Forecasting Tourism Demand Using Linear and Nonlinear Prediction Models

Athanasios Koutras
Technological Educational Institute of Western Greece, Department of Informatics and Mass Media, Greece
koutras@teiwest.gr

Alkiviadis Panagopoulos
Technological Educational Institute of Western Greece, Department of Tourism Management, Greece
panagopa@teiwest.gr

Ioannis A. Nikas
Technological Educational Institute of Western Greece, Department of Tourism Management, Greece
nikas@teiwest.gr

In this paper, we propose and evaluate linear and nonlinear prediction models based on Artificial Neural Networks (ANN) for tourism demand in the accommodation industry. For efficient forecasting, the Multilayer Perceptron (MLP), Support Vector Regression (SVR) and Linear Regression (LR) methods that utilize two different feature sets for training have been used. The major contribution of the proposed models is focused mainly on better forecasting accuracy and lower cost effort. The relative accuracy of the Multilayer Perceptron (MLP) and Support Vector Regression (SVR) in tourism occupancy data is investigated and compared to simple Linear Regression (LR) models. The relative performance of the MLP and SVR models are also compared to each other. Data collected over a period of eight years (2005–2012) showing tourism occupancy and the number of overnight stays in the hotels of the Western Region of Greece is used. Extensive experiments have shown that for time series describing a subset of the number of overnight stays, the SVR regressor with the RBF kernel (SVR-RBF), as well as simple LR models, and the MLP regressor for occupancy time series respectively, outperform other forecasting models, when tested for a wide range of forecast horizons (1–24 months) and present very small and stable prediction errors.

Keywords: support vector regression, multilayer perceptron, artificial neural networks, tourism demand forecasting, forecasting model, time-series

Introduction

An essential factor in the tourist industry is travellers, which in short can be defined as consumers of tourist product. Travelers, usually, spend money on products such as air services, food and beverage services, overnight stays, visits to various places of interest, and similar. This consumption is the result of travelling the world for business, educational and entertainment reasons. In addition, travellers, the consumers of tourism products, seem to develop a more perceptive personality in choosing their destinations; their consumption decisions become less predictable and more spontaneous, motivated mostly by their need and desire for new experiences (Burger, Dohnal, Kathrada, & Law, 2001).

The above are only some of the many characteristics of the tourism industry. The tourism phenomenon is a dynamic system which depends strongly on many untenable and uncertain characteristics and delivers a changeable and perishable product. More now than ever, the important role that the tourism industry plays in the global economy is clear. This has become a great necessity for small countries with a significant percentage of their revenue coming from tourism, such as Greece. Especially with the development and exploitation of new technologies, everyone has the opportunity, especially many local economies, to promote their products globally, with low cost, expecting a bigger share not only from the local but also from the global tourist market. One of the key factors in expanding the tourist market is, somehow, to define the future tourism consumers of the tourist product.

So, how can the present experience be exploited to obtain better decisions for tomorrow? In the tourism industry, this question is a crucial step for success. Obviously, any indisposed tourist 'merchandise' cannot be stocked to be offered again in the next season in the tourist market. For example, empty rooms or unsold airplane tickets consist of lost revenue and probably are a strong indication of bad planning. Hence, it is of great necessity for the tourist industry to have an *a priori* knowledge of the expected tourist arrivals so as to be able to schedule the flights, the hotel and room availability, the necessary employees, and other factors. Although we are not able to know the future, we can adapt forecasting processes to predict the behaviour of future events (Makridakis & Hibon, 1979; Frees, 1996; Franses, 2004).

The developing of reliable and accurate forecasting models is an essential step for decision makers. What matters are 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 essential for the effective formulation and implementation of tourism strategies (Song, Gao, & Lin, 2013) in various tourist organizations and business, in both the public and private sectors. Accurate and reliable forecasting models are the key to the success of the whole tourism industry (Gunter & Önder, 2015).

Generally, for the problem of forecasting time-series, different methods and techniques have been proposed, covering a broad range of different countries and locations, as well as diverse time intervals. The most widely used models (especially using monthly data) are univariate or time-series models (Gunter & Önder, 2015) and, in this framework, the developing of such models is usually based on 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, Snyder, & Grose 2002; Hyndman, Koehler, Ord & Snyder 2008), 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 (Claveria & Torra, 2014; Palmer, Montaño & Sesé, 2006; Kon & Turner, 2005; Hernández-López & Cáceres-Hernández, 2007; Chena & Wang, 2007). An exhaustive review on forecasting time series can be found in the work of Song and Li (2008); according to their work on tourism demand modelling and forecasting, no single model that can be used in all situations in terms of performance and forecasting accuracy exists.

Furthermore, Coshall and Charlesworth (2010) report that forecasting tourism demand can be achieved with causal econometric models (ECM models, VAR models, LAIDS models, etc.), and non-casual time series models (Goh & Law, 2002; Cho, 2003). However, in recent years, Artificial Neural Networks (ANN) have made their appearance in 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 led us to evaluate the forecasting performance of the most significant ANNS. Thus, in this work, forecasting models for the tourist occupancy are presented, using 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 region consists of three dissimilar prefectures (Achaia, Ilia, and Etoloakarnania) regarding the type of the visiting tourists, the available resources and infrastructure, 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 main reason for choosing the region of Western Greece is the financial crisis in Greece and a question about the viability of the local tourist industry, as well as the lack of research made in this area concerning the future and potentials of tourism. The only known work in the literature concerning the area of Western Greece (Panagopoulos & Panagopoulos, 2005) proposes a forecasting model for predicting the tourist occupancy in the Western Greece region using the Box-Jenkins Method (Box & Jenkins, 1976) and monthly data from January 1990 to December 1999. Hence, the study of the Western Greece region constitutes a strong research motivation and any suggestions in the direction of modelling the overall local tourist product circulation remains a well-timed issue for both researchers and local authorities.

