Modeling and Forecasting of Tourist Arrivals in Crete Using Statistical Models and Models of Computational Intelligence:
A Comparative Study

Stefanos K. Goumas, International Hellenic University, Greece

Stavros Kontakos, Democritus University of Thrace, Greece

iD https://orcid.org/0000-0001-9287-1775

Aikaterini G. Mathheaki, International Hellenic University, Greece

Sofoklis Xristoforidis, International Hellenic University, Greece

ABSTRACT

In the past few decades, tourism has clearly become one of the most prominent economic trends for many countries. For many destinations, this trend will continue to rise, and tourism will become the most dynamic and fastest growing sector of the economy. Thus, the reliable and accurate forecasting of tourism demand is necessary in making decisions for effective and efficient planning of tourism policy. The objective of this paper is the modeling and forecasting the international tourist arrivals to four prefectures of Crete in the year 2012, based on the actual tourist arrivals data over the period 1993 – 2011, using one-step-ahead forecast. In particular, this paper presented a comparative study of time series forecasts of international travel demand for the four prefectures of Crete using a variety of statistical quantitative forecasting models along with neural networks and fuzzy models.

KEYWORDS

Artificial Intelligence, Fuzzy C-Means, Fuzzy Models, Grid Partitioning, Neural Networks Models, SARIMA Models, Subtractive Clustering, Time Series, Tourism Demand

INTRODUCTION

Tourism is one of the most dynamic sectors of the global economy. Being the major economic activity in many countries, it has multiple and significant effects on many other sectors of the economy and it contributes to the overall economic development of the countries. Nowadays, tourism is not only an economic issue but also a social, cultural and political activity that involves both local societies and countries and relates to all social classes. International tourism creates jobs, contributes to the increase of the income, that of the savings, that of the investment and generally leads to an economic growth (Lim, 1997). Thus, its apparent that tourism is an economic activity of primary importance for many

countries. A tourist host country, which wishes to play a leading role in the competitive international tourism market, should be able to formulate a good policy in the field of tourism. Undoubtedly, marketing decisions and strategic planning of tourism benefits primarily require knowledge of the factors that influence the choice of destination of tourists (Seddighi & Theocharous, 2002).

Greece is nowadays a popular tourist destination, recording a continuous increase in arrivals of tourists since 1950 (Bouhalis, 2001). It's position in the world ranking of the arrivals is particularly high according to the data from the World Tourism Organization (W.T.O). In recent years Greece has a steady position between the 13th and 15th place on the list of countries with the highest number of arrivals, a position which automatically makes it one of the most important destinations internationally.

Particularly, Crete is one of the most popular Greek destinations. Over the 15% of the total arrivals to the ports and airports of the country are made via the city of Heraklion. At the year 2006 charter flights at Heraklion Airport are accounted for more than the 20% of all charter flights in the country and more than two million tourists visited Crete in the same year. This preference of tourists to Crete reflected on the number of hotel rooms, which increased in Crete by 53% from 1986 to 1991, while at the rest of Greek regions showed an increase of only 25%. Today's tourist infrastructure in Crete satisfies all the preferences. There is housing of every possible category from large, luxury hotels to smaller privately-owned family apartments or organized camping. The tourists' access to the island is made by air through the international airport of Heraklion and the airports in Chania and Sitia or by ferry, to the ports of Heraklion, Chania, Rethymno, Agios Nikolaos, Sitia and Kastelli Kissamos.

THE FORECAST OF TOURIST ARRIVALS

Because of the aforementioned significant contribution of tourism to the local economy and the nature of the tourist industry, the design of the tourism policy is crucial to the whole process of tourism development. It is widely accepted that any business design regarding tourism requires some form of decision-making, based on future forecasting (Palmer et al., 2006, Chandra & Menezes, 2001). Thus, the reliable and the accurate forecast of the tourism demand is essential in making decisions for effective and efficient planning of tourism policy. Moreover, the accurate forecasting of tourism demand is necessary for the successful design of the supply of goods, services and the infrastructures in tourism and in other service industries. Optimal forecasts of tourist arrivals in a country or in a particular geographical area, can effectively help business managers and investors to make operational, tactical and strategic decisions. These may include decisions on scheduling and on personnel management, on preparing tourist booklets and hotel investments etc. Furthermore, government councils need accurate tourist forecasting in order to design the required tourist infrastructure, such as new hotels in touristic places, as well as the creation of appropriate transport and communications. It is also necessary to predict the consumption of goods and services by tourists, in order to assess the relative contribution of tourism to production, income and employment in tourist destinations (Palmer et al., 2006).

The forecast configuration based generally on two basic conditions. The first one regards the quantitative method that is used and the second is how the economic variable behaves. The better the way the variables are created, the better predictions are expected to be formed. However, for this to be the case, the existing structure of the behavior of the variable's value should remain constant or almost stable in the near future.

The success of a forecast depends primarily on the availability of appropriate data sets (Montgomery et al., 1990; Hanke & Wichern, 2009). Data collection may, in some cases, be a difficult and time-consuming process and therefore it should be given particular attention in order to address possible problems related to availability and accuracy. Research on tourism demand forecast is based primarily on secondary data for the construction of the forecasting model (Song & Li, 2008). Secondary data is typically collected for some purpose and stored somewhere. Their sources of origin may be internal, such as e.g. balance sheets, income statement, invoices, etc. or governmental organizations,

such as the Ministry of Finance, the Ministry of Commerce, the National Statistical Service (NSI) and international organizations such as the International Monetary Fund (IMF) and the World Bank (WB). ICAP, chambers, stock exchange, cooperatives are also sources from which data can be drawn for statistical processing. In recent years, access to the above-mentioned sources of secondary data can be quickly accessed via the Internet.

Overview of Forecasting Methods

The forecasting methods that extensively used are, mainly, time series and econometric models. With econometric models, researchers make predictions of a (dependent) variable based on the economic and statistical relationship between this variable and a number of other variables (predictors). Thus, an econometric model is an extension of multiple regression methods and typically comprises several relationships to one or more variables, which appear in more than one model equations.

