Evaluating the Performance of Linear and Nonlinear Models in Forecasting Tourist Occupancy in the Region of Western Greece

Athanasios Koutras, Alkiviadis Panagopoulos, and Ioannis A. Nikas

Abstract Accurate tourism demand forecasting systems are very important in tourism planning, especially in high tourist countries and regions within. In this paper we investigate the problem of accurate tourism demand prediction using nonlinear regression techniques based on Artificial Neural Networks (ANN). The relative accuracy of the Multilayer Perceptron (MLP) and Support Vector regression (SVR) in tourist occupancy data is investigated and compared to simple Linear Regression (LR) models. The relative performance of the MLP and SVR models is also compared to each other. For this, the data collected for a period of 8 years (2005–2012) showing tourism occupancy of the hotels of the Western Region of Greece is used. Extensive experiments have shown that the SVM regressor with the RBF kernel (SVR-RBF) outperforms the other forecasting models when tested for a wide range of forecast horizon (1–24 months) presenting very small and stable prediction error compared to SVR-POLY, MLP, as well as the simple LR models.

Keywords Support Vector Regression • Multilayer Perceptron • Artificial Neural Networks • Tourism demand forecasting • Forecasting model • Western Greece tourism • Time-series

1 Introduction

Nowadays, being a traveler means, in a few words, a consumer of tourist product. The tendency for expending in such products, as, e.g., airplane tickets, food and beverage services, overnight stays and museum visits emerges by the need of modern man to travel along the world for business, educational and entertainment reasons.

A. Koutras (✉)
Department of Informatics and Mass Media, Technical Educational Institute of Western Greece, Patras, Greece
e-mail: koutras@teipat.gr

A. Panagopoulos • I.A. Nikas
Department of Tourism Management, Technical Educational Institute of Western Greece, Patras, Greece

V. Katsoni, A. Stratigea (eds.), *Tourism and Culture in the Age of Innovation*, Springer Proceedings in Business and Economics,
DOI 10.1007/978-3-319-27528-4_26

Nevertheless, the nature of tourism, as a dynamic system, depends strongly in many unequable and uncertain characteristics delivering a changeable and perishable product.

It is clear now, more than never, the importance of tourism industry globally, let alone in small countries with a significant percentage of their revenue be coming from tourism, like Greece. Moreover, new technologies gave the opportunity to many local economies to promote their local products globally, expecting a bigger share not only from the local but, also, from the global tourist market.

At the opposite side, the travelers, the consumers of tourism product, seem to develop a more perceptive personality in choosing their destinations; their consumption decisions are turned into less predictable and more spontaneous, driving their choices mostly from the need and desire for new experiences (Burger, Dohnal, Kathrada, & Law, 2001).

So, the most important issue arising from all the above is how we can exploit the present experience in order to make better decisions for tomorrow; in the tourism industry this is a crucial step for success. Evidently, any undisposed tourist "merchandise" cannot be stocked in order to be offered again in a next season in the tourist market. For example, empty rooms or unsold airplane tickets consist in lost revenue and probably in a strong indication of a bad planning. Hence, it is a great necessity for the tourist industry to have an a priori knowledge for the expected tourist arrivals to be able to schedule the flights, the hotel and room availability, the necessary employees etc. Unfortunately we are not able to know the future, but we can adapt forecasting processes to predict the behavior of future events (Franses, 2004; Frees, 1996; Makridakis & Hibon, 1979).

The importance of developing reliable and accurate forecasting models is a crucial step for the decision makers. What really matters is the knowledge of the size, directions and characteristics of future international tourist flows (Shahrabi, Hadavandi, & Asadi, 2013). Accurate forecasting models in both short- and long-term periods are considerable important for the effective formulation and implementation of tourism strategies (Song, Gao, & Lin, 2013) in various tourist organizations and business, in both public and private sector. Actually, accurate and reliable forecasting models are the key to the success of the whole tourism industry (Gunter & Önder, 2015).

In this work a forecasting model for the tourist occupancy in the area of Western Greece is presented. The Western Greece region consists of three dissimilar prefectures (Achaia, Ilia, Aitoloakarnania) regarding the type of the visiting tourists, the available resources and infrastructures and the level of development and employment (Panagopoulos & Panagopoulos, 2005). However, despite the heterogeneous geographic morphology and economic activity, the overall region retains the same characteristics of a tourist destination, that is, the suggestibility in various exogenous factors as well as the considerable contribution to the local and country economy.

