & Witt, 2001a, 2001b), tourist arrivals by air (Coshall, 2005; Rosselló, 2001), tourist expenditure in the destination (Li, Song, & Witt, 2004, 2006; Li, Wong, Song, & Witt, 2006), tourist expenditure on particular tourism product categories as meal expenditure (Au & Law, 2002), sightseeing expenditure (Au & Law, 2000) and shopping (Law & Au, 2000). Other tourism demand variables used in the literature include tourism revenues (Akal, 2004), tourism employment (Witt, Song, & Wanhill, 2004) and tourism import and export (Smeral, 2004).

3. Methodology

The research is based on the Box–Jenkins methodology (Box & Jenkins, 1976) as a suitable technique for short-run forecasting. It is an algebraic model that is commonly applied in forecasting and is known as autoregressive integrated moving average (A.R.I.M.A.) models. The Box–Jenkins models are very popular for forecasting and enjoy great success in academic research (Bigovic, 2012; Chu, 2008; Coshall, 2005; Goh & Law, 2002; Huang & Min, 2002; Kulendran & Shan, 2002; Kulendran & Witt, 2001; Law, 2000, 2004; Lim & McAleer, 2002; Preez & Witt, 2003; Qu & Zhang, 1996).

The A.R.I.M.A. models applied in the current research follow the standard expression (1) of Zou and Yang (2004) and Fernandes, Teixeira, Ferreira, and Azevedo (2008):

$$(1 - \Phi_1 B - ... - \Phi_p B^p)(1 - B)^d Y_t = (1 - \theta_1 B - ... - \theta_q B^q)e_t \qquad (1)$$

In expression (1), the A.R.I.M.A. model is (p,d,q), in which p is the order of the autoregressive (A.R.) process, d is the number of differences or integrations and q is the order of the moving average (M.A.) process. A.R.I.M.A. models can be used for short-run estimation based on annual, quarterly, monthly or even weekly, daily or hourly data.

4. Analysis, results and discussion

The main objective of this research is to make a short-run estimation of international tourism demand in F.Y.R. Macedonia in terms of international tourist arrivals by introducing several specifications of A.R.I.M.A. models. The research is based on the available sources of secondary data – in this case from the State Statistical Office – and these are processed by the software E-views version 6.0. The original time series is the number of international tourists per year, which is considered to be significant indicator of tourism development – in this case of F.Y.R. Macedonia. The data observed cover the period from 1956 to 2013 on an annual basis due to the unavailability of monthly data, so the sample consists of 58 observations.

Figure 1 visually evaluates the potential stationarity as a first basic assumption for applying the A.R.I.M.A. modelling. One can observe two pronounced trends in the movement of the time series, suggesting that some of its characteristics imply non-stationarity. Specifically, it is obvious that the series has different average values and variances in various sub-periods of the sample.

However, for precautionary reasons, the above-noted intuitive conclusion is tested by a correlogram of the series, showing the autocorrelation function (A.C.F.) and autocorrelation coefficients (ρ_k) calculated for 24 lags (Table 1). Due to facts that the correlogram starts
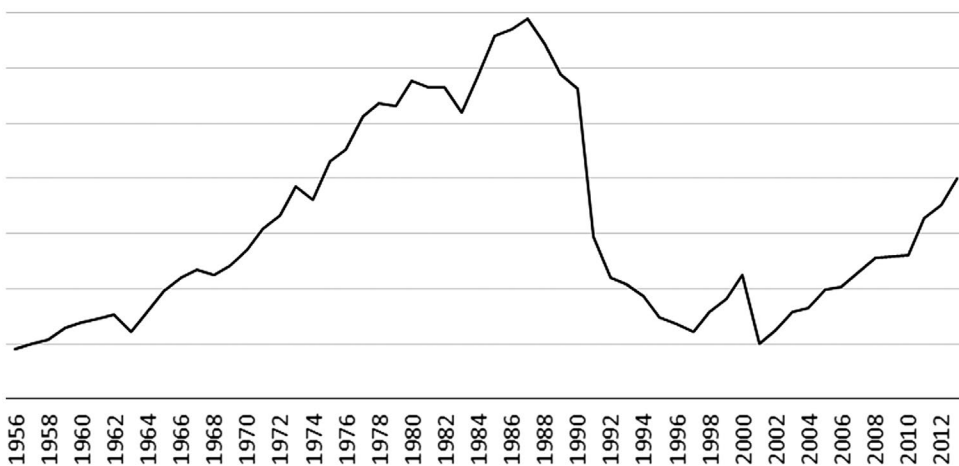

Figure 1. International tourist arrivals in F.Y.R. Macedonia, 1956–2013. Source: State Statistical Office 2014.