By time series, researchers mean a series of continuous data, characterized by a defined order of the observations over time, and a defined dependency between the observations of the series. This specified configuration and dependency guarantees the creation of reliable forecasts (Song & Li, 2008). In time series models, unlike econometric models, which are based on multiple regression techniques, the forecast relies solely on previous values of the same time series of interest. That is, the future behavior of a time series model is predicted not in relation to the changes of other variables that may affect it, but only by looking at its past behavior or the "history".

The advantages of time series models against econometric methods are that the former have lower cost of provisioning and are less complex. On the other hand, they have the disadvantage of not relying on a theory that explains how the values of the time series models are formed. On the contrary, they take into account the hypothetic scenario that whatever happened in the past will continue to happen in the future. Thus, they do not have a mechanism to allow for changes, that might influence their future values (for example, changes in government policy). This entails a reduction in the accuracy of forecasts, especially for long periods in the future, which is undesirable (Dimeli, 2002). For the above reasons, time series methods are more suitable for short-term forecasting, while econometric methods for long-term forecasting (Dimeli, 2002).

One of the greatest advantages of econometric approaches over time series lies in their ability to analyze the causal relation between the tourist demand (dependent variable) and the factors that affect it (independent variables), (Song & Li, 2008). Econometric analysis plays an important role, also in other fields, beyond forecasting. For example, econometric models consolidate the existing empirical and theoretical knowledge of how the economies of different countries operate, providing a framework for a progressive research strategy, and helping to explain their malfunctions (Clements & Hendry, 1998). Regarding tourist demand, econometric analysis has an empirical utility in interpreting any change in tourist demand from the point of view of economists, by providing policy proposals as well as evaluating the efficiency of the existing tourism policies. On the contrary, a time series model cannot help under special circumstances, i.e. where the interdependence between tourism demand and other relevant factors is of paramount importance to businesses and governments (Song & Li, 2008).

The most commonly used time series models are *Winter's Exponential Smoothing*, the method of *Moving Average – MA*, *Naive Methods*, the method of *the Decomposition* and the *Autoregressive Integrated Moving Average* (*ARIMA*) models of Box – Jenkins (Goh & Law, 2002; Song & Li, 2008; Lee et al., 2008). Different versions of ARIMA models have been applied to more than two-thirds of tourism forecasting studies since 2000. The last few years, the *Seasonal ARIMA* (SARIMA) models are more popular, since seasonality is an important feature of the tourist industry and therefore the seasonal variation of tourist demand for destinations is of direct interest to managers of tourism businesses (Song & Li, 2008).

In addition to time series and econometric models, a number of new quantitative forecasting methods, predominantly Artificial Intelligence (AI) methods, have appeared in the tourism forecasting literature (Cho, 2005; Palmer et al., 2006; Song & Li, 2008). The Artificial Intelligence is rapidly

developing as a research field in a variety of scientific areas in recent years. Old Artificial Intelligence methods used rule-based systems, as well as logic programming systems. In recent years, the research interest has focused on Computational Intelligence or Soft Computing techniques, particularly Artificial Neural Networks (ANN), Fuzzy Logic, and Genetic Algorithms, as well as in hybrid models of the aforementioned such as Neural-Fuzzy Systems and Fuzzy-Neural Systems (Principe et al., 2000; Palit & Popovic, 2005; Haykin, 2009). The main advantage of Artificial Intelligence techniques is that they do not require any preliminary or additional information about the data such as, for example, their distribution.

The method of Artificial Neural Networks (ANN) is a computational technique, which attempts to simulate the learning process of the human brain (Law, 2000). The unique features of ANNs, such as their ability to adapt to incomplete data, to non - linearity, and to an arbitrary depiction of a function, make this method a useful alternative to classical (statistical) forecasting models. ANNs were introduced to tourism forecasts in the early 1990s and some improved ANN techniques continued to appear in the literature until nowadays (Song & Li, 2008).

The fuzzy time series method has a strong performance in analyzing small time series, using only few observations of the past. Wang (2004) has applied the fuzzy time - series and the fuzzy theory by developing three different models for forecasting tourist arrivals to Taiwan from Hong Kong, Germany and USA. He found out that the forecasting performance of the three models depends on the country of origin and does not outperform other forecasting techniques in all cases.

RESEARCH METHODOLOGY

The purpose of this paper is to predict the arrivals of tourists from all over the world in the four prefectures of Crete, for the year 2012, based on the actual data, i.e. the arrivals for the years 1993 to 2011, for each month, with one-step ahead. The data have been collected per Crete prefecture and the word arrivals include only the number tourists that used any kind of hotel accommodation, except camping. The data was obtained from the National Statistical Service (NSS). Then, the predicted values will be compared to the actual data of the arrivals of the year 2012. The objective of the survey is to compare the various statistical time series and neural network models, in order to identify the appropriate method for predicting the tourist demand for each one of the four prefectures of Crete.

The time series analysis was performed with the statistical software Minitab 17.0. Initially, from the time series diagram shown in Fig.1 seems that the time series format is similar at the four prefectures and shows that the time series are generated by the composition of a trend component, a 12-month seasonal component and the random component. Consequently, it cannot be assumed that the specific time series is stationary. The synthesis model of the three components is the "Multiplication Model".

Therefore, the forecasting methods used are suitable for seasonal time series and these are the Winter's method and SARIMA. Also, a method using Neural Networks is used. Moreover, three methods of *Adaptive Network-based Fuzzy Inference System* (ANFIS) with Neural Networks (ANFIS NN) are used namely Fuzzy C-Means (FCM), Subtractive Clustering, and Grid Partitioning. These methods allow classifying the input data into groups. Each one of the groups has similar properties, that helps to understand the correlation between the data and, as a result, to simplify the forecasting process. A brief description of each one follows.