The basic reason, motivating this work, is from the one hand the finance crisis in Greece and a question about the viability of local tourist industry and from the other hand the poor research made in this area about the future and the perspectives of tourism. The only known work for the area of Western Greece is the paper of Panagopoulos and Panagopoulos, who proposed a forecasting model for predicting the tourist occupancy in the West Greece area using the Box-Jenkins Method (Box & Jenkins, 1976) and monthly data from January 1990 to December 1999, forecasting for 2 years (Panagopoulos & Panagopoulos, 2005). Hence, the study of the West Greece region constitutes a strong research motivation and any suggestions in the direction of modeling the overall local tourist product circulation remains a well-timed issue for both researchers and local authorities.

Generally, for the problem of forecasting 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 & Önder, 2015). The most widely used technique in this framework is the (Seasonal) Autoregressive (Integrated) Moving Average models (Box & Jenkins, 1976). Recently, some new, well performed, time-series models have been proposed such as the Exponential Smoothing models (Hyndman, Koehler, Ord, & Snyder, 2008; Hyndman, Koehler, Snyder, & Grose, 2002), and a low cost inferential model (Psillakis, Panagopoulos, & Kanellopoulos, 2009); multivariate or Econometric models are also employed, such as Autoregressive Distributed Lag Models (Dritsakis & Athanasiadis, 2000; Ismail, Iverson, & Cai, 2000), Error Correction Models (Kulendran & Witt, 2003; Roselló, Font, & Roselló, 2004), Vector Autoregressive models (Shan & Wilson, 2001; Witt, Song, & Wanhill, 2004) and Time-Varying Parameter models (Li, Song, & Witt, 2006; Song & Witt, 2006); some artificial intelligence methods were, also, used (Chena & Wang, 2007; Claveria & Torra, 2014; Hernández-López & Cáceres-Hernández, 2007; Kon & Turner, 2005; Palmer, Montaño, & Sesé, 2006). An exhaustive review on forecasting time series can be found in (Song & Li, 2008).

According to the review work of Song and Li (Song & Li, 2008) on tourism demand modeling and forecasting, there exists no single model that can be used in all situations in terms of performance and forecasting accuracy. Furthermore, Coshall and Charlesworth (2010) report that tourism demand forecasting can be achieved with causal econometric models (ECM models, VAR models, LAIDS models etc), and non-casual time series models. In the latter case, the most widely used techniques are the ARIMA (Goh & Law, 2002) and the Exponential Smoothing (ES) models (Cho, 2003). However, during the last years, Artificial Neural Networks (ANN) have made their appearance into solving the tourism forecasting problem (Kon & Turner, 2005; Palmer et al., 2006).

The increasing interest in more advanced prediction models, together with the fact that tourism is a leading industry worldwide, contributing to a significant proportion of world production and employment, has lead us to evaluate the forecasting performance of the most significant ANNs. In this work we have used different forecasting horizons and compare the performance of the different ANNs architectures on the prediction problem of tourism demand as it is described by the occupancy of hotels in the Region of Western Greece.

The purpose of the paper is to evaluate the forecasting performance of the Multilayer Perceptron (MLP), and the Support Vector Regressor based on polynomial (SVR-POLY) as well as Radial Basis Functions Kernels (SVR-RBF), two of the most widely known network architectures in the literature. The Support Vector Regressor network uses a structural risk minimization principle that attempts to minimize the upper bounds of the generalization error rather than minimizing the training error as classic neural networks do (Vapnik, Golowich, & Smola, 1996). The generalization error is defined as the expected value of the square of the difference between the learned function and the exact target (mean square error), while the training error is calculated as the average loss over the training data. In this work we have used official statistical monthly data of the hotels occupancy in the Western Greece Region from 2005 until 2012 taken from the official records of the Hellenic Statistical Authority. Then the MAPE and the RSME is computed for different forecast horizons ranging from 1 to 24 months (2 years prediction).

The structure of the paper is as follows: In the next section we briefly present the theoretical background of the utilized neural network forecast models. The experimental setup as well as the data set is described in Sect. 3. In Sect. 4 the results of the forecasting are presented and discussed and finally in Sect. 5 some conclusions and remarks are given.

2 Methodology

2.1 Multilayer Perceptron Regressor

An Artificial Neural Network (ANN) is a non-linear black box statistical approach. The most commonly used ANN structure is the feed-forward multilayer perceptron (MLP). This structure is composed of at least three layers: an input layer, one or more hidden layers and an output layer. The network consists of a set of neurons connected by links and normally organized in a number of layers. The number of neurons in the input and output layer is equal to the number of input and output variables respectively. The number of neurons in the hidden layer(s) is usually selected by trial-and-error. The output of this network can be calculated by the following equation:

$$Y_j = f\left(\sum_i w_{ij} X_{ij}\right), \tag{1}$$

where Y_j is the output of node j, $f(.)$ is the transfer function of the network, w_{ij} are the connection weights of the network that need to be estimated between nodes j and i and X_i is the input. The MLP uses the well-known Back Propagation learning algorithm to estimate adaptively the values of the network's weights. In order to do this, it minimizes the square error between the calculated and the desired network's output based on a steepest descend technique with the addition of a momentum weight/bias function, which calculates the weight change for any given neuron at each iteration step. By considering that the prediction error is given by the following equation

