ICOT2015

LONDON, UK, 24-27 JUNE


http://iatour.net/icot2015
facebook.com/groups/355746984498748

PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON TOURISM (ICOT2015)

From Tourism Policy into Practice: *Issues and Challenges in Engaging Policy Makers and End Users*

London, UK, 24-27 JUNE

ISSN 2241-9314
ISBN 978-618-81503-0-0

EDITOR: Konstantinos Andriotis

Organized by: IATOUR, Middlesex University London

CONFERENCE PROCEEDINGS OF THE

INTERNATIONAL CONFERENCE ON TOURISM (ICOT 2015)

From Tourism Policy into Practice: Issues and Challenges in Engaging Policy Makers and End Users

London 24-27June 2015

EDITOR

Konstantinos Andriotis
Middlesex University London, U.K.

k.andriotis@mdx.ac.uk

A HEURISTIC TECHNIQUE FOR SELECTING SARIMA MODELS FOR FORECASTING TOURISM TIME SERIES

IOANNIS A. NIKAS AND ALKIVIADIS PANAGOPOULOS
Technical Educational Institute of Western Greece, Greece

The problem of forecasting tourism demand is a well-timed issue in the tourism sector. One of the methods widely used is the well-known Box-Jenkins method. Another approach is the adoption of a "try-and-error" methodology for the analytical testing of all possible SARIMA models until a proper one is found. In this work, the problem of selecting automatically SARIMA models for tourist time-series is considered. In particular, a heuristic algorithm is proposed to test, exhaustively, all possible models until a SARIMA one is found. The stated algorithm is accelerated by proposing efficient termination criteria. These criteria are variations of classic metrics for the evaluation of forecasting models. The selected model is re-evaluated for the goodness of fit using some well-known evaluators. Finally, the performance of the algorithm is tested on a set of time-series data describing the tourism demand from the region of Western Greece.

1. Introduction

If someone tries to give an interpretation in the word “traveler” he could say that "traveler is a person that consumes the tourism product". This interpretation emerges from the tendency of modern people to consume such products, like e.g., airplane tickets, food and beverage services, overnight stays and museum visits, for traveling along the world for business, educational and entertainment reasons.

For all the specialists in the field of tourism beyond their major activities, there exists a crucial question that they should answer: what shall be the next place where these travelers will consume the tourism product? It is evident from this simple question that prediction plays a major role in tourist planning and policy making. So, the possibility of analyzing the current and past tourist trends, and making some accurate and reliable predictions for the future of tourism product, would be a great asset in the hands of specialist in the field of tourism.

Generally, for the problem of forecasting tourism demand in the form of time-series different methods and techniques have been proposed covering a wide range of different countries and locations, as well as different time intervals. The most widely used models (especially using monthly data) are univariate or time-series models (Gunter, and Önder, 2015). The most widely used technique in this framework is the (Seasonal) Autoregressive (Integrated) Moving Average (SARIMA) models (Box and Jenkins, 1976) because it can handle flexibly different types of data, it can produce reliable predictions when the appropriate model is chosen and it can be easily found and applied in many computational and statistical packages. An exhaustive review on forecasting time series can be found in (Song, and Li, 2008).

In this work we consider the problem of automatically selecting SARIMA models for time-series describing tourism demand. A naive but time-consuming implementation for this idea is the adoption of an exhaustive "try-and-error" methodology which tests analytically all the possible SARIMA models until a proper forecasting model is found.

Recently, Hyndman and Khandakar (Hyndman, and Khandakar, 2008) proposed an automatic way for identifying a SARIMA model for a given time-series, based on a series of tests and rules.

Under the above "try-and-error" framework, an algorithm is proposed for identifying SARIMA models. In particular, a heuristic algorithm is proposed to test exhaustively all feasible SARIMA models, employing time-series of tourism data. The proposed algorithm is accelerated by proposing some termination criteria. These criteria are based on some variations of classic metrics for the evaluation of forecasting models. The first model satisfying all the proposed termination criteria is selected and then is re-evaluated using some well-known evaluators for the goodness of fit of time-series models.

The performance of the algorithm is tested on a data set describing the tourism demand from the prefectures of the region of Western Greece.