First, the FCM algorithm is used widely in clustering methods. It separates the data into groups by optimizing an objective function (Dunn, 1973; Bezdek, 1981). In this case, the tourist arrivals time series presents a non-linearity, which makes the clustering process more difficult. Hence, phase space reconstruction is used in order to understand the underlying dynamics of this time series (MacQueen, 1967; Benmouiza et al., 2016). Second, the Subtractive Clustering technique is applied when there is not a clear idea about the number of centers for the distribution of data. Subtractive method is an extension of the classification method proposed by Yager and Filev (1994). In this algorithm, each data point is considered as a cluster center candidate, and then it calculates the potential of each data

Figure 1. The time series scatterplot for each one of the four prefectures of Crete
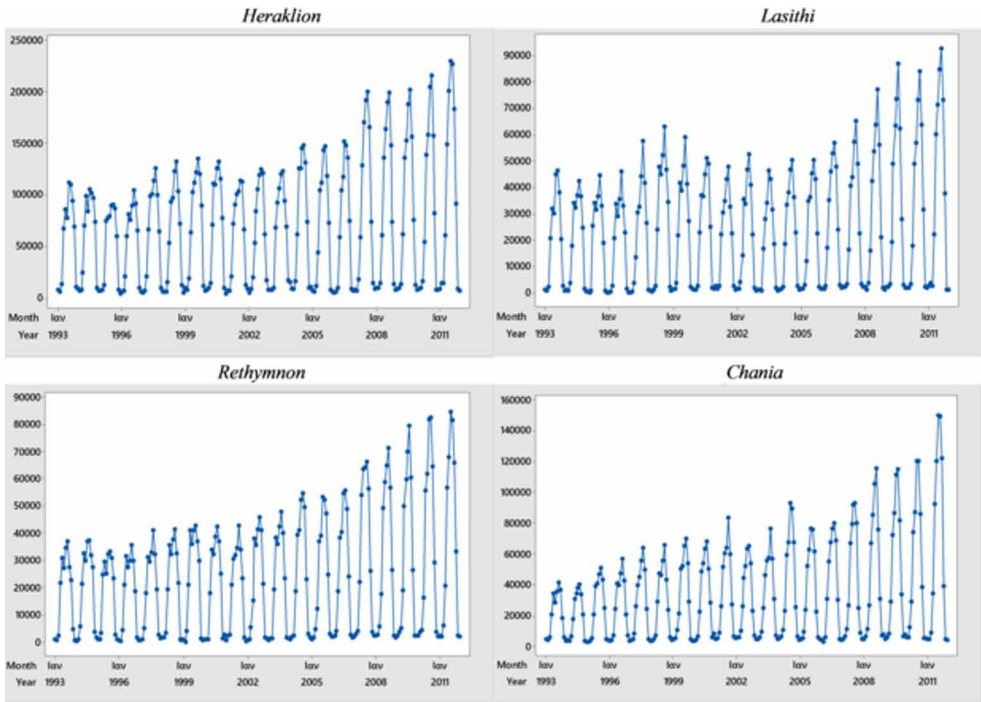
point by measuring the density of points data surrounding it. Finally, in Grid Partitioning method, the input data space is divided into rectangular subspace using an axis-paralleled partition; each input is partitioned into identically shaped membership functions. The number of the fuzzy if-then rules is equal to m^n, where n is the input dimension and m is the number of partitioned fuzzy subsets for each input variable. The grid is constructed without taking any physical meaning or data density repartition, and each part of the grid is used to generate fuzzy rules based on system input–output training data, which allows fast learning processes and calculation time optimization. However, the performance of this method depends strongly on the size of the inputs and the grid; generally, a finer grid leads to high performance. More information and detailed steps, as well as the mathematical formulation of the above procedures can be found, among others, in Wang (2004) and Benmouiza, et al. (2016).

As mentioned before, a comparative study between the three abovementioned methods of ANFIS NN is carried out. The aim of the research is to determine whether these methods are equally effective for the prediction of the arrivals at all prefectures of Crete or the element of locality is involved.

RESULTS

Initially, the method of *Exponential Smoothing of Trend and Seasonality (Winter's Method)* was applied. The smoothing parameters Alpha (α), Gamma (γ) and Delta (δ) which account for Level, Trend and Seasonality respectively, where calculated by the Solver Programmer of Microsoft Excel by minimizing the Mean Absolute Deviation (MAD) accuracy measure. The starting values were 0.1 for all the parameters and the optimal values obtained by the program are $\alpha=0.2$, $\gamma=0.18$ and $\delta=0.22$, for all the prefectures. The monthly forecasts for this method, as well as the real arrivals of year 2012, are presented in Table 1.

Table 1. Forecast of tourist arrivals in four prefectures of Crete for the year 2012 with Winter's Method and real arrivals

Month	Heraklion		Lasithi		Rethymnon		Chania	
	Forecast	Arrivals 2012	Forecast	Arrivals 2012	Forecast	Arrivals 2012	Forecast	Arrivals 2012
Jan/2012	9.898	8.198	2.314	1.364	2.277	2.719	7.014	5.081
Feb/2012	10.945	10.608	2.187	1.685	2.720	3.302	7.750	4.679
Mar/2012	17.363	17.680	3.225	1.787	5.112	3.400	13.310	7.708
Apr/2012	69.314	62.868	18.540	19.707	21.015	19.426	39.413	34.879
May/2012	142.921	144.840	41.885	56.318	51.883	48.117	90.072	85.358
Jun/2012	163.007	199.843	46.191	68.351	56.132	65.296	105.225	123.693
Jul/2012	186.484	237.323	53.603	77.713	65.939	78.811	124.679	130.561
Aug/2012	184.813	237.954	57.482	95.596	67.043	75.481	124.348	135.272
Sep/2012	143.992	192.701	41.657	74.689	53.897	62.952	91.819	102.888
Oct/2012	76.156	97.718	20.271	35.489	27.318	28.239	36.754	35.094
Nov/2012	11.096	10.508	1.797	1.744	2.773	2.550	7.799	5.994
Dec/2012	8.089	6.915	1.112	1.250	1.987	2.097	6.943	4.755

From Table 1, it can be seen that Winter's method forecasts are close enough to actual data, almost for all prefectures, but in some cases, there are divergencies for the real arrivals. Therefore, Winter's method gives good results and could be considered successful, in general. This result is consistent with the theoretical potential of the Winter's model, which accounts for trend and seasonality in a time series model. From Table 1 it is also clear that, despite the similar structure of the time series of tourist arrivals at the four prefectures, the forecasts of tourist arrivals for the Lasithi and Chania are better than the forecasts regarding prefectures of Heraklion and Rethymnon, for the year 2012. Moreover, is apparent from both Table 1 and Fig.2, that forecasts with Winter's method have a significant divergence from real arrivals in summer months for the preferences of Lasithi and Rethymnon. In Table 2 is presented the accuracy assessment measures Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE) and MAD of the Winter's method for each one of the four prefectures of Crete.