$$E = \frac{1}{2} \sum_{p} \sum_{j} \left[O_j^p - Y_j^p \right]^2, \tag{2}$$

the adaptation rule for estimating the values of the weights is given by:

$$\Delta w_{ij}^p(n) = -n \frac{\partial E(n)}{\partial w_{ij}^p} \tag{3}$$

The above equation after applying the chain rule of differentiation leads to the following rule

$$\begin{aligned} \Delta w_{ij}^p(n) &= \mu e_j^p(n) X_i^{p-1}(n) + m \Delta w_{ij}^p(n-1) \\ \Delta w_{ij}^p(n+1) &= w_{ij}^p(n) + \Delta w_{ij}^p(n), \end{aligned} \tag{4}$$

where $e_j^p(n)$ is the nth error signal at the jth neuron in the pth layer, $X_i^{p-1}(n)$ is the output signal of neuron i at the layer below, μ is the learning rate, and m is the momentum factor. The last two parameters are specified at the start of the training procedure and affect the speed and stability of the convergence of the steepest descend algorithm.

In brief, the procedure to set-up a MLP neural network to solve the regression problem is:

(1) Select the number of the input data points and define the input layer.
(2) Select the number of the output points and define the output layer.
(3) Determine the number of the hidden layers as well as the number of the nodes in each layer. There is no rule for this task; this may depend on trial and error.
(4) Perform learning from a set of known data. This step results in estimating the weights of the connections between the nodes of all layers of the network.
(5) Test the neural network using known data that have never been presented to the network in step (4). In this way we can measure the accuracy as well as the efficacy of the network using various metrics (mean square error, mean absolute percentage error etc).

2.2 Support Vector Regression

The support vector regression (SVR) is a recent adaptation of the classification scheme based on support vector machines. The general regression problem can be formulated as follows: Consider a set of data points $D = \{(\boldsymbol{x}_i, q_i)\}_{i=1}^{n}$, where $\boldsymbol{x}_i$ is a vector of model inputs, q_i is the actual value that is a scalar and n the total number of data patterns. The purpose of the regressor is to estimate a function $f(\boldsymbol{x})$ that can predict the desired values q_i given a set of input samples.

A regression function is given in the form of $q_i = f(\boldsymbol{x}_i) + \delta$, where δ is the error that follows the normal distribution. Support Vector regression deals with the most general and difficult non-linear regression problem. In order to solve the non-linear regression problem, the SVR maps non-linearly the inputs into a high dimensional space where they are linearly correlated with the outputs. This is described by:

$$f(\boldsymbol{x}) = \left(\boldsymbol{v} \cdot \Phi(\boldsymbol{x})\right) + b, \quad (5)$$

where $\boldsymbol{v}$ is a weight vector, b is a constant, $\Phi(\boldsymbol{x})$ denotes the non-linear function. So, in SVR, the problem of nonlinear regression in the lower dimension space is transformed into an easier linear regression problem in a higher dimension feature space.

For solving this problem the most commonly used cost function is:

$$L_\varepsilon(f(\boldsymbol{x}), q) = \begin{cases} |f(\boldsymbol{x}) - q| - \varepsilon, \; if \; |f(\boldsymbol{x}) - q| \geq \varepsilon \\ 0, \; otherwise \end{cases}, \quad (6)$$

where ε is the precision parameter that represents the radius of the tube located around the regression function $f(\boldsymbol{x})$ and t is the target value.

The weight vector $\boldsymbol{v}$ as well as the constant b can be estimated by minimizing the following risk function:

$$R(C) = C\frac{1}{n}\sum_{i=1}^{n} L_e(f(\boldsymbol{x}_i), q_i) + \frac{1}{2}|\boldsymbol{w}|^2 \quad (7)$$

where $L_e(f(\boldsymbol{x}), q_i)$ is the loss function, $\frac{1}{2}|\boldsymbol{w}|^2$ is the regularization term which controls the trade-off between the complexity and the approximation accuracy of the model, C is the regularization constant. Both C and ε are determined by the user in a trial and error manner.