The purpose of the paper is twofold: Firstly, we evaluate the forecasting performance of linear and non-linear prediction methods using the Multilayer Perceptron Regressor (MLP), the Support Vector Regressor with polynomial kernels (SVR-POLY), and the Support Vector Regressor with Radial Basis Functions Kernels (SVR-RBF), three of the most widely known neural network architectures in the literature, as forecasting models for tourism demand. In addition, two different feature sets are proposed to train the networks, based on the extraction of the essential characteristics of a time series, such as trend and seasonality.

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 conventional neural networks do (Vapnik, Golowitch, & 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 hotel occupancy in the Western Greece Region from January 2005 until December 2012, taken from the official records of the Hellenic Statistical Authority. Next, the most commonly used metric based on the Root Mean Square Error (RMSE) is computed for different forecast horizons, ranging from 1 to 24 months (2-year 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 Section 3. In Section 4, the results of the forecasting are presented and discussed, and in Section 5 some conclusions and remarks are given.

Methodology

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 using the following equation:

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

where Y_j is the output of node j, $f(\cdot)$ 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 Y_j and the desired network's output O_j based on the 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 an 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 were not 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.).

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:

$$f(\boldsymbol{x}) = (\boldsymbol{v} \cdot \phi(\boldsymbol{x})) + b, \tag{5}$$

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:

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

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

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

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

where $L_e(f(\boldsymbol{x}), q_i)$ is the loss function, $1/2|\boldsymbol{w}|^2$ 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 by trial-and-error.

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

$$R_{reg}(f) = \frac{1}{2}|\boldsymbol{w}|^2 + C\sum_{\iota=1}^{n}(\xi_i + \xi_i^*), \tag{8}$$

subject to:

$$\begin{cases} q_i - (\boldsymbol{w} \cdot \phi(\boldsymbol{x}_i)) - b \leq \varepsilon + \xi_i \\ (\boldsymbol{w} \cdot \phi(\boldsymbol{x}_i)) + b - q_i \leq \varepsilon + \xi_i^* \\ \xi_i \xi_i^* \geq 0, i = 1, \dots, n \end{cases}. \tag{9}$$

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

$$L_d(\alpha, \alpha^*) = -\varepsilon\sum_{i=1}^{n}(a_i^* + a_i) + \sum_{i=1}^{n}(a_i^* - a_i)q_i \\ -\frac{1}{2}\sum_{i=1}^{n}(a_i^* - a_i)(a_j^* - a_j) \\ \cdot\boldsymbol{K}(\boldsymbol{x}_i, \boldsymbol{x}_j), \tag{10}$$

subject to the constraints,

$$\begin{cases} \sum_{i=1}^{n}(a_i^* - a_1) = 0 \\ 0 \leq a_i \leq C, i = 1, \dots, n \\ 0 \leq a_i^* \leq C, i = 1, \dots, n \end{cases}, \tag{11}$$

where $\boldsymbol{K}(\boldsymbol{x}_i, \boldsymbol{x}_j)$ is the kernel function.

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

$$\boldsymbol{v}^* = \sum_{i=1}^{n}(a_i - a_i^*)\boldsymbol{K}(\boldsymbol{x}_i, \boldsymbol{x}_j). \tag{12}$$

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

$$f(x, v) = f(x, a_i, a_i^*) \\ = \sum_{i=1}^{n}(a_i - a_i^*)\boldsymbol{K}(\boldsymbol{x}_i, \boldsymbol{x}_j) + b. \tag{13}$$

The values of the kernel function equal the inner product of the vectors $\boldsymbol{x}_i$, $\boldsymbol{x}_j$ in the feature space $\phi(\boldsymbol{x}_i)$, $\phi(x_j)$.

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

$$\boldsymbol{K}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(\frac{-\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2}{2\sigma^2}\right)$$

and the polynomial kernel (SVR-POLY) function defined as $\boldsymbol{K}(\boldsymbol{x}_i, \boldsymbol{x}_j) = (\boldsymbol{x}_i^T \cdot \boldsymbol{x}_j + c)^d$, with d the degree of the polynomial and $c \geq 0$ is a free parameter trading off the influence of higher-order versus lower-order terms in the polynomial.

Experimental Setup

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 tourism occupancy and overnights stays of the hotels in the Region of Western Greece, we have used two Feature Sets (FS) with different dimensionality in order to assist the predictor. The number of the neurons in the hidden layer is selected using the 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. Extensive experiments were performed for each different case in order to find the optimal number of neurons in the hidden layer that result in the best performance (smaller error) of the predictor. For training, this type of neural network we have used the Levenberg-Marquardt optimization algorithm.

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 of the polynomial (SVR-POLY) as well as the RBF (SVR-RBF) kernel function presented in the previous section. The performance of the proposed RBF regressors 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 no general rule for selecting these parameters exists, 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 error is considered as the best parameter set and used throughout the experiments.