Table 2 shows that, based on the MAPE criterion, the Winter's method gives better predictions for the prefectures of Heraklion and Chania (i.e. the mean percentage of error is smaller), and based on the MAD and RMSE criterion (i.e. in absolute numbers of arrivals), this method gives better predictions for the prefectures of Lasithi and Rethymnon. Thus, the evaluation of the specific method, based on two out of three accuracy measures shows similar results for the prefectures of Crete.

Table 2. The evaluation of Winter's method based on accuracy assessment measures

Prefecture	MAPE	MAD	RMSE
Heraklion	16	8879	13655
Lasithi	23	3377	5220
Rethymnon	29	3478	5084
Chania	15	4608	7945

Table 3. Forecast of tourist arrivals in four prefectures of Crete for the year 2012 with SARIMA Model and real arrivals

Month	Heraklion		Lasithi		Rethymnon		Chania	
	Forecast	Arrivals 2012	Forecast	Arrivals 2012	Forecast	Arrivals 2012	Forecast	Arrivals 2012
Jan/2012	9.893	8.198	2.722	1.364	2.761	2.719	4.408	5.081
Feb/2012	13.404	10.608	3.315	1.685	2.806	3.302	4.566	4.679
Mar/2012	15.495	17.680	3.276	1.787	6.299	3.400	8.860	7.708
Apr/2012	59.978	62.868	21.135	19.707	20.176	19.426	31.558	34.879
May/2012	145.275	144.840	56.460	56.318	56.468	48.117	83.781	85.358
Jun/2012	185.637	199.843	67.041	68.351	67.042	65.296	106.508	123.693
Jul/2012	218.596	237.323	80.711	77.713	83.666	78.811	135.732	130.561
Aug/2012	221.002	237.954	89.689	95.596	81.806	75.481	136.105	135.272
Sep/2012	173.953	192.701	69.832	74.689	65.815	62.952	106.672	102.888
Oct/2012	87.958	97.718	35.254	35.489	32.841	28.239	37.231	35.094
Nov/2012	9.962	10.508	1.955	1.744	3.070	2.550	4.398	5.994
Dec/2012	8.024	6.915	1.580	1.250	2.392	2.097	4.267	4.755

Because, the above time series have a strong seasonality, for more accurate forecasts, the Multiplier Model SARIMA (p, d, q) (P, D, Q)$_S$, is also used, in order to predict the tourist arrivals in the four prefectures of Crete. After several experimental SARIMA models from SARIMA (0,0,0) (0,0,0)$_{12}$ to SARIMA (1, 1, 10) (1, 1, 10)$_{12}$ the model SARIMA (0, 1, 1) (0, 1, 1)$_{12}$ was chosen, which gave the best forecasts. The results of the SARIMA forecasts are presented in Table 3.

Results presented in Table 3 show that the model SARIMA (0, 1, 1) (0, 1, 1)$_{12}$ gives very good forecasts for all the four prefectures of Crete, especially for Lasithi, Rethymnon and Chania. For the prefecture of Heraklion, there seems to be a slight underestimation mainly in summer months, in September and in October. The RMSE measure for the evaluation of the SARIMA model to all the prefectures of Crete is presented in Table 4.

Table 4 shows that, based on the RMSE criterion, the SARIMA method gives better predictions for the prefectures of Lasithi and Rethymnon since the values of the RMSE of these two prefectures are smaller. The RMSE value for the Heraklion is the greatest one and almost reaches the 10500 arrivals.

A Comparative Study between Winter's, SARIMA and Neural Network Methods

Next, a comparative assessment was made in the results of time series forecasts provided by seasonal models Winter, SARIMA, and the model with the use of Neural Networks, for the tourist arrivals in

Table 4. The evaluation of SARIMA model based on Root Mean Square Error

Prefecture	RMSE
Heraklion	10456
Lasithi	2552
Rethymnon	3822
Chania	5475

Figure 2. The predictions of the arrivals of tourists, in each one of the prefectures of Crete, by the time series (Winter and SARIMA) and the Neural Network models
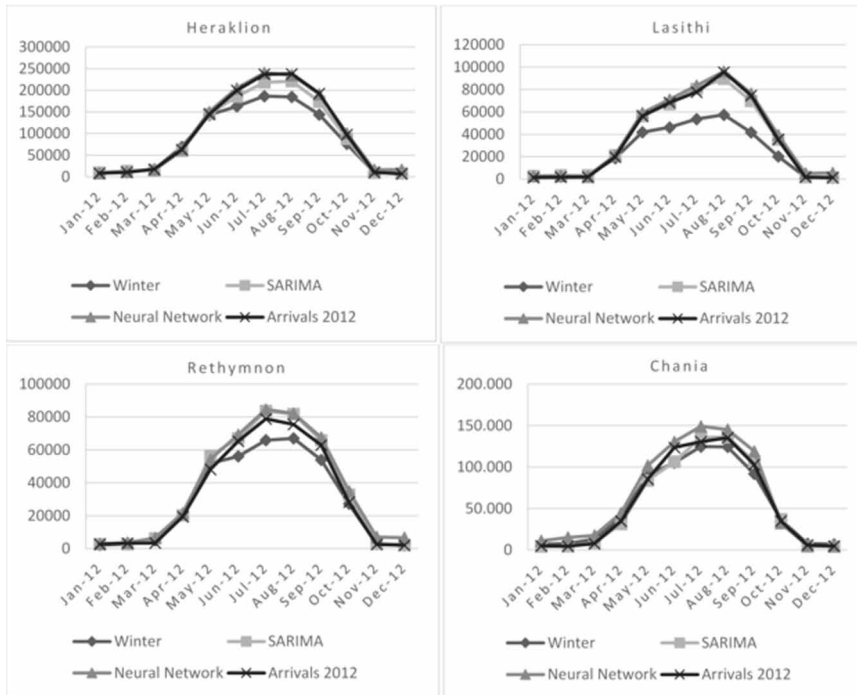
each one of the four prefectures of Crete. In the study, a Feed-Forward Multilayer Neural Network is used, with the backpropagation algorithm. More specific, the Neural Network used in the analysis has two neurons to the input layer, 10 neurons to the hide layer and one neuron to the output layer. The simulation of the Neural Network has made via Neural Networks Toolbox of MATLAB 2016. The comparative graphs that were produced show the curves of both the actual data of the arrivals of the year 2012 and the corresponding forecasts, using the above-mentioned models, for the four prefectures of Crete. The comparison is made for each month of the year 2012. The graphs are presented in Figure 2.