By using slack variable ξ_ι and ξ_ι^*, the previous equation is transformed into the constrained form:

minimize:

$$R_{reg}(f) = \frac{1}{2}|\boldsymbol{w}|^2 + C\sum_{\iota=1}^{n} \left(\xi_\iota + \xi_\iota^*\right) \quad (8)$$

subject to:

$$\begin{cases} q_i - \left(\boldsymbol{w} \cdot \boldsymbol{\Phi}(\boldsymbol{x_i})\right) - b \leq \varepsilon + \xi_\iota \\ \left(\boldsymbol{w} \cdot \boldsymbol{\Phi}(\boldsymbol{x_i})\right) + b - q_i \leq \varepsilon + \xi_\iota{}^* \\ \xi_\iota, \xi_\iota{}^* \geq 0, i = 1 \ldots n \end{cases} \tag{9}$$

By using Laplace multipliers and the Karush-Kuhn-Tucker conditions to the equation, it results to the following dual Lagrangian form,

maximize:

$$\begin{aligned} L_d(a, a^*) = &-\varepsilon \sum_{i=1}^{n} (a_i^* + a_i) + \sum_{i=1}^{n} (a_i^* - a_i) q_i \\ &- \frac{1}{2} \sum_{i=1}^{n} (a_i^* - a_i)(a_j^* - a_j) \boldsymbol{K}(\boldsymbol{x_i}, \boldsymbol{x_j}), \end{aligned} \tag{10}$$

subject to the constrains,

$$\begin{cases} \sum_{i=1}^{n} \left(a_i^* - a_i\right) = 0 \\ 0 \leq a_i \leq C, i = 1 \ldots n \\ 0 \leq a_i^* \leq C, i = 1 \ldots n \end{cases} \tag{11}$$

The Lagrange multipliers satisfy the equality $a_i^* a_i = 0$. The Lagrange multipliers, a_i^*, a_i are calculated and an optimal desired weight vector of the regression hyperplane is

$$\boldsymbol{v}^* = \sum_{i=1}^{n} \left(a_i - a_i^*\right) \boldsymbol{K}\left(\boldsymbol{x_i}, \boldsymbol{x_j}\right). \tag{12}$$

Hence the general form of the regression function can be written as

$$f(\boldsymbol{x}, \boldsymbol{v}) = f\left(\boldsymbol{x}, a_i, a_i^*\right) = \sum_{i=1}^{n} \left(a_i - a_i^*\right) \boldsymbol{K}\left(\boldsymbol{x_i}, \boldsymbol{x_j}\right) + b \tag{13}$$

where $\boldsymbol{K}(\boldsymbol{x_i}, \boldsymbol{x_j})$ is the kernel function. The values of the kernel function equals the inner product of the vectors $\boldsymbol{x}_i$, $\boldsymbol{x}_j$ in the feature space $\boldsymbol{\Phi}(\boldsymbol{x_i})$, $\boldsymbol{\Phi}(\boldsymbol{x_j})$.

Several choices for the kernel function exist, the two most widely known and used in the literature are the radial basis function (SVR-RBF) defined as $\boldsymbol{K}\left(\boldsymbol{x_i}, \boldsymbol{x_j}\right) = exp\left(\frac{-\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2}{2\sigma^2}\right)$ and the polynomial kernel (SVR-POLY) function defined as $\boldsymbol{K}\left(\boldsymbol{x_i}, \boldsymbol{x_j}\right) = \left(\boldsymbol{x}_i^T, \boldsymbol{x}_j + c\right)^d$, with d the degree of the polynomial.

3 Experimental Setup

3.1 Multilayer Perceptron Regressor

The main objective in designing the MLP model's architecture is to find the optimal architecture that will model the relationship between input and output (forecasted) values. The number of neurons in the input layer equals the number of the dimensionality of the input data, while the number of neurons in the output layer is equal to the number of the output data. In forecasting the tourist occupancy of the hotels in the Region of Western Greece, we have used a 12-dimensional input vector that holds the occupancy values in a history window of a year (12 months). The number of the neurons in the hidden layer is selected by trial-and-error procedure. In this paper, we have tested the efficacy of the MLP network using a wide range of neurons in the only hidden layer from 1 to 50. Experiments have shown that the MLP network performs best (smaller error) with 22 nodes in the hidden layer. For training this type of neural network we have used the Levenberg-Marquardt algorithm.

3.2 Support Vector Regressor

The first step in using the SVR is the selection of the Kernel function. In this paper we have tested the forecasting performance for the polynomial (SVR-POLY) as well as the RBF (SVR-RBF) kernel function presented in the previous section. The performance of the proposed RBF regressor depends on the values of the kernel function parameters. Thus the selection of three parameters, regularization constant C, loss function ε and σ (the width of the RBF) of a SVR-RBF regressor, as well as the selection of the regularization constant C, loss function ε and d (the degree of the polynomial) of a SVR-POLY regressor is crucial for accurate forecasting. As there exists no general rule for selecting these parameters, this is usually based on the grid search method proposed by Lin, Hsu, and Chang (2003). The grid search method is a straightforward method that uses exponentially growing sequences of C and ε to estimate the best parameter values. The parameter set C, ε that generates the minimum forecasting RMSE and MAPE error is considered as the best parameter set and used throughout the experiments. In our work, we have used the (100, 0.1) for the SVR-RBF and (100, 1) for the SVR-POLY regressor.