Experimental Dataset

For evaluating the performance of the utilized forecasting methods, (a) the occupancy of all tourist accommodations (except from camping sites) and (b) the number of overnight stays in the Region of Western Greece that includes data from the Prefectures of Etoloakarnania, Achaia, and Hlia from January of 2005 to December 2012 were used. All data employed in this study were obtained from the official records of the Hellenic Statistical Authority. It must be emphasized that the Hellenic Statistical Authority has not released any similar data for the period 2013 until now.

There is a total of 96 data points in the dataset, and the monthly occupancy series and the number of overnight stays are plotted in Figures 1 and 2. Both plots exhibit a long-term 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 regressors, the collected data is divided into two sets, training data, and testing data set. 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 (forecasting horizon of 2 years).

Performance criteria

According to Tay and Cao (2001) and Thomason (1999), the prediction performance of our method is evaluated using measures of the root mean square error (RMSE). RMSE is 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 (P_i) are to the actual values (A_i).

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - A_i)^2}. \tag{14}$$

Experimental Results

The proposed method's performance was tested by using the first 72 data points (72 months; 6 years from 2005–2010) for training purposes and the remaining 24 data points (24 months; 2 years, 2011 and 2012) were forecasted using four different types of regressors: SVR-POLY, SVR-RBF, MLP as well as the LP forecasting networks. The performance of these methods was compared using the prediction error measurements (RMSE).

Occupancy Prediction

The performance error of the occupancy prediction was estimated using different values of prediction

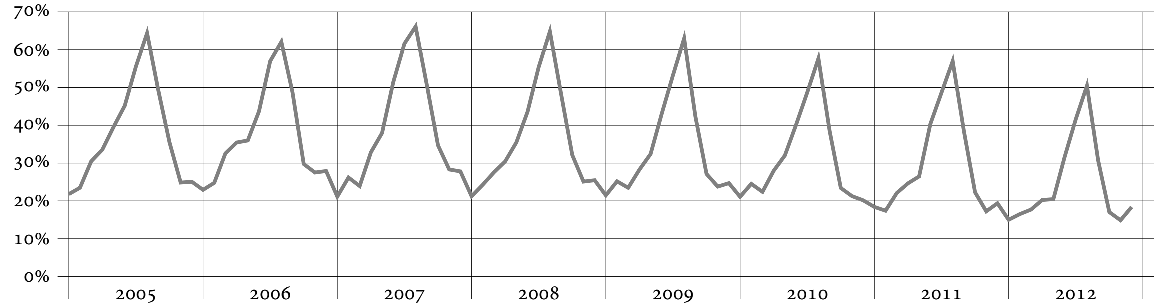


Figure 1 Monthly Occupancies of All Tourist Accommodations (Except Camping Sites) from 2005 (1) to 2012 (12)

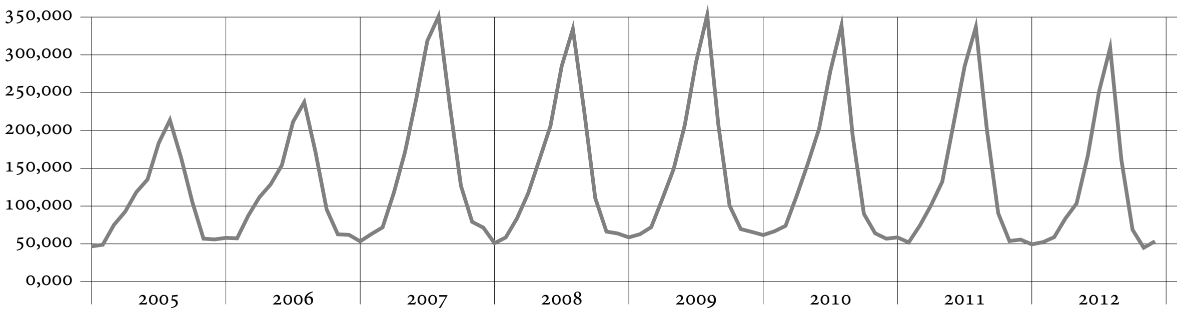


Figure 2 Monthly Overnight Stays of All Tourist Accommodations (Except Camping Sites) from 2005 (1) to 2012 (12)

horizon ranging from 1 to 24 months for all four types of the tested regressors, using two different Feature Sets as inputs:

- Feature Set 1 (FS1): This set was used to capture the pattern of the tourism data as it changes in the last 12 months (one year), and thus takes into consideration 12 previous monthly occupancy time-series values $X_{t-1}, X_{t-2}, \ldots, X_{t-12}$ in order to predict X_t.
- Feature Set 2 (FS2): The selection of this set is based on the fact that business and economic cycles usually last five years and all seasonal data are typically related to their predecessor and successor ones. Therefore, it is adequate to use only seven time series values from the past 60 ones (Psillakis, Panagopoulos & Kanellopoulos, 2009). FS2 takes into consideration $X_{t-11}, X_{t-12}, X_{t-13}, X_{t-24}, X_{t-36}, X_{t-48}, X_{t-60}$ occupancy time-series values, in order to predict X_t.

The prediction results for both feature sets are presented in a comparative plot in Figure 3. From this figure, it is clear that FS2 presents the smallest prediction error for all used predictors with a stable behaviour regardless of the forecast's time horizon (1–24 months). In Figure 4, we present a comparison of the four regressors' prediction accuracy for FS2. It is clear that the MLP performs the best, with the LR and the SVR-RBF performing worse after 15 months of prediction.