Table 1. Correlogram of international tourist arrivals.

Lags (k)	A.C.F. (ρ_k)	L.B. statistic	p-value
1	0.944	54.377	0.000
2	0.869	101.32	0.000
3	0.778	139.62	0.000
4	0.680	169.40	0.000
5	0.575	191.12	0.000
6	0.462	205.43	0.000
7	0.354	214.00	0.000
8	0.252	218.41	0.000
9	0.150	220.00	0.000
10	0.049	220.18	0.000
11	−0.061	220.46	0.000
12	−0.159	222.37	0.000
13	−0.252	227.27	0.000
14	−0.349	236.89	0.000
15	−0.427	251.68	0.000
16	−0.494	271.87	0.000
17	−0.533	295.99	0.000
18	−0.552	322.48	0.000
19	−0.568	351.25	0.000
20	−0.567	380.65	0.000
21	−0.551	409.21	0.000
22	−0.517	435.00	0.000
23	−0.469	456.91	0.000
24	−0.435	476.30	0.000

Source: Own estimates.

with a very high correlation coefficient of 0.94, that the ρ_k slowly decays and that there is a highly expressed autocorrelation in the series even for several lags (ρ_k at lag 6 is still high at almost 0.5), one may conclude that the series is non-stationary.

Furthermore, the research continues by checking the statistical significance of the sample autocorrelation coefficients in order to determine whether they really represent the true ρ_k of the population. As stated in the statistical theory, if dealing with a random process, then the ρ_k are approximately characterised by the normal distribution, with a zero mean and variance of $1/n$, where n is the sample size (Gujarati, 1995, p. 717). Hence, the standard error of the ρ_k can be calculated as: $\sqrt{1/58} = 0.131$. According to the tabular values for the normal distribution, the 95% confidence interval for the ρ_k is $\pm 1.96 \times 0.131 = \pm 0.257$.

If the calculated ρ_k lies within the confidence interval, this means that the null hypothesis that the true ρ_k of the population is zero (H_0: $\rho_k = 0$) cannot be rejected. From Table 1, it can be seen that the first seven ρ_k are statistically significant (i.e., different from zero), the coefficients from lag 8 to lag 13 lie within the confidence interval, and then these coefficients are again different from zero. The large number of statistically significant coefficients once again confirms that the series is non-stationary.

However, given the problems with the individual testing of the significance of ρ_k, the joint hypothesis that all ρ_k are equal to zero is tested by employing the Ljung–Box (L.B.) statistic. This tests the null hypothesis that there is no autocorrelation for all of the coefficients at certain numbers of time lags. Then, if the null hypothesis is true, the L.B. statistic asymptotically follows the χ^2 distribution, with degrees of freedom equal to the number of autocorrelation coefficients. The values of the L.B. statistic are presented in Table 1, showing that for all time lags, the L.B. statistic by far exceeds the critical values. For instance, even at lag 24, the L.B. statistic is statistically highly significant (476.3). In that respect, this test

shows that the null hypothesis can be rejected, which is a proof that the analysed time series is non-stationary.

Nevertheless, it is known that the L.B. statistic has low power because the significant coefficients can be neutralised by the insignificant ones. Hence, the evidence obtained by the L.B. statistic is additionally tested by employing two unit root tests: the augmented Dickey–Fuller (A.D.F.) and the Phillips–Perron (P.P.) test.

In the first row of Table 2, the values of the A.D.F. test are shown in its three variants, and in all cases, the null hypothesis for the presence of a unit root cannot be rejected. Consequently, this test suggests that the series is non-stationary. Yet, it is known that the A.D.F. test has low power and, hence, the results are additionally checked with the P.P. test. As shown in the second row of Table 2, all of the variants of the P.P. test show that the null hypothesis of a unit root cannot be rejected. Hence, this test, too, suggests that the series is non-stationary. As mentioned previously, if the time series is non-stationary, then the Box–Jenkins methodology cannot be applied. This means that it is necessary to transform the series in order to make it stationary, which is done by differencing the original series (Table 3).