The rest of the paper is structured as follows: In the next two sections the proposed framework and algorithm are described. In the Section 4 a short description of the region of Western Greece is given and then the numerical results of the proposed algorithm are presented. Finally, in Section 5 some conclusions and issues for discussions are given.

2. The Proposed Framework

As described in the previous section, in this work an exhaustive “try-and-error” algorithm will be presented for identifying SARIMA models using tourism data. The adopted framework can be described by the following Algorithm 1 (Table 1).

The space model **S** is defined as the set of all feasible SARIMA models. Since the model space has already been defined, the parameters of all SARIMA models are estimated. The model that has the lowest value of an Information Criterion (**IC**) is considered as a candidate model. The Information Criteria are methods for model selection and in this work the Akaike's Information Criterion (AIC), the Corrected Akaike's Information Criterion (AICc) and the Schwarz Bayesian Information Criterion (BIC) are used (Hyndman, and Athanasopoulos, 2015). Although in the most cases the model with minimum value of AIC, AICc or BIC is selected as the best model, in this work we prefer to check further the goodness of fit by using some well-known tests, as they will be described below.

Table 1: Algorithm 1 - The proposed algorithmic framework

Step 1.	Define the model space S and estimate all the corresponding SARIMA models.
Step 2.	Select the model with the lowest IC value and name this model as candidate model.
Step 3.	Check the goodness of fit of the candidate model.
Step 4.	If the candidate model is not satisfying the goodness of fit criteria then remove it from S and go to Step 2.
Step 5.	Evaluate the performance of the accepted model.
Step 6.	Forecast using the accepted model.

The selected model will be a candidate model. If the candidate model fails to pass the goodness of fit tests then it is discarded from the model space **S**. The next model is chosen based on the same criterion and, if it is necessary, all the models will be examined until the goodness of fit criteria are satisfied. Finally, the performance of the accepted model is evaluated and a forecast for the given data is estimated.

The main advantage of the above Algorithm 1 is that, eventually, the best model will be delivered. However, its complexity indicates that with a moderate space model **S**, more than 2000 models should be examined. So the challenge in the above framework is how to keep as complete as possible the size of **S** and at the same time to shorten the total number of examined models.

3. The Proposed Algorithm

In this section a step-by-step algorithm will be presented to optimize the proposed framework of the previous section. For this reason a time-series $\boldsymbol{Y = \{Y_t : t=0,1,2,...,N\}}$ of $\boldsymbol{N}$ observations is considered.

3.1 Data preparation

The collected data in the form of a time-series are distinguished into two sets: the **training set**, the set that contains the initial $\boldsymbol{k}$ of $\boldsymbol{N}$ observations, where $\boldsymbol{k < N}$, and the **test set**, the set that contains all the observations of the given time-series. The excluded observations of the training set will be used for the evaluation of the selected model. The forecast will be done using the test set.

3.2 Stationarity

The usability of SARIMA models depends on the stationarity of the given time-series. A common way to check the non-stationarity of a time-series is to plot the series and the corresponding autocorrelation function, and then, visually, someone to examine the graph for a visible trend or clear variable changes over time. Since, the purpose of this work is to identify a model in an automatic way, a series of tests are adopted to check the non-stationarity of the given time-series instead of a visual examination. Therefore, the number of needed normal and seasonal differences is estimated using some well-known unit roots tests (the Kwiatkowski-Phillips-

Schmidt-Shin (kpss) test, the Augmented Dickey-Fuller (adf) test, and the Phillips-Perron (pp) test, while the seasonal unit root tests are the Canova-Hansen (ch) test, and the Osborn-Chui-Smith-Birchenhall (ocsb) test) (Inside-R, 2015).

It is clear that these tests can reduce significantly the size of model space.

3.3 A pre-processing step

The Step 2 of the Algorithm 1 says that the model with the lowest **IC** value is chosen for further consideration. However, this step requires the estimation of all models of the space model **S**. So, it would be very convenient if we could define a threshold **L** for **IC** values, such that the models with **IC** value less than this threshold **L** can be considered as candidate models.

For this reason, a subspace $\mathbf{S_p}$ of **S** is defined with all the arguments of a SARIMA model to lie in {0, 1}, and then the Algorithm 1 is applied on $\mathbf{S_p}$. The stationarity arguments of SARIMA model are estimated using the test of the previous paragraph.