Overnight Stays Prediction

The prediction errors of the overnight stays were also estimated for different values of prediction horizon that range from 1 to 24 months and all four different types of regressors using the aforementioned two Feature Sets as inputs. The results are presented in Figure 5. From this figure, it is shown that also in this prediction task, FS2 shows the smallest error for all predictors, with the only exception of the SVR-RBF. In Figure 6, we present a comparison of the four regressors' prediction accuracy for FS2. It is clear that the SVR-RBF, as well as the LR, performs the best, showing great robustness regardless of the length of the forecasting horizon.

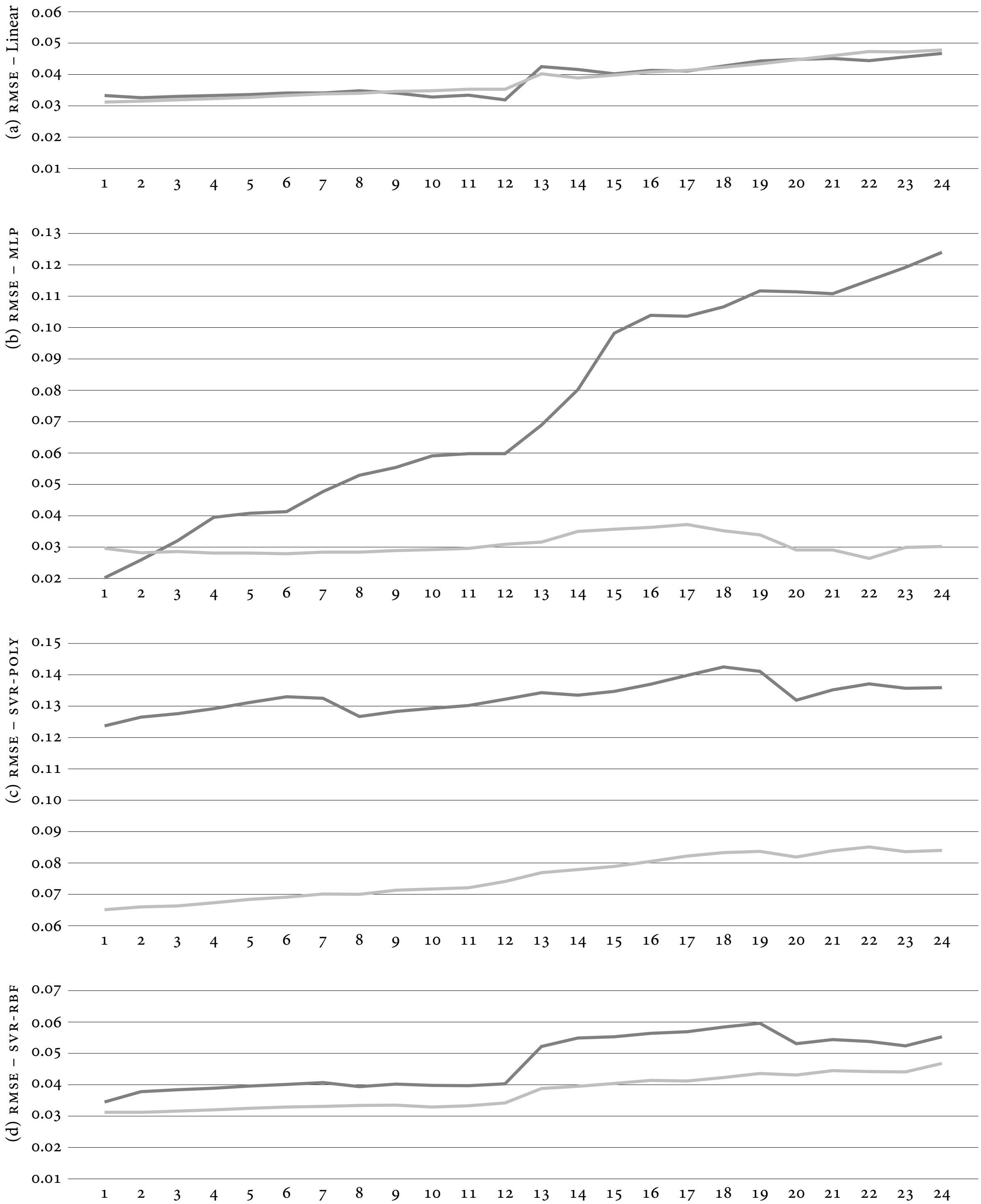


Figure 3 Monthly occupancy prediction performance (2011–2012) for FS1, FS2 and (a) Linear Regressor, (b) Multilayer Perceptron Regressor, (c) Support Vector Regressor (Polynomial Kernel), and (d) Support Vector Regressor (RBF Kernel) (gray – FS1, light gray – FS2)

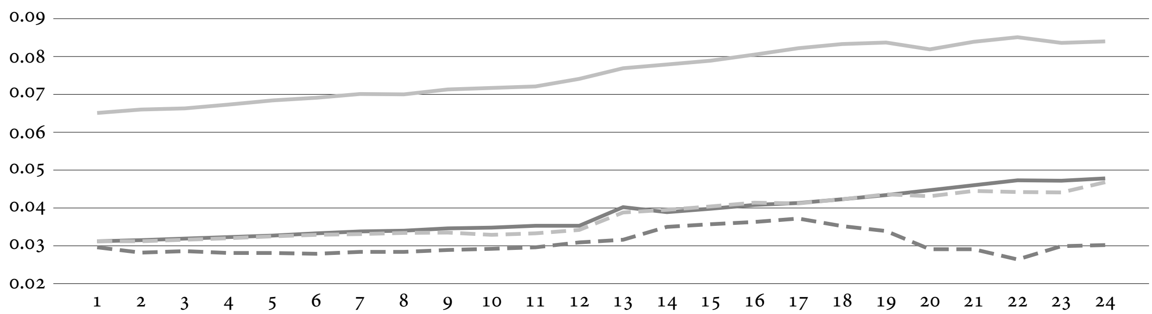


Figure 4 Comparison of the Proposed Four Models Occupancy Prediction Performance Using FS2 (gray – LINEAR, gray dashed – MLP, light gray – SVR-POLY, light gray dashed – SVR-RBF)

