Advances in Management & Applied Economics, vol. 3, no.6, 2013, 1-18
ISSN: 1792-7544 (print version), 1792-7552(online)
Scienpress Ltd, 2013

Tourism Demand Forecasting: Econometric Model based on Multivariate Adaptive Regression Splines, Artificial Neural Network and Support Vector Regression

Chang-Jui Lin[1] and Tian-Shyug Lee[2]

Abstract

This paper develops tourism demand econometric models based on the monthly data of tourists to Taiwan and adopts Multivariate Adaptive Regression Splines (MARS), Artificial Neural Network (ANN) and Support Vector Regression (SVR), MARS, ANN and SVR to develop forecast models and compare the forecast results. The results showed that SVR model is the optimal model, with a mean error rate of 3.61%, ANN model is the sub-optimal model, with a mean error rate of 7.08%, and MARS is the worst model, with a mean error rate of 11.26%.

JEL classification numbers: M31
Keywords: Tourism demand, Econometric models, Multivariate adaptive regression splines, Artificial neural network, Support vector regression.

1 Introduction

The data of The World Tourism Organization show that the number of international tourists was 435 million in 1990, 674 million in 2000, 940 million in 2010, and 983 million in 2011, growing continuously year-by-year. Tourism revenue was USD 927 billion in 2000 and grew to USD 1,030 billion in 2011, reflecting an increase of 11%. It shows that tourism has become an important industry for every country; thus, countries are paying more attention in to developing tourism. In 2009, the Tourism Bureau, Ministry of Transportation and Communications, R.O.C. proposed a five-year

[1]Fu Jen Catholic University.
[2]Fu Jen Catholic University.

Article Info: *Received* : July 14, 2013. *Revised* : August 12, 2013.
Published online : November 15, 2013

"Project Vanguard for Excellence in Tourism" and invested NTD 30 billion (USD 1 billion) to construct tourism facilities, train tourism industry workers, and promote tourism to increase Taiwan's international tourists and foreign exchange earnings from international tourists. The number of passengers arriving to Taiwan increased from 3,845,000 in 2008 to 6,087,000 in 2011, and the foreign exchange earnings from international tourists increased from USD 6.82 billion (accounting for 1.48% of GDP) in 2008 to USD 8.72 billion (accounting for 2.37% of GDP) in 2011. It shows that tourism industry is gaining increasing attention in Taiwan.

The development of tourism entails investment in many aspects, including the construction of traffic infrastructure, international airports, public transport facilities, tourist hotels, as well as recreational facilities in tourist spots, which all require a long period of planning and years of construction. Good long-term planning entails forecasting of the demand of international tourists to avoid inadequate investment and construction or waste due to excessive construction. Therefore, the correctness of forecasting the demand of international tourists of a country is very important. A model that can correctly forecast the tourism demand would help a country develop its tourism. Witt and Witt (1995) reviewed 114 papers on tourism demand forecasting models proposed during the 30 years before 1995. Studies on quantitative forecasting utilized econometric models and time-series models. Econometric model is a causal model that forecasts the tourism demand through many variables, such as GDP of the country of origin, tourist expenses, accommodation expenses of hotels, and transport expenses of travelling, the destination's consumption level, and the exchange rate (Witt & Witt, 1995). Time-series model uses the number of tourists in the previous periods to forecast the number of tourists in the current period. Song and Li's (2008) study reviewed 121 papers on forecasting of tourist models, out of which 72 papers adopted time-series models, 71 papers adopted econometric models, and 30 papers adopted both time-series models and econometric models. Thus, the tourist demand model for quantitative forecasting can be divided into (1) time-series models and (2) econometric models.

Over the past 40 years, time-series models adopted mostly ARIMA models for tourism demand forecasting, followed by Naïve 1, Naïve 2, exponential smoothing models, and simple autoregressive models (Song, & Li, 2008). The less often adopted models were artificial neural network (Cho, 2003; Huarng et al., 2006; Lin et al., 2011), support vector regression (Chen, & Wang, 2007), and multivariate adaptive regression splines (Lin et al. 2011). Econometric model has been frequently used to analyze the relation between "tourism demand" variable and its affecting factors. The most frequently used analytical model was traditional regression analysis model (Ordinary least squares, OLS) (Li et al. 2005; Song, & Li, 2008), less frequently used models were ADLM, ECM, VAR, TVP, and AIDS (Song, & Li, 2008), and the least frequently used model was structural equation (Song, & Li, 2008). Few studies adopted ANN model for econometric model analysis (Law, 2000; Moutinho *et al.*, 2008).