3.3 Experimental Dataset

For evaluating the performance of the utilized forecasting methods, the occupancy of all tourist accommodations (except from camping sites) in the Region of Western

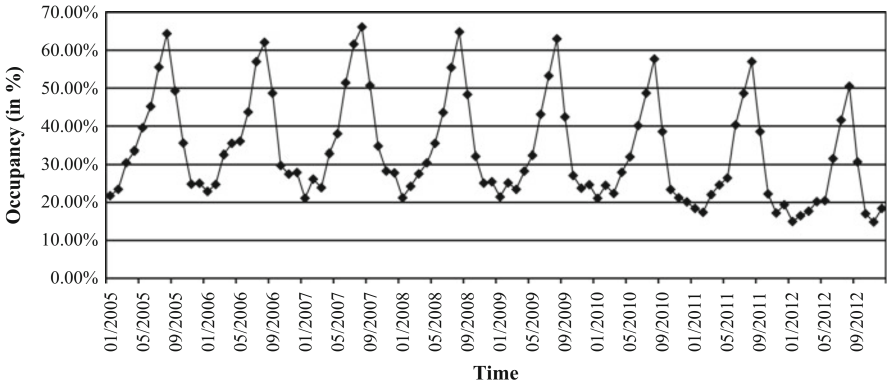

Fig. 1 Monthly occupancies of all tourist accommodations (except from camping sites) from 2005 (1) to 2012 (12)

Greece that includes data from the Prefectures of Aitoloakarnania, Achaia and Hlia from January of 2005 till December 2012 was used. 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. There are a total of 96 data points in the dataset and the monthly occupancy series is plotted in Fig. 1. The plot exhibits a long-term of downward trend as well as a strong seasonality of 12 months with the maxima of the occupancy occurring during the high touristic summer season (maximum in August for every year).

In order to test the performance of the proposed method, the collected data is divided into two sets, training data and testing data. In order to further test the efficacy of the linear and non-linear prediction methods, we have calculated the prediction accuracy with a prediction step ranging from 1 to 24 months (2 years).

3.4 Performance Criteria

According to Tay and Cao (2001) and Thomason (1999), the prediction performance of our method is evaluated using measures including mean absolute percentage error (MAPE), and root mean square error (RMSE). MAPE and RMSE were used to measure the correctness of the prediction in terms of levels and the deviation between the actual and predicted values. The smaller the values, the closer the predicted values are to the actual values.

Table 1 Performance indices and their calculations

Metrics	Calculation
MAPE*	$\frac{1}{n}\sum_{i=1}^{n}\left\lvert\frac{P_i - A_i}{A_i}\right\rvert$
RMSE	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - A_i)^2}$

*A_i and P_i present the actual and the predicted values

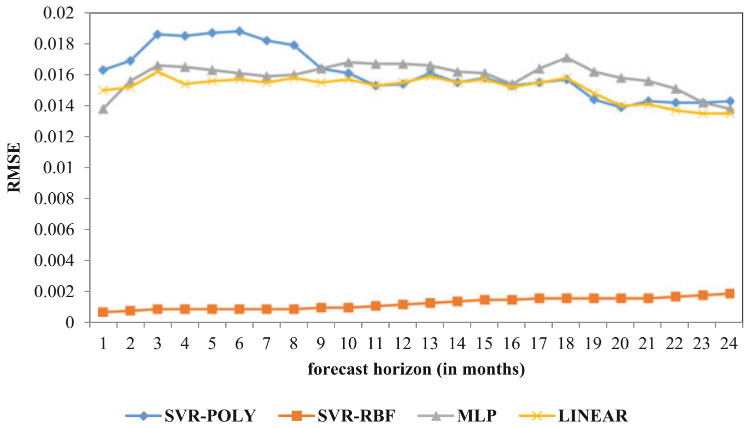

Fig. 2 The RMSE error for various values of forecast horizon (1–24 months, for years 2011–2012)

4 Experimental Results

The proposed method's performance was tested by using the first 72 data points (72 months—6 years from 2005 to 2010) for training purposes and the remaining 24 data points (24 months—2 years, 2011–2012) were forecasted using the SVR-POLY, SVR-RBF, MLP as well as the LP forecasting networks. The performance of these methods was compared using the prediction error measurements (MAPE & RMSE) presented in Table 1. The errors were estimated for different values of prediction horizon ranging from 1 to 24 months for all regressors and the results are presented in Figs. 2 and 3. From these Figures it is clear that the SVR-RBF shows the best performance as it is clear that it can forecast accurately and robustly with very small forecasting error, almost constant that does not depend on the forecast's time horizon (1–24 months) compared to the other three models.