When the series is differenced, one cannot observe a regular movement of the autocorrelation coefficients, which begin with low values, decreasing quickly to zero and then, up to lag 12, moving in a wave style (i.e., increasing and decreasing). In order to check the stationarity of the differenced series, the ρ_k are individually tested with the confidence interval, which in this case is ±0.257. It was then shown that the null hypothesis that the true ρ_k of the population are equal to zero cannot be rejected. Specifically, the value of L.B. statistic with 24 degrees of freedom is 34.919, which is not sufficient to reject the null hypothesis. The above results show that by differencing the original time series, stationarity is obtained. Yet, once again, in order to verify the results, the A.D.F. test and the P.P. test are used. From Table 3, it can be concluded that the values of the statistics are highly significant, so once again, it can be concluded that the first-differenced series is stationary.

After performing the additional tests, it can be concluded that the Box–Jenkins methodology can be applied. The first step is to identify the appropriate model that will explain the time series movement. Here, crucial instruments are the sample A.C.F. and the partial A.C.F. (P.A.C.F.). The detailed analysis of both functions (presented in Table 4) reveals that both A.C.F. and P.A.C.F. decay gradually up to lag 3, and then there is no regularity in the movement of the autocorrelation coefficients, from which the model could be identified. What the correlogram suggests is that we have a mixed process (i.e., a combination of A.R. and M.A. processes).

Given the unclear character of the time series, two alternative specifications were used in order to model the original series – A.R.I.M.A.(1,1,1) and A.R.I.M.A.(2,1,2) – which represent the original time series in an adequate manner. The models' estimates along with the accompanying diagnostics are shown in Tables 5 and 6, respectively. As can be seen, the A.R.I.M.A.(2,1,2) model has a slightly lower adjusted coefficient of determination

Table 2. Stationarity tests.

Test	Constant	Constant + trend	None
A.D.F.	−1.610721	−1.589757	−0.03342
	(0.4705)	(0.7846)	(0.6774)
P.P.	−1.674346	−1.636771	−0.346957
	(0.4387)	(0.7658)	(0.5557)

Source: Own estimates.

Table 3. Correlogram of international tourist arrivals (first differences).

Lags (k)	A.C.F. (ρ_k)	P.C.F. (ρ_{kk})	L.B. statistic	p-value
1	0.259	0.259	4.0221	0.045
2	0.218	0.162	6.9340	0.031
3	0.154	0.071	8.4024	0.038
4	0.089	0.008	8.9030	0.064
5	0.107	0.054	9.6379	0.086
6	−0.036	−0.103	9.7242	0.137
7	−0.098	−0.116	10.374	0.168
8	0.046	0.108	10.517	0.231
9	0.062	0.089	10.786	0.291
10	0.153	0.139	12.452	0.256
11	−0.109	−0.214	13.322	0.273
12	−0.064	−0.074	13.629	0.325
13	0.005	0.025	13.631	0.400
14	−0.156	−0.159	15.532	0.343
15	−0.083	0.005	16.081	0.377
16	−0.270	−0.170	22.081	0.141
17	−0.180	−0.029	24.807	0.099
18	−0.069	−0.008	25.218	0.119
19	−0.128	−0.048	26.673	0.113
20	−0.155	−0.105	28.868	0.090
21	−0.146	−0.035	30.868	0.076
22	−0.159	−0.104	33.293	0.058
23	−0.016	0.025	33.317	0.076
24	−0.125	−0.012	34.919	0.070

Source: Own estimates.

Table 4. Stationarity tests (first differences).

Test	Constant	Constant + trend	None
A.D.F.	−5.582111	−5.545318	−5.582888
	(0.0000)	(0.0001)	(0.0000)
P.P.	−5.684429	−5.650428	−5.683326
	(0.0000)	(0.0001)	(0.0000)

Source: Own estimates.

Table 5. A.R.I.M.A.(1,1,1).