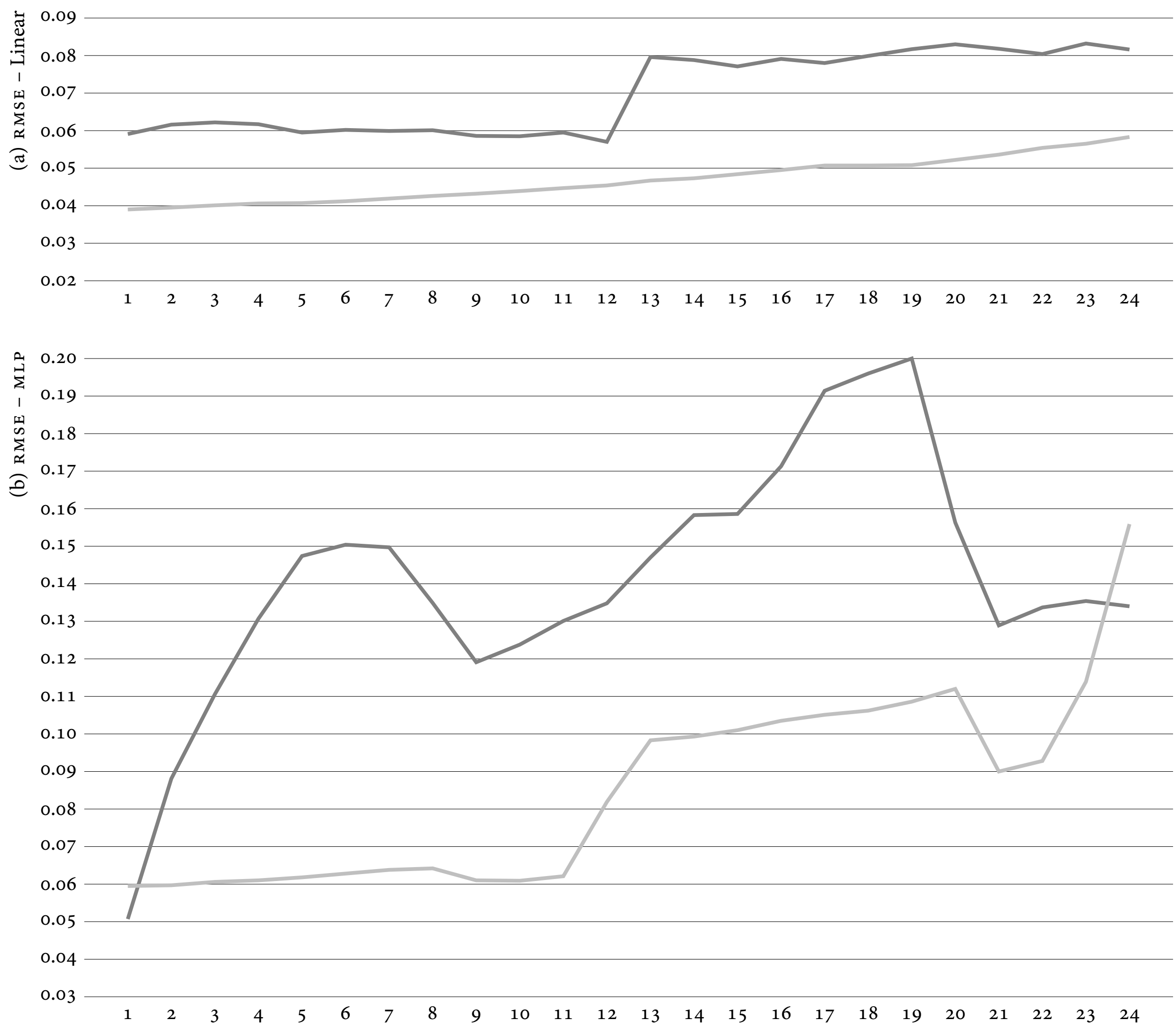
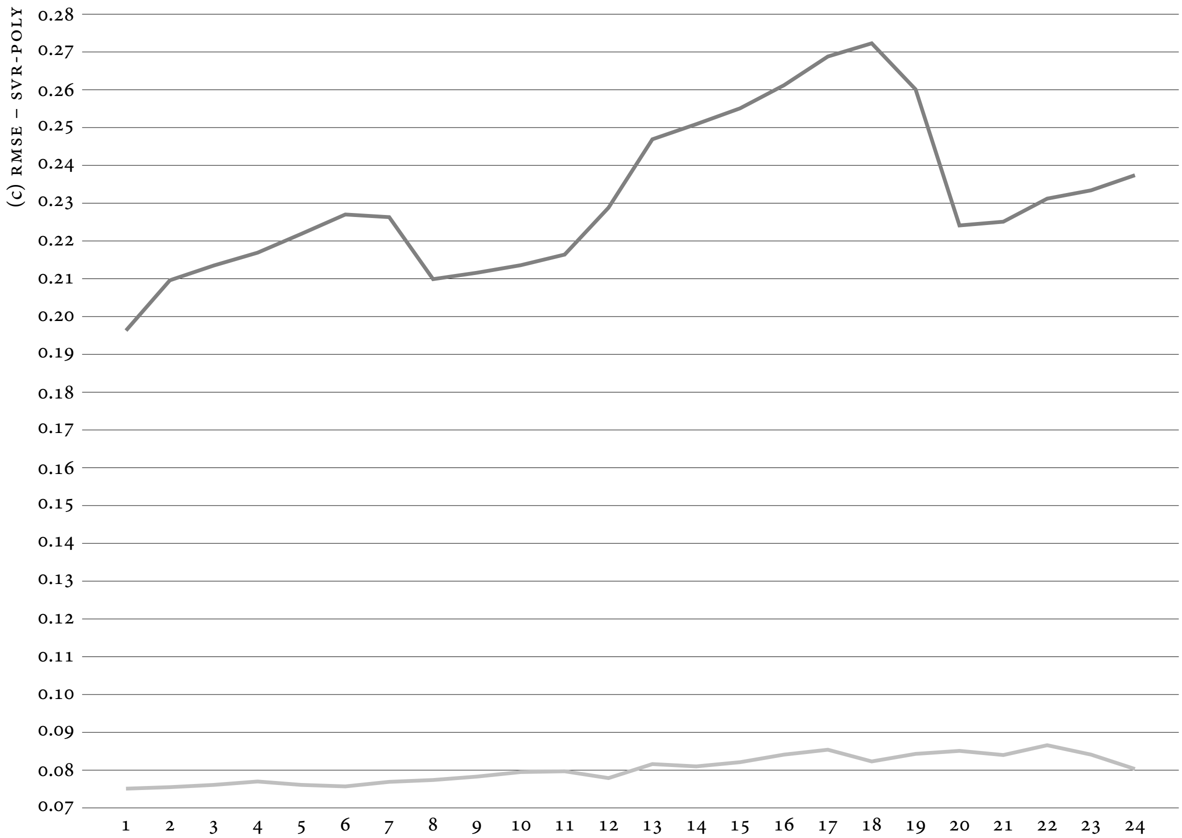