From Figure 2, it can be seen that the SARIMA model overtakes the Winter model in forecasting arrivals of tourists in the Prefectures of Heraklion, Lasithi and Rethymnon. In the Prefecture of Chania, the forecasts of the arrivals by Winter and SARIMA models are very close (although Winter's method seems to give a slight better forecasts) and both models can be considered to predict with great precision the real arrivals of the tourists in the specific prefecture. Moreover, the SARIMA model for the Prefectures of Rethymno and Chania seems to be better, compared to the neural network, while the neural network gives a better prediction for the Prefectures of Heraklion and Lasithi. Finally, it is concluded that the SARIMA model has a clear lead over the Winter model, in most cases, in predicting tourist arrivals, in the four prefectures of Crete, while the SARIMA and neural network models appear to be equivalent, as results from this study. Moreover, the accuracy measure of RMSE for the NN model, presented in Table 5, verifies that these two models seem to have similar results. The errors made for the prefectures of Lasithi and Rethymnon are smaller than the corresponding errors for the other two prefectures of Crete.

Table 5. The evaluation of Neural Network model based on Root Mean Square Error

Prefecture	RMSE
Heraklion	8125
Lasithi	3955
Rethymnon	3526
Chania	4724

A Comparative Study Between The Three ANFIS NN Methods

The first evaluation of the three ANFIS NN methods, as they described analytically in research methodology session was carried out for the Prefecture of Heraklion. The results are presented in the following Figure 3.

Figure 3 shows that ANFIS NN with Grid Partitioning method gives the worst forecast of the tourist arrivals. In spite of the fact that from January to July there seems to be no difference between the three methods and the real arrivals, for the months after August, the divergence of the predictions of Grid Partitioning from both the real arrivals and the other two ANFIS methods is significant high. Thus, the method of Grid Partitioning is not used for the other three prefectures of Crete for the prediction of tourist arrivals, in order for the results to be comparable. Figure 4 presents the forecast of tourist arrivals for the other prefectures of Crete (Lasithi, Rethymnon, Chania), with Subtractive Clustering and FCM methods, together with the actual arrivals for the year 2012.

Figures 3 and 4 present the full picture of ANFIS NN. From the specific graphs it is obvious that the two ANFIS network methods (Subtractive Clustering and FCM) give the best results, i.e. its forecasts are very close to the real arrivals of tourists at year 2012, for the prefecture of Heraklion. Moreover, comparing the prefectures of Rethymnon and Chania (Figure 4), the forecast lines seem to have the same pattern, i.e. the three lines are almost the same in all months, except May and July. In May, the Subtractive Clustering overestimates the arrivals and FCM underestimates them, whereas

Figure 3. The predictions of the arrivals of tourists, in Prefecture of Heraklion, by each one of the ANFIS NN models
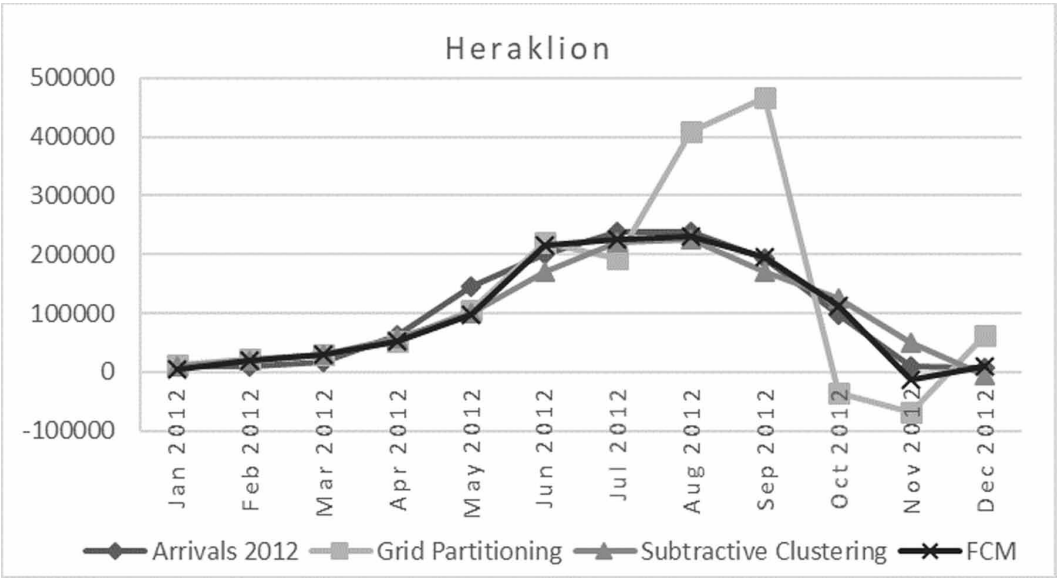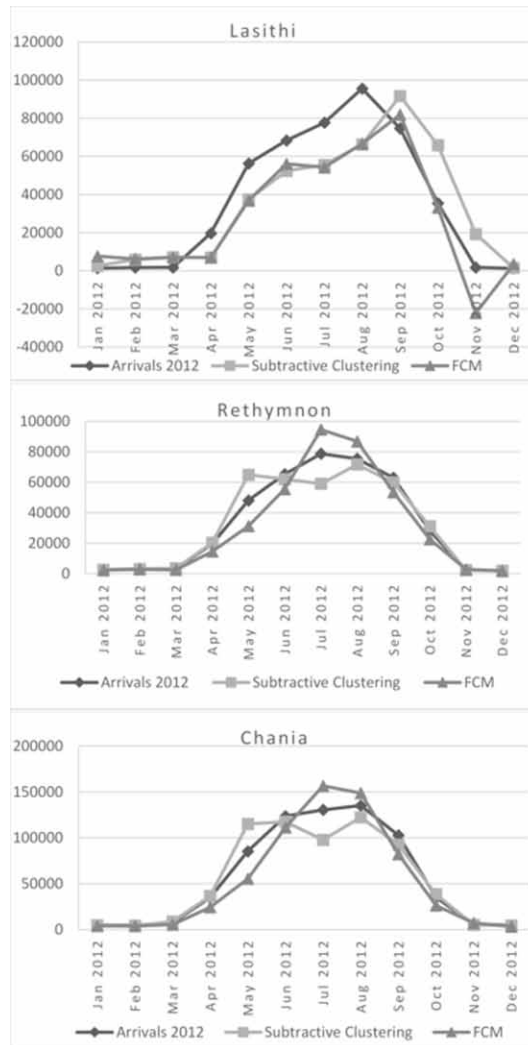
Figure 4. The predictions of the arrivals of tourists, in Prefectures of Lasithi, Rethymno and Chania, by each one of the ANFIS NN models: Subtractive Clustering and FCM