Dependent variable: international tourist arrivals (first differences)				
Method: least squares				
Sample: 1956–2013 (58 observations)				
Variable	Coefficient	Standard error	t-statistic	Probability
A.R.(1)	0.743985	0.251603	2.956979	0.0046
M.A.(1)	−0.515492	0.318801	−1.616968	0.1117
R^2	0.092518	Akaike information criterion		24.52577
Adjusted R^2	0.75712	Schwarz criterion		24.59810
Standard error of regression	50,332.97	Durbin–Watson statistics		2.026473
Inverted A.R. roots 0.74				
Inverted M.A. roots 0.52				

Source: Own estimates.

comparing to the competing model. However, all of the parameters in this model are statistically significant in contrast to A.R.I.M.A.(1,1,1), in which the M.A. term is not significant even at 10%. In both models, the inverted A.R. roots have modulus values lower than one (i.e., they lie within the unit circle). This is true for the M.A. terms, too, whose roots are

Table 6. A.R.I.M.A.(2,1,2).

Dependent variable: international tourist arrivals (first differences)				
Method: least squares				
Sample: 1956–2013 (55 observations)				
Variable	Coefficient	Standard error	t-statistic	Probability
A.R.(1)	1.303123	0.226527	5.752623	0.0000
A.R.(2)	−0.726533	0.199709	−3.637954	0.0006
M.A.(1)	−1.196601	0.185019	−6.467451	0.0000
M.A.(2)	0.798118	0.141648	5.634502	0.0000
R^2	0.109851	Akaike information criterion		24.59853
Adjusted R^2	0.057490	Schwarz criterion		24.74451
Standard error of regression	51,295.14	Durbin–Watson statistics		1.814606
Inverted AR roots	0.65−0.55i	0.65+0.55i		
Inverted MA roots	0.60+0.66i	0.60−0.66i		

Source: Own estimates.

imaginary, but lie within the unit circle with values much lower than one. These results confirm that the two models are covariance stationary and invertible, thus making them appropriate for forecasting. Finally, the values of the Akaike and the Schwarz information criteria are very similar, suggesting that both models perform well (i.e., we can barely discriminate between the two specifications based on the information criteria).

5. Forecast results and evaluation of models

This section presents the results of the estimation of the A.R.I.M.A. models of international tourism demand in F.Y.R. Macedonia and the adequacy testing of the proposed models. In these regards, the figures presented in the appendices show that the two A.R.I.M.A. specifications produce fitted values that very closely track the actual values of the time series. As for their out-of-sample forecast performance, Table 7 reveals that there are noticeable similarities in the forecasted values across the two A.R.I.M.A. specifications. For instance, according to the A.R.I.M.A.(1,1,1) model, it can be expected that by 2018, international tourist arrivals in F.Y.R. Macedonia will reach 455,107 foreign tourists, thus representing an increase of 13.9% compared to 2013. On the other hand, the A.R.I.M.A.(2,1,2) model predicts 439,247 foreigners visiting F.Y.R. Macedonia in 2018, representing an increase of 9.9% in the number of foreign tourists as compared to 2013. Hence, both models imply an upward trend in international tourist arrivals in Macedonia.

Due to the fact that the primary purpose of building a forecasting model is to clearly discern the future of a phenomenon, the most important criterion is how accurately the model does this. This means attempting to disentangle how closely the estimations provided by the model conform to the actual events being forecasted. In this respect, we support the accuracy of the suggested model by within-sample forecasts. To this end, some of the standard indicators are employed, such as root mean squared error (R.M.S.E.), mean absolute error (M.A.E.), absolute percentage error (A.P.E.), mean absolute percentage error (M.A.P.E.) and the Theil inequality coefficient (T.I.C.).

Table 8 presents the values of the evaluation of the suggested models. Generally, both models present good accuracy according to the M.A.P.E. classification for analysing the quality of the values forecast since the results lie within the interval 10–20% (Lewis, 1982).

Table 7. Estimating international tourism demand in F.Y.R. Macedonia, 2014–2018.

Model	2014	2015	2016	2017	2018
A.R.I.M.A.(1,1,1)	418,060	431,734	441,907	449,476	455,107
A.R.I.M.A.(2,1,2)	420,058	438,887	448,618	447,619	439,247

Source: Own estimates.

Table 8. Evaluation of models.