From the application of this algorithm, the lowest values of AIC, AICc and BIC information criteria are collected. These values can be used to define the threshold **L**.

It is noted that for this part of the algorithm, the training set is used.

3.4 Termination criteria

The Algorithm 1 should be accelerated, especially for the case of Step 2 of the algorithm, and for this reason two termination criteria are proposed. The first one says that if during the search process a model **M** reveals

$$M_{IC} < L \tag{3.1}$$

then this model should be processed in the next Step 3 of Algorithm 1. However, since the threshold **L** is a user-defined parameter the criterion is turned into a more relaxed one. Therefore, for the enhancement of the current model selection, a second, complementary, criterion is proposed which says that if

$$M_{p_{index}} < \varepsilon \tag{3.2}$$

then this model should be processed in the next Step 3 of Algorithm 1, where $\boldsymbol{\varepsilon}$ is a small, user-defined value and $\boldsymbol{p_{index}}$ is a performance index that, in general, could be any of the metrics used to measure the performance of a model, e.g. the mean absolute error (MAE) or the mean absolute percentage error (MAPE).

Actually, the above criteria (3.1) and (3.2) can be interpreted as a single termination criterion and may be combined. Therefore, if

$$M_{IC} < L \quad \text{and} \quad M_{p_{index}} < \varepsilon$$

then the current model is considered as a candidate model and should be processed in the next step of Algorithm 1.

3.5 Normality of residuals

According to the Step 3 of the Algorithm 1, if the candidate model satisfies the tests of goodness of fit, this model should be accepted as the proposed model for the given data. In this work three tests were utilized and the satisfaction of, at least, one of them is required. These tests are checking the normality of the residuals.

The used normality tests are the Kolmogorov-Smirnov (ks) test, the Anderson-Darling (ad) test, and the Shapiro-Wilk (sw) test (R Core Team, 2015; Gross and Ligges, 2015).

3.6 Prediction and forecasting accuracy

After all the above steps a SARIMA model should be delivered and its performance must be measured. There exist a series of metrics that can be employed, however, in this work the prediction performance will be evaluated using the mean absolute percentage error (MAPE):

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{\tilde{Y}_t - Y_t}{Y_t}\right|$$

and the root mean square error (RMSE):

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\tilde{Y}_t - Y_t\right)^2}$$

where $\boldsymbol{Y_t}$ represents the actual value of the time-series and the corresponding tilde $\boldsymbol{Y_t}$ the predicted value. In addition, the performance is, also, measured using the mean error (ME):

$$ME = \sum_{t=1}^{N} \left(\tilde{Y}_t - Y_t \right).$$

MAPE and RMSE are used to measure the correctness of the prediction in terms of levels and the deviation between the actual and predicted values, while ME is giving us some evidence about the over-fitting or under-fitting of the proposed model. For all three metrics the smaller the values, the closer the predicted values are to the actual values. Finally, the delivered model is used to produce a prediction, usually, over a 12 month horizon.

Summarizing all the above notes, heuristics and observations we conclude to the following Algorithm 2 (Table 2).

Table 2: Algorithm 2 - The proposed algorithm

Step 1	Collect the data and distinguished them to training and test sets.
Step 2	Define the model space S.
Step 3	If it is necessary make the time-series stationary using the kpss, adf and pp tests for normal stationarity and ch and ocsb tests for seasonal stationarity.
Step 4	Apply Algorithm 1 to a sample model space S_p to estimate the heuristic parameter L. L will depend on the value of ICs (AIC, AICc or/and BIC).
Step 5	Pick a model M inside S and estimate the corresponding IC value. If $M_{IC} < L$ and $M_{p_{index}} < \varepsilon$ then go to next step. Otherwise remove M from S and repeat Step 5.
Step 6 ks,	Check the goodness of fit of the candidate model using the normality ad and s tests. If any of them confirms the goodness of fit of the processed model then go to next step. Otherwise remove M from S and go to Step 5.
Step 7	Evaluate the performance of the accepted model.
Step 8	Forecast using the accepted model.

4. Numerical Experiments

For the performance of the proposed algorithm a set of tourism data from the Region of Western Greece were used. Specifically, the overnight stays from the three prefectures (Achaia, Etoloakarnania and Ilia) of the Region were used, from January of 2005 till December 2012. All data employed in this study were obtained from the official records of the Hellenic Statistical Authority. It is underlined that Hellenic Statistical

Authority has not released any similar data for the period 2013 until now.