In Fig. 4, the forecasting of the 2011–2012 years is presented for the four aforementioned methods as well as the actual monthly data. It is clear that the SVR-RBF network works well and significantly better than the other three techniques and manages to predict the actual occupancy values more accurately. For comparison reasons, in Table 2 we present analytically the forecasted occupancy

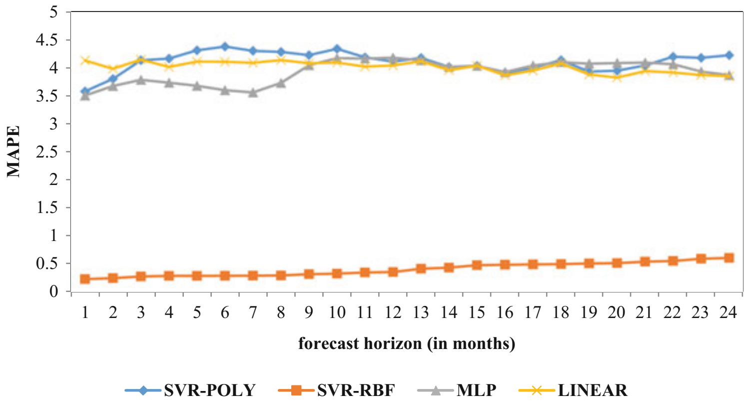

Fig. 3 The MAPE error for various values of forecast horizon (1–24 months, for years 2011–2012)

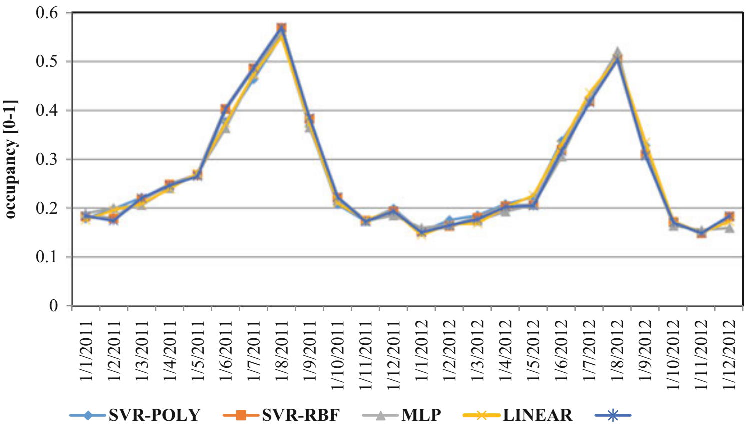

Fig. 4 Occupancy prediction for the years 2011–2012

values for the years 2011–2012 estimated by the SVR-RBF, SVR-POLY, MLP, LR, together with the real values for this period of time.

Additionally, all four forecasting models were further tested to predict the next 2 year's hotel occupancy for the years 2013–2014. The results are shown in Fig. 5, while in Table 3 we present analytically the forecasted occupancy values for this time period.

Table 2 Predicted occupancy values for the years 2011–2012 (24 month prediction)

	SVR-POLY	SVR-RBF	MLP	LR	REAL
1/1/2011	0.1783	0.1826	0.1888	0.176	0.184
1/2/2011	0.1983	0.1783	0.2002	0.1955	0.174
1/3/2011	0.2215	0.2175	0.2061	0.2085	0.2207
1/4/2011	0.2435	0.2479	0.2408	0.2393	0.2463
1/5/2011	0.2669	0.2676	0.2707	0.2715	0.2647
1/6/2011	0.3767	0.4026	0.3636	0.3736	0.403
1/7/2011	0.4637	0.4858	0.474	0.4731	0.4863
1/8/2011	0.5524	0.5689	0.5626	0.5501	0.5697
1/9/2011	0.3676	0.3824	0.365	0.3729	0.385
1/10/2011	0.2074	0.2212	0.2156	0.2101	0.2227
1/11/2011	0.1733	0.1741	0.1735	0.1762	0.1723
1/12/2011	0.1989	0.1935	0.1848	0.1938	0.1937
1/1/2012	0.1517	0.1522	0.1595	0.1445	0.15
1/2/2012	0.1755	0.163	0.1655	0.1668	0.165
1/3/2012	0.1851	0.18	0.1739	0.1682	0.177
1/4/2012	0.2082	0.204	0.193	0.202	0.2023
1/5/2012	0.2223	0.206	0.2064	0.2258	0.2047
1/6/2012	0.3372	0.3195	0.3055	0.3305	0.3157
1/7/2012	0.4238	0.4186	0.4247	0.436	0.416
1/8/2012	0.5047	0.5054	0.5215	0.5069	0.5047
1/9/2012	0.3283	0.3094	0.3077	0.3346	0.307
1/10/2012	0.1679	0.1716	0.1637	0.1721	0.1703
1/11/2012	0.149	0.1483	0.1549	0.1498	0.1487
1/12/2012	0.1825	0.1823	0.1595	0.1725	0.184