Model	R.M.S.E.	M.A.E.	A.P.E.	M.A.P.E.	T.I.C.
A.R.I.M.A.(1,1,1)	49,426	30,364	732.424	12.628	0.068097
A.R.I.M.A.(2,1,2)	49,395	34,339	841.348	14.506	0.067685

Source: Own estimates.

Based on the analysis of the error values (Table 8), it may be concluded that the models produce good and satisfactorily accurate forecasts for the observed period (Appendix A, Table A.1 and Table A.2; Appendix B, Table B.1 and Table B.2). Yet, the majority of the estimated values of the forecasted errors (M.A.E., A.P.E. and M.A.P.E.) are in favour of the A.R.I.M.A.(1,1,1) model, pointing to the conclusion that this model is slightly better compared to the A.R.I.M.A.(2,1,2) model.

When evaluating the forecast performance of econometric models, the issue of model stability arises naturally. As revealed in Figure 1, international tourist arrivals in F.Y.R. Macedonia show an upward trend until the end of the 1980s, followed by a steep decline during the 1990s. Then, since the 2000s, one can observe another reversal in the dynamics of the series. Naturally, this raises the issue of the presence of potential structural breaks during the sample period. Though some preliminary conclusions can be obtained on the basis of mere visual inspection of Figure 1, we provide the results of several formal stability tests, such as the Chow breakpoint test, the Chow forecast test and the Quandt–Andrews test for unknown breakpoint dates. Though we have run the stability tests for the two A.R.I.M.A. specifications, for the sake of preserving space, we report and discuss the results for the A.R.I.M.A.(1,1,1) model. Note that the results are very similar for the A.R.I.M.A.(2,1,2) model and they are available upon request.

Appendix C presents the two variants of the Chow test for possible structural breaks dated in 1986, 1991 and 2002, as well as the Quandt–Andrews test for unknown breakpoint dates. The main findings from these tests are as follows: regarding the possible structural breaks occurring in 1986 and 1991, both the Chow breakpoint test and the Chow forecast test reject the null hypothesis at the 1% significance level. Only the first Chow test points to a possible structural break in 2002 at the 10% significance level. All of the test statistics associated with the Quandt–Andrews test confirm the presence of multiple structural breaks within the sample period. These results suggest that some caution is needed when applying the above A.R.I.M.A. models for forecasting. However, this qualification can be partially counterbalanced by the fact that the selected models consist of few A.R. and M.A. terms, which could be appropriate for tracking the dynamics of many economic time series characterised by considerable degrees of persistence.

The Box–Jenkins methodology is commonly applied on a regular basis when dealing with tourism demand estimations. In this case, the evaluated results provide good accuracy, highlighting that A.R.I.M.A.(1,1,1) is more suitable for modelling and forecasting due to

slightly better and satisfactory statistical and adjustment qualities. Furthermore, it should be pointed out that the anticipated values must be interpreted with caution, because the model does not indicate the reasons for these effects on the estimated results. This is very important, as these indicators have significant influence on identifying and implementing measures and activities in order to create appropriate tourism policy for the country.

6. Conclusion and future work

Tourism in F.Y.R. Macedonia plays a significant role in strengthening the economy, particularly in terms of job creation. International tourist arrivals have shown persistent increases during the past few years, due to some government support measures (subsidies, larger budgets for tourism promotion, raising awareness of the positive impacts of tourism, aggressive promotion in international regional markets, etc.). Hence, anticipating international tourism demand is necessary for creating favourable conditions and images of the country.

The specification of the Box–Jenkins methodology, the A.R.I.M.A.(1,1,1) model, is constructed and proposed for analysing and short-run forecasting of international tourism demand in F.Y.R. Macedonia based on annual data. This research forecasts that the upward trend in the international tourism demand will continue in the near future, pointing to positive impulses and expectations of a continuous increase by 2018.

Additionally, the paper explains that the suggested model does not provide 'the solution', but only assists in finding it. Even though the model results are essential elements in the preparation of well-coordinated policies, they cannot do the job all by themselves. The research outcomes may be presented only as a framework, while the rest needs to be achieved with a lot of common sense and knowledge of the details. So, the projected values cannot explain the factors underlying these trends, but on the other hand may serve as a solid base for mitigating any potential negative impacts and preparing the tourism development plan in F.Y.R. Macedonia.