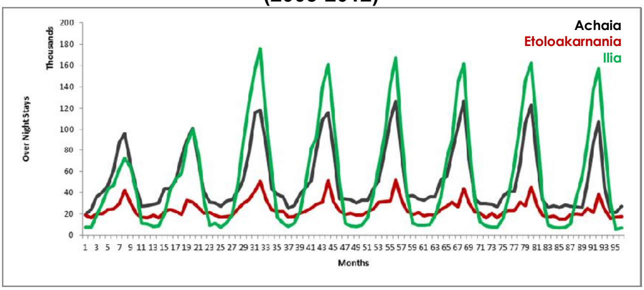
4.1 The Region of Western Greece

The basic reason for selecting the Region of Western Greece is the inherent dissimilarities of the prefectures of the Region, the different type of visiting tourists, the available resources and infrastructures and the level of development and employment.

The region of Western Greece consists of three prefectures: the prefectures of Achaia, Etoloakarnania and Ilia. The land in these areas is mostly mountainous but with extensive coastline at all prefectures. The main economic activities in the region are agriculture and tourism services, with significant wine and olive oil production. The region accommodates many, various and significant sensitive ecosystems. Finally, the region has significant tourism infrastructures which include 2 airports and 6 ports.

The following time-series chart (Figure 1) describes the fluctuation of overnight stays over time for the three prefectures of the region.

Figure 1: Overnights stays for the Region of Western Greece (2005-2012)



For all three prefectures a strong seasonality with considerable variation in the number of overnight stays between summer and winter months is revealed. The maximum stays are observed in August and for the three summer months a significantly high number of overnight stays is noticed. For the winter months the overnight stays are very low. The greater volatility in overnight stays occurs in prefecture of Ilia, while the smallest fluctuation in overnight stays between summer and winter months is observed in the prefecture of Etoloakarnania.

The entire region of Western Greece discloses a decreasing trend over the last 3 years of the observed data (2010-2012): Etoloakarnania with 35%, Achaia with 17.91% and Ilia with 6.06% in overnight stays.

4.2 The implementation

The implementation of the proposed algorithm needs some settings to be considered, as well as some of the adopted heuristics to be estimated. The used settings are the result of a series of test over the given data.

Firstly, a log transformation is applied to the used time-series to improve the model adaptation of SARIMA models. Next, we have to define the model space **S**. Studying the only previous work in this region (Panagopoulos and Panagopoulos, 2005) we decided that the model space **S** should have models with arguments lying in {0,1,2,3}. In addition the search process should examine SARIMA models with this priority order for their arguments: number of non-seasonal differences needed for stationarity, number of moving average terms, number of autoregressive terms, number of seasonal differences needed for stationarity, number of seasonal moving average terms and number of seasonal autoregressive terms.

In the next, the value of heuristic parameter **L** should be defined. Thus, in this work the threshold **L** was decided to be set with a more relaxed value: $L = 0.9 X BIC_{\min}$

where BIC_{min} is the minimum value of BIC criterion occurred after the preprocessing step (Algorithm 2: Step 4).

As it mentioned in the previous subsection, one of the characteristics of the Region of Western Greece is that the maximum overnight stays occurs in August. So, we would like models to perform well at the forecasting period and especially at August, where the maximum overnight stays are observed. So, the main idea behind the termination criterion (3.2) is that the algorithm should suggest models that "catch" as accurate as possible these months of the time-series. Therefore, the following performance index is defined, named as Partial Mean Absolute Error (**pMAE**), and it is based on the metric Mean Absolute Error

$$pMAE = \frac{1}{N}\sum_{t\in A}\left|\tilde{Y}_t - Y_t\right|$$

where $\boldsymbol{A}$ is the set of indices corresponding to all Augusts of the test set.

The user defined value $\boldsymbol{\varepsilon}$ seems to depend strongly on the studied time-series and the used performance index. The proposed algorithm was tested using various values of $\boldsymbol{\varepsilon}$ ($\boldsymbol{\varepsilon}$ varies between 0.025 and 0.1).