Figure 5 Monthly Overnight Stay Prediction Performance (2011–2012) for FS1, FS2 and (a) Linear Regressor, (b) Multilayer Perceptron Regressor, (c) Support Vector Regressor (Polynomial Kernel), and (d) Support Vector Regressor (RBF Kernel) (gray – FS1, light gray – FS2) *Continued on the next page*



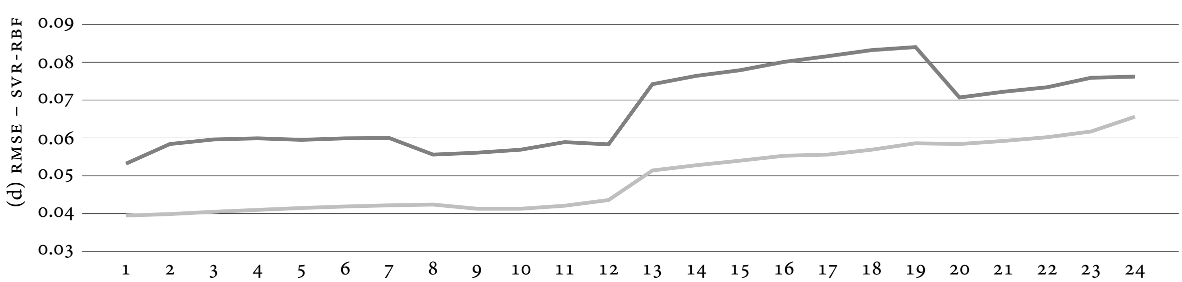


Figure 5 Continued from the previous page

Additionally, for comparison reasons, in Tables 1 and 2, we present analytically the forecasted occupancy and the overnight stays values for the years 2011 and 2012 estimated by the SVR-RBF, SVR-POLY, MLP, LR, together with the real values for this period.

Finally, Figures 7 and 8 make point-to-point comparisons of actual and predicted values of the best models presented in this (overnight stays) and the previous section (occupancy), respectively. As shown in these Figures, the MLP Regressor for the occupancy and the SVR-RBF for the overnight stays forecasting, work efficiently and manage to capture the trend of data extremely well, especially in the high-demand seasons (June–September).

Conclusion

In this paper we have presented an evaluation of the forecasting performance of four of the most widely