in July seems to happen the opposite. In Lasithi, the two methods of ANFIS NN models gave almost the same results, and both seem to underestimate the real arrivals from April to August. However, comparing these results of Subtractive Clustering (ANFIS) with the corresponding ones with the method of Backpropagation Neural Network (Figure 2), Backpropagation NN, seems to be better in the prediction of tourist arrivals in all the prefectures of Crete.

Finally, Figure 5 presents the comparison of the three ANFIS NN models with SARIMA model for each one of the four preferences of Crete.

Figure 5 shows that for the prefecture of Heraklion the ANFIS NN models have similar forecasts with SARIMA model and both are extremely close to the real arrivals. On the other hand, for the other three prefectures, SARIMA model seems to have much better forecasts than the two ANFIS NN models. The RMSE criterion for the evaluation of the two ANFIS NN methods, for all of the prefectures of Crete, is presented in Table 6.

Figure 5. The prediction of the arrivals of tourists by the two ANFIS NN and the SARIMA model for each one of the four prefectures of Crete
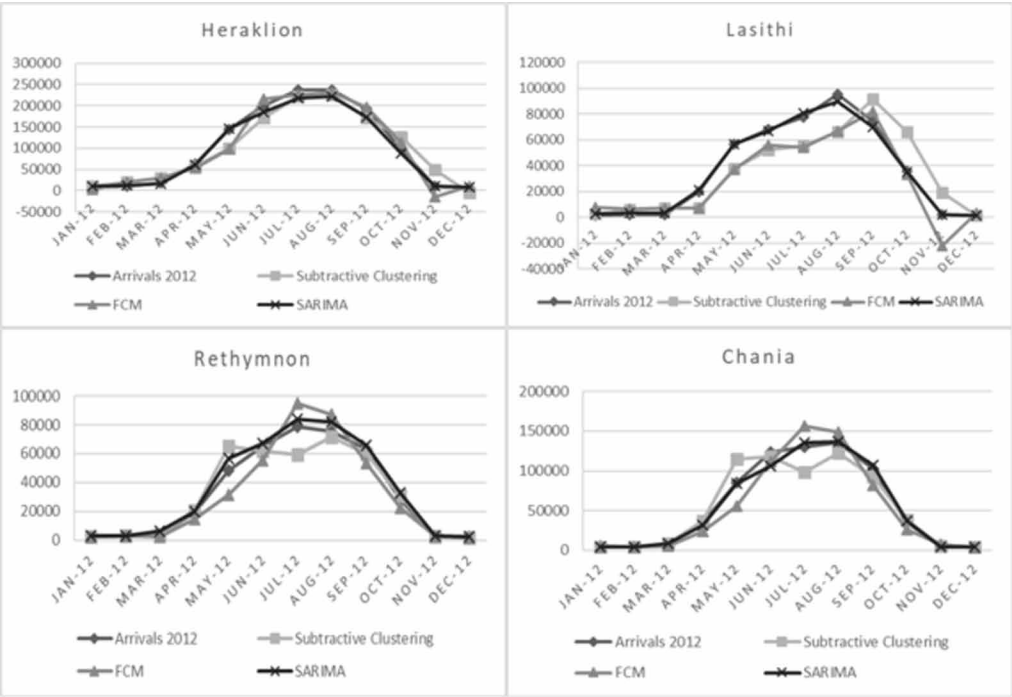

Table 6 shows that Rethymnon has the smallest and the Heraklion the largest RMSE among the three prefectures of Crete with both ANFIS NN methods. Moreover, regarding the evaluation of the specific methods FCM seems to have smaller errors than Subtractive Clustering even if in Rethymnon and Chania the values are almost equal.

CONCLUSION

Tourism is at the forefront of global development policy, and especially of the Greek economy. Greece has the comparative advantage of both natural and cultural wealth, which are combined with its tradition of hospitality. Moreover, there is multiple interconnection of tourism with many other sectors of the national economy. Crete's tourism, in particular, makes a significant contribution to the Greek tourist industry, as Crete is a popular tourist destination with an annually increasing tourist

Table 6. The evaluation of ANFIS NN models based on Root Mean Square Error

	RMSE	
Prefecture	**Subtractive Clustering**	**FCM**
Heraklion	23453	18056
Lasithi	17534	15273
Rethymnon	7722	7670
Chania	13799	13519

arrivals from places, almost all around the world. Optimal forecasts of tourist arrivals in a country or in its particular geographical area can effectively help business executives, investors and government officials in order to make operational, tactical and strategic decisions on the design of tourism policy.