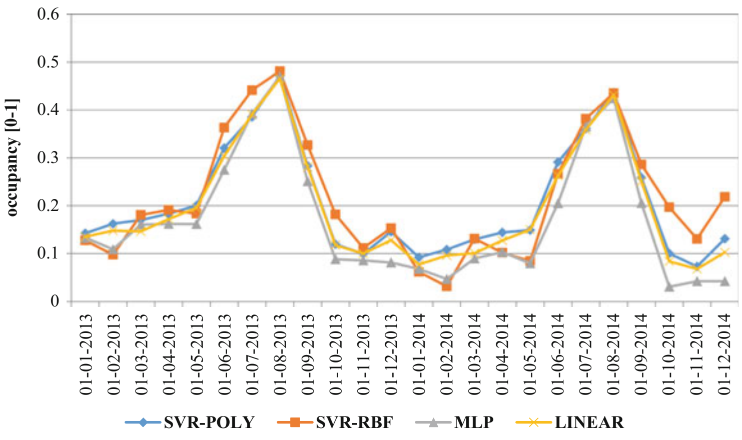

Fig. 5 Occupancy prediction for the years 2013–2014

Table 3 Predicted occupancy values for the years 2013–2014 (24 month prediction)

	SVR-POLY	SVR-RBF	MLP	LR
01-01-2013	0.1424	0.1277	0.1333	0.1353
01-02-2013	0.1623	0.0974	0.109	0.148
01-03-2013	0.1702	0.1807	0.1606	0.1464
01-04-2013	0.1835	0.1912	0.1618	0.1719
01-05-2013	0.2004	0.1834	0.1616	0.1956
01-06-2013	0.3206	0.3632	0.2748	0.3039
01-07-2013	0.3855	0.4414	0.3922	0.3905
01-08-2013	0.4715	0.4812	0.4701	0.4654
01-09-2013	0.2833	0.3266	0.2511	0.281
01-10-2013	0.1195	0.1818	0.0882	0.1174
01-11-2013	0.0994	0.1118	0.0858	0.0999
01-12-2013	0.1456	0.1531	0.0812	0.1282
01-01-2014	0.0921	0.0618	0.068	0.0774
01-02-2014	0.1084	0.0315	0.0469	0.0959
01-03-2014	0.1296	0.1313	0.0898	0.1008
01-04-2014	0.1441	0.102	0.103	0.1275
01-05-2014	0.149	0.0847	0.0794	0.1501
01-06-2014	0.2913	0.2666	0.2051	0.2624
01-07-2014	0.3613	0.3818	0.3653	0.3586
01-08-2014	0.4362	0.4352	0.4244	0.4311
01-09-2014	0.2585	0.286	0.2057	0.2507
01-10-2014	0.1002	0.1969	0.0307	0.0837
01-11-2014	0.0731	0.1306	0.0422	0.0679
01-12-2014	0.131	0.2186	0.0422	0.1025

5 Conclusion

In this paper we have presented an evaluation of the forecasting performance of two of the most widely known Artificial Neural Networks ANN architectures, the Multilayer Perceptron (MLP), and the Support Vector Regressor based on polynomial (SVR-POLY) as well as Radial Basis Functions Kernels (SVR-RBF). For our experiments we have used official statistical monthly data of the hotels occupancy in the Western Greece Region from 2005 until 2012 taken from the official records of the Hellenic Statistical Authority. Then the MAPE and the RSME is computed for different forecast horizons ranging from 1 to 24 months (2 years prediction, 2011–2012). The SVR that uses RBF kernels provides better performance than the other artificial neural networks architectures tested in this paper, when applied to solve the tourism forecasting problem. In the tourism industry, tourism service providers should assess the costs and benefits of each model before choosing one for forecasting. This has significant managerial implications when it comes to constructing a strategic plan for marketing. With the accurate forecasted trends and patterns that indicate the sizes of tourist demand, the government and private sectors can have a well-organized tourism strategy and provide a better infrastructure to serve the visitors and develop a suitable marketing strategy to gain benefit from the growing tourism (Shahrabi et al., 2013). Moreover, armed with accurate estimates of demand for tourism, tourism authorities and decision makers in the hospitality industries would be better able to perform strategic planning.

Acknowledgements 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).

References

Box, G., & Jenkins, G. (1976). *Time series analysis: Forecasting and control* (2nd ed.). San Francisco: Holden-Day.

Burger, C. J. S. C., Dohnal, M., Kathrada, M., & Law, R. (2001). A practitioners guide to time-series methods for tourism demand forecasting—A case study of Durban, South Africa. *Tourism Management, 22*, 403–409.