Lastly, the proposed algorithm was implemented in R (R Core Team, 2015). All the SARIMA models were constructed using the corresponding implementation from the forecast package (Hyndman, and Khandakar, 2008; Hyndman, 2015). The resulting models were compared with the resulting model of the auto.arima algorithm as it is implemented in the forecast package.

4.3 Numerical results

The proposed algorithm, along with the suggested settings was applied on the time-series from each prefecture of the Region of Western Greece. It is clear from Table 3 that the models from the proposed algorithm depend strongly on the selection of parameter $\boldsymbol{\varepsilon}$.

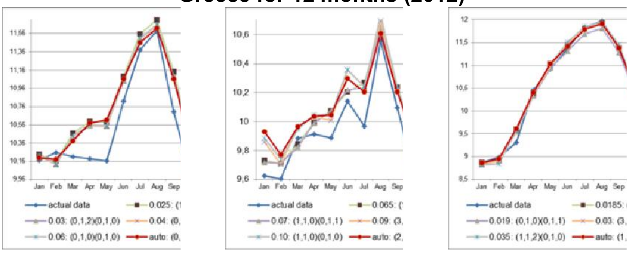
The Figure 2a shows that every model derived from the proposed algorithm seems to be an over fitting in respect of the auto.arima procedure for the prefecture of Achaia. However, for the prefecture of Etoloakarnania (Figure 2b) we can observe that the proposed technique produces two models ($\boldsymbol{\varepsilon}$=0.065, $\boldsymbol{\varepsilon}$=0.07), which try to overcome the intense fluctuation of the time-series and produce a smoother approximations. On the other hand (Figure 2c), both procedures showed similar behavior for the case of prefecture of Ilia.

The performance measurement was realized in two ways: the metrics were applied in the whole test set and then only over the last 12 months, as it shown in the following Tables 2 and 3. For all prefectures the metrics for the last 12 months shows that there exists at least a model that behaves slightly better than the auto.arima's one. It is, also, noticed, that the proposed algorithm behaves better when a lot of fluctuations exist.

Table 3: The resulting models of the proposed algorithm vs auto.arima

	ε	SARIMA	Models checked	auto.arima
Achaia	0.025	(1,0,2)(0,1,0)	13	(0,1,1)(0,1,1)
	0.03	(0,1,2)(0,1,0)	6	
	0.04	(0,1,1)(0,1,0)	4	
	0.06	(0,1,0)(0,1,0)	2	
Etoloakarnania	0.065	(1,1,0)(0,1,2)	74	(2,0,0)(1,0,0)
	0.07	(1,1,0)(0,1,1)	42	
	0.09	(3,1,2)(0,1,0)	30	
	0.10	(1,1,0)(0,1,0)	10	
Ilia	0.0185	(2,1,2)(1,1,0)	150	(1,1,1)(0,1,1)
	0.019	(0,1,0)(0,1,1)	34	
	0.03	(3,1,0)(0,1,0)	26	
	0.035	(1,1,2)(0,1,0)	14	

Figure 2: Ex-post forecasting for the Region of Western Greece for 12 months (2012)



a. Achaia b. Etoloakarnania c. Ilia

Table 2: ME, RMSE and MAPE errors on test set

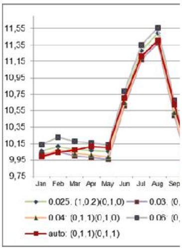
Achaia	ε	0.025	0.03	0.04	0.06	auto
	ME	0,0008	-0,0149	-0,0125	-0,0059	-0,0199
	RMSE	0,1079	0,1091	0,1100	0,1234	0,1028
	MAPE	0,7125	0,7396	0,7427	0,8401	0,6875
Etoloakarnania	ε	0.065	0.07	0.09	0.10	auto
	ME	-0,0019	-0,0022	-0,0008	-0,0005	0,0045
	RMSE	0,1032	0,1124	0,1315	0,1388	0,1327
	MAPE	0,7577	0,8169	0,9463	1,0295	1,0366
Ilia	ε	0.0185	0.019	0.03	0.035	auto
	ME	-0,0131	-0,0042	-0,0046	-0,0154	-0,0142
	RMSE	0,1478	0,1530	0,1592	0,1591	0,1459
	MAPE	1,0306	1,1032	1,0994	1,1382	1,0378

Lastly, the forecasting charts (Figure 3) for 2013 are presented for all prefectures of the Region of Western Greece.