In the present study, a variety of statistical quantitative forecasting models were applied in order to analyze the chronological series of international tourist arrivals of the four prefectures of Crete (i.e. Heraklion, Lasithi, Rethymno and Chania), during the period 1993 – 2011. The main purpose was to predict the arrivals for the year 2012 in the aforementioned prefectures and to compare them with the real arrivals. In addition, different techniques of the area of artificial intelligence were used, together with methods based on fuzzy logic, in order to determine to what extent all these methods can give satisfactory predictions, both in comparison with each other and also among the four prefectures of Crete.

Initially, the Winter's method was used, in order to predict the tourist arrivals. The results showed that there were significant divergences of the forecasts made with this method, with the corresponding actual arrivals. The smallest deviations are presented for the time series regarding prefectures of Heraklion and Chania. The greatest deviation of the arrivals, that this method gave, appeared to the prefecture of Lasithi and especially during the summer months. For the Heraklion, the number of the arrivals at the summer of the year 2012 are higher, than the corresponding forecasts. This shows that Winter's method underestimated the summer arrivals of the tourists in Heraklion and Lasithi.

At the next stage, SARIMA models were used for the prediction of tourist arrivals. These models have better prediction results, than the Winter's model, in general. Their main disadvantage is that they use several parameters and moreover, these parameters are not known in advance. Thus, for safe predictions, many experimental trials have to be made, in order to calculate those parameters, and this is a time-consuming procedure. From the results obtained from the application of the SARIMA models, and compared to the corresponding results of the Winter method, someone can conclude that the model SARIMA $(0,1,1)(0,1,1)_{12}$ prevails the Winter's model in the forecast of tourist arrivals in Heraklion, Lasithi and Rethymnon, i.e. in three out of four Crete Perfectures. On the contrary, the Winter's model has a slight advantage over the model SARIMA $(0,1,1)(0,1,1)_{12}$ in forecasting of the arrivals in the Prefecture of Chania. Finally, it is concluded that the model SARIMA $(0,1,1)(0,1,1)_{12}$ seems to be better than the Winter's model, in the majority of the tourist arrivals in the four prefectures of Crete.

In addition to statistical time series prediction models, neural network models were also used, in order to predict tourist arrivals. The innovation of neural networks models lies in their ability to model nonlinear relationships between variables without the necessity to know exactly the nature of the process they represent. Artificial neural networks are not bound by any kind of limiting hypotheses, as in other predictive techniques, and therefore no tests are required to formulate hypotheses. The comparative analysis between the SARIMA model and the Neural Network model, showed that the SARIMA model gave better forecasts for the prefectures of Rethymno and Chania, whereas the Neural Network model gave slightly better forecasts for the Heraklion and Lasithi. Therefore, these two models seem to be equivalent, at least as arises from the present study. Finally, in addition to the aforementioned methods, neural network methods with fuzzy logic theory (ANFIS NN) were used to predict tourist arrivals. A comparative analysis was performed between three such models, named FCM, Subtractive Clustering and Grid Partitioning. The analysis showed that FCM and Subtractive Clustering methods can predict very well the tourist arrivals and give similar results with SARIMA models.

Finally, the survey results are very satisfactory and consistent with the conclusions drawn from similar empirical studies of the international literature. Therefore, it is also confirmed in the case of Crete, as a popular tourist destination, that there is no universal method for forecasting tourist demand. Tourism is a dynamic phenomenon that is continuously evolving and influenced by many different factors, which can also be changed, in turn, with the changes in society. For this reason, forecasting methods should also be updated and evolved, in order to provide safer conclusions, as

far as possible. Forecasting not with one, but with many different methods will be a safe way, of estimating forecasts of tourist arrivals.

Future research could focus on more modern methods, regarding both time series and artificial intelligence methods. These methods can be applied not only to the prediction of tourist arrivals but also in many other fields of the sciences, as for example, to predict arrivals at the emergence rooms. Aladeemy et al. (2016) used ARIMA models in order to predict those arrivals and the results are very interesting. *Vector Autoregressions Models* (*VAR models*), for example, could be used in order to analyze both tourist and patient arrivals via time series models (Song & Witt, 2006; Diebold, 2007; Song & Li, 2008). These models are an extension of the univariate autoregressive models AR(p) and are mainly used today in the place of econometric models of equation systems. VAR models have been proven to be more effective and successful in predicting interconnected systems, on one hand, and easier and faster to use, on the other (Diebold, 2007). Also, a combination of ANN and classic methods could be used, such as the SARIMABP hybrid model, which is a combination of the SARIMA model and a Backpropagation Network model, proposed by Tseng et al. (2002). Finally, Box-Jenkins ARMA/ARIMA models could be combined with Holt's exponential smoothing model, using a fuzzy model as *nonlinear forecasts combiner* (Palit & Popovic, 2005).

ACKNOWLEDGMENT

REFERENCES

Aladeemy, M., Shan, X., Poranki, S., & Khasawneh, M. T. (2016). *Daily Emergency (ED) Arrival Forecast for Reducing ED Holds*. *Healthcare Systems Process Improvement Conference*, Houston, TX.

Benmouiza, K., Tadj, M., & Cheknane, A. (2016). Classification of hourly solar radiation using fuzzy c-means algorithm for optimal stand-alone PV system sizing. *International Journal of Electrical Power & Energy Systems*, *82*, 233–241. doi:10.1016/j.ijepes.2016.03.019

Bezdek, J. C. (1981). *Pattern recognition with fuzzy objective function algorithms*. Springer. doi:10.1007/978-1-4757-0450-1

Buhalis, D. (2001). Tourism in Greece: Strategic Analysis and Challenges. *Current Issues in Tourism*, *4*(5), 440–480. doi:10.1080/13683500108667898

Chandra, S., & Menezes, D. (2001). Applications of Multivariate Analysis in International Tourism Research: The Marketing Strategy Perspective of NTOs. *Journal of Economic and Social Research*, *3*(1), 77–98.