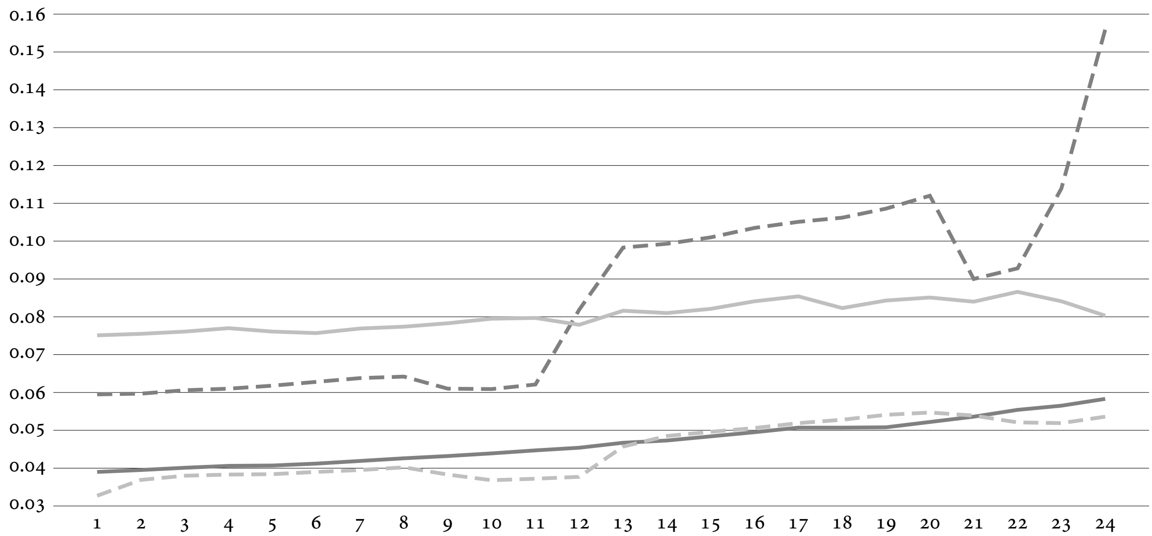


Figure 6 Comparison of the Prediction Performance of the Proposed Four Models Overnight Stays Using FS2 (gray – LINEAR, gray dashed – MLP, light gray – SVR-POLY, light gray dashed – SVR-RBF)

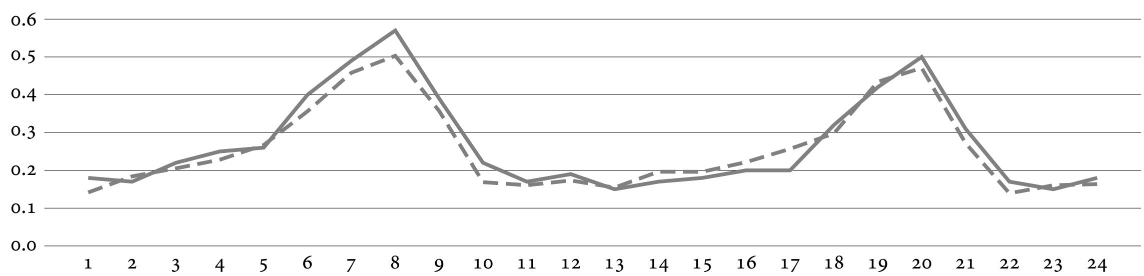


Figure 7 Illustration of Actual and Forecasting Values of Occupancy Using the Best Model (MLP Regressor – FS2; gray – actual, gray dashed – MLP)

known linear and non-linear regressors based on Artificial Neural Networks, the Multilayer Perceptron (MLP), and the Support Vector Regressor based on polynomial (SVR-POLY) as well as Radial Basis Functions Kernels (SVR-RBF) as forecasting models for tourism demand. For our experiments, we have used official statistical monthly data of the hotel occupancy and overnight stays in the Western Greece Region from 2005 to 2012 taken from the official records of the Hellenic Statistical Authority. Then the RMSE is computed for different forecast horizons, ranging from 1 to 24 months (2-year prediction, 2011 and 2012).

Extensive experiments have shown that forecasting tourism demand in the Western Greece Region can be accomplished with a small error when using features sets that take into account the trend the seasonality of the data, even if a small number (7) of observations are used. Furthermore, the difference in the behaviour of the data between the occupancy and the overnight stays was also highlighted as different regressors must be used in different forecasting problems.

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 tourism demand, the government, and private sec-

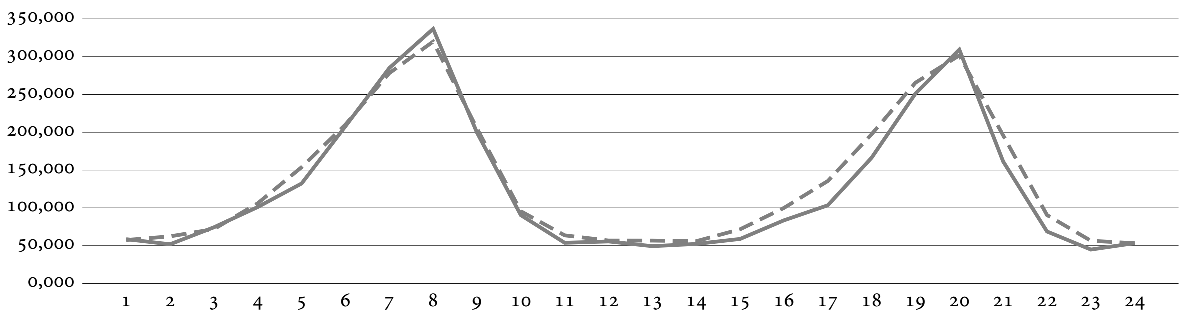


Figure 8 Illustration of Actual and Forecasting Values of Occupancy Using the Best Model (SVR-RBF – FS2; gray – actual, gray dashed – SVR-RBF)

tors can have a well-organized tourism strategy and provide 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

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, CA: 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, University Press.

Frees, E. W. (1996). *Data analysis using regression models: The business perspective.* New York, NY: 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, Germany: 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*(3), 301–328.