Although the accuracy of the proposed A.R.I.M.A. model can be regarded as good, valid and satisfactory, and despite the fact that the projected values are classified as reliable forecasts, the model is still not highly accurate, probably due to the presence of several structural breaks during the sample period. This means that the suggested model has some difficulties in producing forecasts with minimal errors, implying that different models could be tested. Therefore, future aspects of this research may address the application and modelling of other models that may obtain forecasts that are closer to the actual time series and with greater accuracy, such as hybrid models, artificial neural network models, conditional volatility models, Holt–Winter additive models, Holt–Winter multiplicative models, vector autoregressive models, error correction models; Gaussian mixture model (G.M.M.), etc.

Even when an ideal forecasting model can be identified, it can only serve as an approximation for complex tourist behaviours, as it is possible that tourists' decisions could reflect changes in preferences, motivations or economic shocks. Hence, the planner should always be prepared to make revisions to the previously identified and defined model, adapting it to any newly created changes.

Disclosure statement

No potential conflict of interest was reported by the author.

References

Akal, M. (2004). Forecasting Turkey's tourism revenues by ARMAX model. *Tourism Management, 25*, 565–580.

Au, N., & Law, R. (2000). The application of rough sets to sightseeing expenditures. *Journal of Travel Research, 39*, 70–77.

Au, N., & Law, R. (2002). Categorical classification of tourism dining. *Annals of Tourism Research, 29*, 819–833.

Bigovic, M. (2012). Demand forecasting within Montenegrin tourism using Box-Jenkins methodology for seasonal ARIMA models. *Tourism and Hospitality Management, 18*(1), 1–18.

Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control*. San Francisco: Holden-Day Inc.

Chu, F. L. (2008). Forecasting tourism demand with ARMA-based methods. *Tourism Management, 10*, 1–12.

Coshall, J. T. (2005). A selection strategy for modelling UK tourism flows by air to European destinations. *Tourism Economics, 11*, 141–158.

Edgell, D. L., DelMastro Allen, M., Smith, G., & Swanson, J. R. (2008). *Tourism policy and planning: Yesterday, today and tomorrow*. Amsterdam: Elsevier.

Fernandes, P., Teixeira, J., Ferreira, J. M., & Azevedo, S. G. (2008). Modelling tourism demand: A comparative study between artificial neural networks and the Box-Jenkins methodology. *Romanian Journal of Economic Forecasting, 3*, 30–50.

Goeldner, C. R., & Ritchie, J. R. B. (2006). *Tourism: Principles, practicies, philosophies*. New Jersey: John Wiley & Sons.

Goh, C., & Law, R. (2002). Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. *Tourism Management, 23*, 499–510.

Gujarati, D. N. (1995). *Basic econometrics* (2nd ed.), New Delhi: McGraw-Hill.

Huang, J. H., & Min, J. C. H. (2002). Earthquake devastation and recovery in tourism: the Taiwan case. *Tourism Management, 23*, 145–154.

Kulendran, N., & Shan, J. (2002). Forecasting China's monthly inbound travel demand. *Journal of Travel & Tourism Marketing, 13*, 5–19.

Kulendran, N., & Witt, S. F. (2001). Cointegration versus least squares regression. *Annals of Tourism Research, 28*, 291–311.

Kulendran, N., & Wong, K. K. F. (2005). Modeling seasonality in tourism forecasting. *Journal of Travel Research, 44*, 163–170.

Law, R. (2000). Demand for hotel spending by visitors to Hong Kong: A study of various forecasting techniques. *Journal of Hospitality and Leisure Marketing, 6*, 17–29.

Law, R. (2004). Initially testing an improved extrapolative hotel room occupancy rate forecasting technique. *Journal of Travel & Tourism Marketing, 16*, 71–77.

Law, R., & Au, N. (2000). Relationship modeling in tourism shopping: a decision rules induction approach. *Tourism Management, 21*, 241–249.

Lewis, C. D. (1982). *Industrial and business forecasting method*. London: Butterworth Scientific.

Li, G., Song, H., & Witt, S. F. (2004). Modeling tourism demand: A dynamic linear AIDS approach. *Journal of Travel Research, 43*, 141–150.