Cho, V. (2005). A comparison of three different approaches to tourist arrival forecasting. *Tourism Management*, *24*(3), 323–330. doi:10.1016/S0261-5177(02)00068-7

Clements, M. P., & Hendry, D. F. (1998). *Forecasting Economic Time Series*. Cambridge University Press. doi:10.1017/CBO9780511599286

Diebold, F. (2007). *Elements of Forecasting*. Thomson.

Dimeli, S. (2002). *Σύγχρονες Μέθοδοι Ανάλυσης Χρονολογικών Σειρών*. Αθήνα: Εκδόσεις Κριτική.

Dunn, J. C. (1973). A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, *3*(3), 32–57. doi:10.1080/01969727308546046

Goh, C., & Law, R. (2002). Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. *Tourism Management*, *23*(5), 499–510. doi:10.1016/S0261-5177(02)00009-2

Hanke, J., & Wichern, D. (2009). *Business Forecasting*. Prentice Hall.

Haykin, S. (2009). *Neural Networks and Learning Machines*. Pearson Education, Inc.

Law, R. (2000). Back-propagation Learning in Improving the Accuracy of Neural Network-Based Tourism Demand Forecasting. *Tourism Management*, *21*(4), 331–340. doi:10.1016/S0261-5177(99)00067-9

Lee, C., Song, H., & Mjelde, J. (2008). The forecasting of International Expo tourism using quantitative and qualitative techniques. *Tourism Management*, *29*(6), 1084–1098. doi:10.1016/j.tourman.2008.02.007 PMID:32287725

Lim, C. (1997). Review of International Tourism Demand Models. *Annals of Tourism Research*, *24*(4), 835–849. doi:10.1016/S0160-7383(97)00049-2

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, Volume 1: *Statistics*. The regents of the University of California.

Montgomery, D., Johnson, L., & Gardiner, J. (1990). *Forecasting and Time Series Analysis*. McGraw-Hill.

Palit, A., & Popovic, D. (2005). *Computational Intelligence in Time Series Forecasting*. Springer.

Palmer, A., Montano, J., & Sese, A. (2006). Designing an artificial neural network for forecasting tourism time series. *Tourism Management*, *27*(5), 781–790. doi:10.1016/j.tourman.2005.05.006

Principe, J., Euliano, N., & Lefebvre, C. (2000). *Neural and Adaptive Systems*. John Wiley & Sons, INC.

Seddighi, H. R., & Theocharous, A. L. (2002). A Model of Tourism Destination Choice: A Theoretical and Empirical Analysis. *Tourism Management*, *23*(5), 475–487. doi:10.1016/S0261-5177(02)00012-2

Song, H., & Li, G. (2008). Tourism Demand Modeling and Forecasting— A Review of Recent Research. *Tourism Management*, *29*(2), 203–220. doi:10.1016/j.tourman.2007.07.016

Song, H., & Witt, S. F. (2006). Forecasting International Tourist Flows to Macau. *Tourism Management*, *27*(2), 214–224. doi:10.1016/j.tourman.2004.09.004

Tseng, F.-M., Yu, H.-Ch., & Tseng, G.-H. (2002). Combining Neural Network Model with Seasonal Time Series ARIMA Model. *Technological Forecasting and Social Change*, *69*(1), 71–87. doi:10.1016/S0040-1625(00)00113-X

Wang, C. H. (2004). Predicting Tourism Demand using Fuzzy Time-Series and Hybrid Grey Theory. *Tourism Management*, *25*(3), 367–374. doi:10.1016/S0261-5177(03)00132-8

Yager, R. R., & Filev, D. P. (1994). Generation of fuzzy rules by mountain clustering. *Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology*, *2*(3), 209–219. doi:10.3233/IFS-1994-2301

Stefanos K. Goumas was born in Athens, Greece, in 1960. He received the B.Sc. degree in mathematics from the National and Kapodistrian University of Athens, in 1984 and the B.Sc. degree in computer science from the University of Crete, Greece, in 1994. Also he received the M.Sc. and Ph.D. degrees from the Technical University of Crete, Greece, in 1999 and 2008, respectively, all in electronic and computer engineering. After holding a visiting Assistant Professor position with the Technological Educational Institute of Crete, from 1991 to 2003, he joined the Technological Educational Institute of Eastern Macedonian and Trace in 2003 where he is currently an Associate Professor with the Business Administration Department. Professor Goumas has been participant investigator in 8 research projects, funded both publicly and privately, in the European Union. He has authored or co-authored over 40 publications in scientific journals and referred conference proceedings. His research interests include automated inspection systems, signal and image processing and analysis, wavelet analysis for data/image processing, time series forecasting and compression, neural networks, fuzzy logic and genetic algorithms in pattern recognition applications, multimedia databases, geographical information systems and educational technology. Professor Goumas is a member of the Greek Computer Society.

Stavros Kontakos is Mathematician, with MSc in Statistics and Operational Research. His PhD is on Applied Statistics and nowadays is a post doctoral student in Democritus University of Thrace. His main field of research is Structural Equation modeling, Time series analysis, and Multivariate Techniques.

Aikaterini Mathheaki was born in Athens, Greece, in 1963. Sh received the B.Sc. degree in economics from the National and Kapodistrian University of Athens, in 1998. Also she received the M.Sc. degree from the Hellenic Open University, Greece, in 2012, in tourism management. From 2003, she joined the Technological Educational Institute of Eastern Macedonian and Trace where she is currently a visiting Assistant Professor Department. Professor Mathheaki has been participant investigator in 2 research projects, funded both publicly and privately, in the European Union. His research interests tourism management, time series forecasting and educational technology. Professor Mathheaki is a member of the Economic Chamber of Greece.