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

Table 1 Predicted Occupancy Values for the Years 2011 and 2012 (24-Month Prediction)

Months	Actual	FS1				FS2			
		LINEAR	NLP	SVR-POLY	SVR-RFB	LINEAR	NLP	SVR-POLY	SVR-RFB
01-11	0.1800	0.2030	0.1701	0.2078	0.2031	0.2036	0.1415	0.2006	0.1997
02-11	0.1700	0.2074	0.1927	0.2294	0.2192	0.2441	0.1841	0.2310	0.2202
03-11	0.2200	0.2141	0.2420	0.2048	0.2170	0.2035	0.2053	0.2020	0.2144
04-11	0.2500	0.2616	0.2618	0.2449	0.2609	0.2669	0.2281	0.2573	0.2545
05-11	0.2600	0.3042	0.2654	0.2797	0.3159	0.3033	0.2669	0.3046	0.3088
06-11	0.4000	0.3911	0.3710	0.3681	0.3891	0.3860	0.3563	0.3914	0.3946
07-11	0.4900	0.4710	0.4351	0.4714	0.4484	0.4712	0.4582	0.4775	0.4747
08-11	0.5700	0.5296	0.4451	0.5372	0.4820	0.5584	0.5031	0.5347	0.5256
09-11	0.3900	0.4046	0.1995	0.3743	0.3727	0.3662	0.3571	0.3664	0.3851
10-11	0.2200	0.2366	0.1357	0.2171	0.2521	0.2147	0.1687	0.2209	0.2464
11-11	0.1700	0.1796	0.1622	0.1809	0.2096	0.2031	0.1607	0.1974	0.2042
12-11	0.1900	0.1802	0.1498	0.1715	0.1920	0.1966	0.1732	0.1897	0.1961
01-12	0.1500	0.1650	0.1564	0.1623	0.1772	0.1763	0.1544	0.1757	0.1826
02-12	0.1700	0.1350	0.1709	0.1390	0.1684	0.1693	0.1964	0.1672	0.1840
03-12	0.1800	0.1915	0.2532	0.1803	0.2037	0.2111	0.1961	0.2007	0.2016
04-12	0.2000	0.2303	0.2273	0.2022	0.2390	0.2437	0.2218	0.2413	0.2404
05-12	0.2000	0.2460	0.2639	0.2053	0.2707	0.2549	0.2568	0.2664	0.2736
06-12	0.3200	0.3639	0.3609	0.3411	0.3707	0.3804	0.2976	0.3743	0.3663
07-12	0.4200	0.4575	0.4308	0.4475	0.4386	0.4650	0.4344	0.4632	0.4548
08-12	0.5000	0.5230	0.4551	0.5146	0.4694	0.5405	0.4709	0.5090	0.4997
09-12	0.3100	0.3843	0.2503	0.3556	0.3637	0.3707	0.2719	0.3633	0.3648
10-12	0.1700	0.2216	0.1440	0.1927	0.2396	0.2055	0.1394	0.2086	0.2229
11-12	0.1500	0.1443	0.1579	0.1340	0.1776	0.1610	0.1604	0.1570	0.1752
12-12	0.1800	0.1484	0.1403	0.1433	0.1715	0.1746	0.1636	0.1567	0.1710

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, England: Elsevier.

Lin, C. J., Hsu C. W., & Chang C. C. (2003). *A practical guide to support vector classification* (Technical report). Taipei, Taiwan: National Taiwan University.

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.

Table 2 Predicted Overnight Stay Values for the Years 2011 and 2012 (24-Month Prediction)

Months	Actual	FS1				FS2			
		LINEAR	NLP	SVR-POLY	SVR-RFB	LINEAR	NLP	SVR-POLY	SVR-RFB
01-11	57481	43059	65609	51636	45204	62254	33998	87536	56409
02-11	50910	67096	65332	58554	53850	66197	49112	87917	61148
03-11	72907	76296	71662	70071	71454	70071	67684	89473	70555
04-11	99226	92205	79685	95768	101959	108668	95560	96425	104103
05-11	129766	129731	89715	130630	144188	148027	148926	116450	150932
06-11	204090	199940	170715	181748	201081	193749	163832	174312	206304
07-11	279660	278138	243414	246181	254136	266068	273331	275613	273020
08-11	330536	326420	205716	298232	271117	325694	361351	312724	314004
09-11	195859	231551	74013	241789	193369	180607	117522	169920	201220
10-11	88712	111919	69379	144949	107527	85219	105141	99434	93762
11-11	52812	40569	59972	76331	63569	63534	43578	89024	62462
12-11	54576	29778	68238	54749	49008	56478	34032	87882	55718
01-12	48420	50011	64364	55441	46587	58346	31162	87363	55683
02-12	51325	72803	66543	61113	54922	52051	59245	87606	54853
03-12	57793	76988	72768	72250	72319	72976	61251	88989	70313
04-12	81968	91306	81657	97808	103273	98604	103031	95768	98050
05-12	101474	134538	94315	135196	148511	128520	129731	113026	132982
06-12	163245	201565	181540	182578	202949	197934	223043	171753	193749
07-12	246527	270771	258701	240405	249605	269353	262609	286093	260638
08-12	303593	311444	173378	286024	261053	317843	297437	317912	296054
09-12	158368	217094	74048	228923	183581	188285	125477	167602	192262
10-12	67477	102132	69621	134642	100783	85807	101924	98915	88816
11-12	44028	36765	59591	69898	60387	52190	46933	88574	55475
12-12	52328	25697	67650	49112	46725	53331	20717	87467	52224

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.