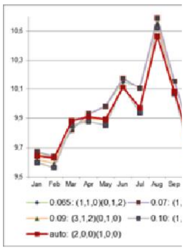
Table 3: ME, RMSE and MAPE errors on the last 12 months

Achaia	*ε*	0.025	0.03	0.04	0.06	auto
	ME	0,2259	0,1773	0,1787	0,1845	0,1817
	RMSE	0,2827	0,2432	0,2444	0,2487	0,2491
	MAPE	0,0233	0,0199	0,0199	0,0203	0,0193
Etoloakarnania	*ε*	0.065	0.07	0.09	0.10	auto
	ME	0,1128	0,1026	0,1343	0,1523	0,1477
	RMSE	0,1445	0,1356	0,1588	0,1715	0,1659
	MAPE	0,0120	0,0116	0,0139	0,0154	0,0150
Ilia	*ε*	0.0185	0.019	0.03	0.035	auto
	ME	0,0740	-0,0001	0,0521	0,0789	0,0819
	RMSE	0,1828	0,1703	0,1568	0,1789	0,1880
	MAPE	0,0119	0,0125	0,0111	0,0133	0,0122

Figure 3: Ex-ante forecasting for the Region of Western Greece for 12 months (2013)



a. Achaia



b. Etoloakarnania

c. Ilia

5. Conclusion

In this paper a heuristic algorithm is proposed for the automatic selection of SARIMA models, using time-series describing tourism demand. The presented algorithm constructs an efficient space model and a search process is initiated. To accelerate the algorithm two termination criteria were proposed based on a heuristic threshold of the used information criteria and a performance index which is a variation of well-known performance measurements. The algorithm was tested on tourism data from the prefectures of Region of Western Greece using overnight stays from January of 2005 till December 2012.

The encouraging result of this work is that the proposed algorithm, under an optimized selection of ***ε***, seems to result more promising models in comparison with the models of auto.arima function, especially for the cases of time-series with intense fluctuations. Or vice versa, the proposed algorithm seems to produce efficient SARIMA models if a proper value of ε is chosen. This issue is a crucial point for the presented work and should be thoroughly investigated. Furthermore, the performance and efficiency of the proposed algorithm should be tested in time-series with different characteristics. Finally, the used performance index in the form of a termination criterion can be extended to other metrics, used for the measurement of the performance of time-series.

Acknowledgement

This research has been co-financed by the European Union (European Social Fund - ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: ARCHIMEDES III, Investing in knowledge society through the European Social Fund. The data that involves the monthly occupancy of all tourist accommodations of both foreign and domestic tourists came from the official records of the Hellenic Statistical Authority (EL. STAT., www.statistics.gr).

6. References

1. Box, G. and Jenkins, G. (1976). *Time Series Analysis: Forecasting and Control*, 2nd ed. San Francisco: Holden-Day.
2. Gross J. and Ligges U. (2015). Nortest: Tests for Normality. R package version 1.0-3. Obtained through the Internet: http://CRAN.R-project.org/package=nortest, [accessed 03/08/2015].
3. Gunter, U. and Önder, I. (2015). Forecasting international city tourism demand for Paris: Accuracy of uni- and multivariate models employing monthly data. *Tourism Management*, 46: 123-135.
4. Hyndman R.J. (2015). Forecast: Forecasting functions for time series and linear models. R package version 6.1, Obtained through the Internet: http://github.com/robjhyndman/forecast, [accessed 03/08/2015].
5. Hyndman, R.J. and Athanasopoulos, G. (2013). Forecasting: Principles and practice. Obtained through

the Internet: http://otexts.com/fpp/, [accessed 03.08.2015].

6. Hyndman, R.J. and Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, 27(3): 1-22.
7. Inside-R (2015). A Community Site for R, Obtained through the Internet: http://www.inside-r.org/packages/cran/forecast/docs/ndiffs, [accessed 03/08/2015].
8. Panagopoulos, An. and Panagopoulos, Al. (2005). A time series method for occupancy forecasting - A case study of West Greece. *Archives of Economic History*, 17(1): 67-78.
9. R Core Team (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
10. Song, H. and Li, G. (2008). Tourism demand modelling and forecasting: A review of recent research. *Tourism Management,* 29: 203-220.