Chena, K.-Y., & Wang, C.-H. (2007). Support vector regression with genetic algorithms in forecasting tourism demand. *Tourism Management, 28*, 215–226.

Cho, V. (2003). A comparison of three different approaches to tourist arrival forecasting. *Tourism Management, 24*, 323–330.

Claveria, O., & Torra, S. (2014). Forecasting tourism demand to Catalonia: Neural networks vs. time series models. *Economic Modelling, 36*, 220–228.

Coshall, J. T., & Charlesworth, R. (2010). A management orientated approach to combination forecasting of tourism demand. *Tourism Management, 32*, 759–769.

Dritsakis, N., & Athanasiadis, S. (2000). An econometric model of tourist demand: The case of Greece. *Journal of Hospitality and Leisure Marketing, 7*, 39–49.

Franses, P. H. (2004). *Time series models for business and economic forecasting*. Cambridge: Cambridge University Press.

Frees, E. W. (1996). *Data analysis using regression models—The business perspective*. New York: Prentice Hall.

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

Gunter, U., & Önder, I. (2015). Forecasting international city tourism demand for Paris: Accuracy of uni- and multivariate models employing monthly data. *Tourism Management, 46*, 123–135.

Hernández-López, M., & Cáceres-Hernández, J. J. (2007). Forecasting tourists' characteristics by a genetic algorithm with a transition matrix. *Tourism Management, 28*, 290–297.

Hyndman, R. J., Koehler, A. B., Ord, J. K., & Snyder, R. D. (2008). *Forecasting with exponential smoothing: The state space approach*. Berlin: Springer.

Hyndman, R. J., Koehler, A. B., Snyder, R. D., & Grose, S. (2002). A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting, 18*, 439–454.

Ismail, J. A., Iverson, T. J., & Cai, L. A. (2000). Forecasting Japanese arrivals to Guam: An empirical model. *Journal of Hospitality and Leisure Marketing, 7*, 51–63.

Kon, S. C., & Turner, W. L. (2005). Neural network forecasting of tourism demand. *Tourism Economics, 11*, 301–328.

Kulendran, N., & Witt, S. F. (2003). Forecasting the demand for international business tourism. *Journal of Travel Research, 41*, 265–271.

Li, G., Song, H., & Witt, S. F. (2006). Forecasting tourism demand using econometric models. In D. Buhalis & C. Costa (Eds.), *Tourism management dynamics: Trends, management and tools* (pp. 219–228). Oxford: Elsevier.

Lin, C. J., Hsu, C. W., & Chang, C. C. (2003). A practical guide to support vector classification. Technical Report, Department of Computer Science and Information Engineering, National Taiwan University, Taipei.

Makridakis, S., & Hibon, M. (1979). Accuracy of forecasting: An empirical investigation. *Journal of the Royal Statistical Society A, 142*, 97–145.

Palmer, A., Montaño, J. J., & Sesé, A. (2006). Designing an artificial neural network for forecasting tourism time-series. *Tourism Management, 27*, 781–790.

Panagopoulos, A., & Panagopoulos, A. (2005). A time series method for occupancy forecasting—A case study of West Greece. *Archives of Economic History, 17*(1), 67–78.

Psillakis, Z., Panagopoulos, A., & Kanellopoulos, D. (2009). Low cost inferential forecasting and tourism demand in accommodation industry. *Tourismos, 4*(2), 47–68.

Roselló, J., Font, A. R., & Roselló, A. S. (2004). The economic determinants of seasonal patterns. *Annals of Tourism Research, 31*, 697–711.

Shahrabi, J., Hadavandi, E., & Asadi, S. (2013). Developing a hybrid intelligent model for forecasting problems: Case study of tourism demand time series. *Knowledge-Based Systems, 43*, 112–122.

Shan, J., & Wilson, K. (2001). Causality between trade and tourism: Empirical evidence from China. *Applied Economics Letters, 8*, 279–283.

Song, H., Gao, B. Z., & Lin, V. S. (2013). Combining statistical and judgmental forecasts via a web-based tourism demand forecasting system. *International Journal of Forecasting, 29*, 295–310.

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. (2006). Forecasting international tourist flows to Macau. *Tourism Management, 27*, 214–224.

Tay, F. E. H., & Cao, L. (2001). Application of support vector machines in financial time series forecasting. *Omega, 29*, 309–317.

Thomason, M. (1999). The practitioner method and tools: A basic neural network based trading system project revisited (parts 1 and 2). *Journal of Computational Intelligence in Finance, 7* (3), 36–45.

Vapnik, V., Golowich, S., & Smola, A. (1996). Support vector machine for function approximation, regression estimation, and signal processing. *Advances in Neural Information Processing Systems, 9*, 281–287.

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