Li, G., Wong, K. F., Song, H., & Witt, S. F. (2006). Tourism demand forecasting: A time varying parameter error correction model. *Journal of Travel Research, 45*, 175–185.

Li, G., Song, H., & Witt, S. F. (2006). Time varying parameter and fixed parameter linear AIDS: An application to tourism demand forecasting. *International Journal of Forecasting, 22*, 57–71.

Lim, C., & McAleer, M. (2002). Time series forecasts of international travel demand for Australia. *Tourism Management, 23*, 389–396.

Mason, P. (2015). *Tourism impacts, planning and management*. London: Routledge.

Petrevska, B. (2012). Forecasting international tourism demand: the Evidence of Macedonia. *UTMS Journal of Economics, 3*, 45–55.

Petrevska, B. (2013). International tourism demand in Macedonia: current status and estimation. *Yearbook of the Faculty of Economics, University in East Sarajevo, 7*, 51–60.

Preez, J. & Witt, S. F. (2003). Univariate versus multivariate time series forecasting: an application to international tourism demand. *International Journal of Forecasting, 19*, 435–451.

Qu, H., & Zhang, H. Q. (1996). Projecting international tourist arrivals in East Asia and the Pacific to the year 2005. *Journal of Travel Research, 35*, 27–34.

Rosselló, J. (2001). Forecasting turning points in international visitor arrivals in the Balearic Islands. *Tourism Economics, 7*, 365–380.

Smeral, E. (2004). Long-term forecasts for international tourism. *Tourism Economics, 10*, 145–166.

Song, H., & Li, G. (2008). Tourism demand modelling and forecasting – A review of recent research. *Tourism Management, 29*, 203–220.

Song, H., Witt, S. F., Wong, K. K. F., & Wu, D. C. (2008). An empirical study of forecast combination in tourism. *Journal of Hospitality & Tourism Research, 33*, 3–29.

State Statistical Office. 2014. Statistical Yearbook of Macedonia, Skopje.

Turner, L. W., & Witt, S. F. (2001a). Factors influencing demand for international tourism: Tourism demand analysis using structural equation modelling, revisited. *Tourism Economics, 7*, 21–38.

Turner, L. W., & Witt, S. F. (2001b). Forecasting tourism using univariate and multivariate structural time series models. *Tourism Economics, 7*, 135–147.

Wilkinson, P. (1997). *Tourism Planning on Islands*. New York, NY: Cognizant Communications.

Witt, S. F., Song, H., & Wanhill, S. P. (2004). Forecasting tourism-generated employment: The case of Denmark. *Tourism Economics, 10*, 167–176.

Wong, K., Song, H., Witt, S. F., & Wu, D. C. (2007). Tourism forecasting: To combine or not to combine? *Tourism Management, 28*, 1068–1078.

Zou, H. & Yang, Y. (2004). Combining time series models for forecasting. *International Journal of Forecasting, 20*, 69–84.
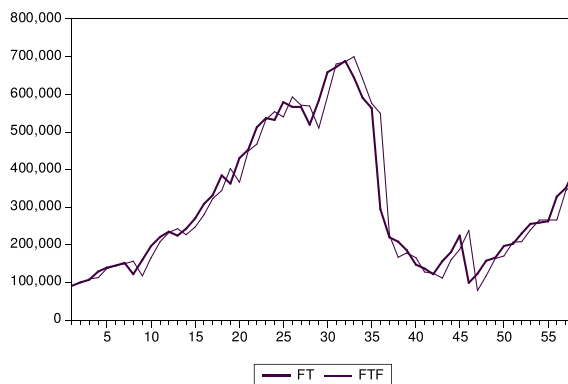
Appendix A

Table A.1. Fitted values of international tourist arrivals by A.R.I.M.A.(1,1,1).
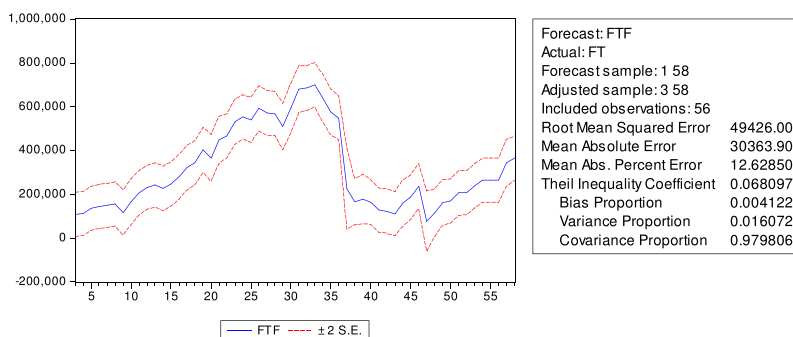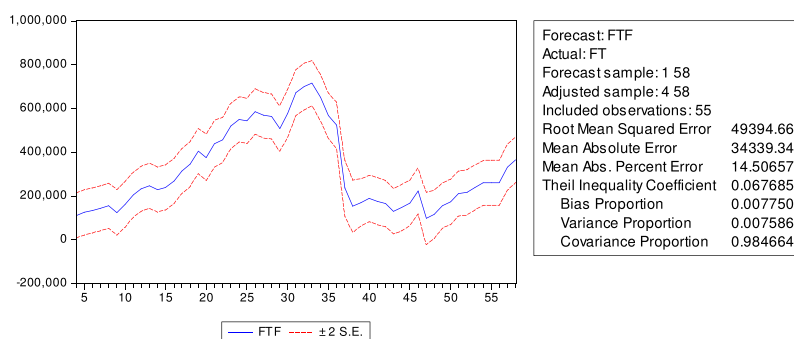

Source: Author's calculations.

Table A.2. Fitted and actual values of international tourist arrivals by A.R.I.M.A.(1,1,1).

Source: Author's calculations.

Appendix B

Table B.1. Fitted values of international tourist arrivals by A.R.I.M.A.(2,1,2).

Forecast: FTF	
Actual: FT	
Forecast sample: 1 58	
Adjusted sample: 4 58	
Included observations: 55	
Root Mean Squared Error	49394.66
Mean Absolute Error	34339.34
Mean Abs. Percent Error	14.50657
Theil Inequality Coefficient	0.067685
Bias Proportion	0.007750
Variance Proportion	0.007586
Covariance Proportion	0.984664

Source: Author's calculations.

Table B.2. Fitted and actual values of international tourist arrivals by A.R.I.M.A.(2,1,2).
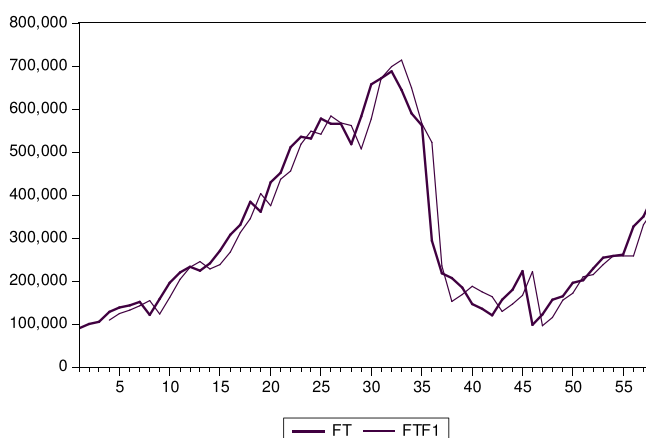

Source: Author's calculations.

Appendix C

Table C.1. Stability tests.

	F-statistic	Likelihood ratio
Chow breakpoint test (1986)	4.430545 (0.0167)	8.811618 (0.0122)
Chow forecast test (1986)	F-statistic 11.44361 (0.0000)	Likelihood ratio 148.8749 (0.0000)
Chow breakpoint test (1991)	F-statistic 6.152948 (0.0040)	Likelihood ratio 11.89483 (0.0026)
Chow forecast test (1991)	F-statistic 3.719238 (0.0004)	Likelihood ratio 77.28688 (0.0000)
Chow breakpoint test (2002)	F-statistic 2.863734 (0.0661)	Likelihood ratio 5.851405 (0.0536)
Chow forecast test (2002)	F-statistic 0.899731 (0.5603)	Likelihood ratio 14.05471 (0.3700)
Quandt–Andrews unknown breakpoint test	Maximum Likelihood ratio (L.R.) F-statistic 24.00269 (0.0000)	Maximum Wald F-statistic 70.43952 (0.0000)

Note: p-values are presented in parentheses.
Source: Author's calculations.