The variables in the econometric model include the population and income of the country of origin of tourists, the travel expenses and accommodation fee in the tourist destination, and the costs of travelling to the tourist destination. At the same time, the tourist destination country's supply for tourism, such as hotel rooms and means of transport, are also considered (Witt, & Witt, 1995). In most studies, the forecasted number of tourists is the total number of tourists from countries of origin (Bicak et al., 2005; Lee, 2011; Song, & Witt, 2006; Wong, et al. 2006) while the number of tourists from main countries of origin is usually calculated or the number of those from a single country of origin is adopted to forecast the number of tourists (Law et al., 2004; Wang, 2009). When the econometric model's tourism demand forecast does not adopt the total number of tourists in the tourist destination for forecasting, the application of the forecast results will be restricted. If the total number of tourists in the tourist destination is combined with other relevant variables with an aim to conduct the analysis of the econometric models, the obtained forecast results, that is, relevant economic variables' forecast of the total number of tourists, will be able to provide more reference value for the destination's tourism management authority in planning its tourism development. Econometric models have too many forecast variables; thus, identifying important ones will help set up a new econometric model that will have better application value.

Thus, it is worth studying how to forecast econometric model's number of tourists through the total number of tourists in the tourist destination and other relevant variables. This paper will discuss the forecasting of the total number of tourists in the tourist destination and adopt MARS, ANN, and SVR models, which are adopted less frequently in econometric model to conduct analysis of tourism demand forecast. The second part of the paper is the literature review, the third part is research method and data, the fourth part presents empirical results, the fifth part is discussion, and the sixth part is conclusion.

2 Literature Review

The quantitative methods for tourist demand models include time-series models and econometric models (Song & Li. 2008). Some scholars have used a single method while other scholars used both time-series and econometric models to study the tourist demand models. This paper adopts econometric models. The main forecast variables of econometric models are the country of origin's population, Gross Domestic Product (GDP), and tourist expenses, which include travelling expenses, accommodation fee and consumption in the tourist destination country. Moreover, the tourist destination country's supply, which includes the number of hotel (accommodation) rooms and the capacity of vehicle, is considered because the quantity of supply cannot increase instantly (Witt & Witt, 1995). This paper makes a new attempt in econometric model's forecast model by adopting less frequently used quantitative methods of MARS, ANN's back-propagation, and SVR.

MARS has advantages that include screening variables automatically, short calculating time, and rapidly setting up classification models or forecast models; thus, it is applied in many fields. For example, De Veaux *et al*. (1993a) applied MARS to survey the terrain around Antarctica; De Veaux *et al*. (1993b) compared the forecast results of MARS with those of ANN and pointed out that MARS had a better forecasting accuracy compared to ANN. De Gooijer *et al*. (1998) applied TSMARS (time series MARS) model to forecast exchange rate and found that TSMARS provided better forecast results compared to the random walk model. Lin *et al*. (2011) applied MARS, ANN, and ARIMA models to forecasting of tourism demand's time-series models and found that ARIMA had the best effect with respect to forecast results, followed by ANN and MARS. Different studies provided different results. This paper applies MARS model to forecasting of tourist demand's econometric models and compares it with ANN's forecast results to investigate whether MARS or ANN provides better forecast results.

The application of ANN in tourism demand's econometric models forecasting is seen in Law and Au's (1999) econometric ANN-model forecasting and analysis of the demand of tourists from Japan to Hong Kong, Law's (2000) econometric ANN-model forecasting and analyzing the demand of tourists from Taiwan to Hong Kong, and Moutinho *et al*.'s (2008) econometric ANN-model forecasting and analysis of the demand of tourists from mainland China to Taiwan. Besides, ANN has also been often used in tourist demand's time-series forecasting models (Chen, & Wang, 2007; Cho 2003; Palmer et al. 2006). Previous scholars also used different time-series analysis methods to forecast tourist demand and to identify a better method to forecast tourist demand model. The results showed ANN has a lower error rate than ARIMA, GMDH (group method of data handling – genetic regression) algorithm, and exponential smoothing (Burger et al., 2001; Cho, 2003).

Support vector regression has been used to build forecast models based on statistical learning theory and the principle of risk minimization (Drucker *et al*., 1997; Vapnik *et al*., 1997). The SVR method has been widely applied as a forecasting tool in many fields because it is able to handle the interaction between variables and does not require too many assumptions for data. For example, forecasting of stock market index, forecasting of Nikkei 225 (Lu *et al*., 2009), S&P500 and NASDAQ index forecasting (Wang & Zhu, 2010), and natural gas price forecasting (Viacaba *et al*., 2012). Chen and Wang (2007) adopted time-series, SVR, ANN, and ARIMA forecast models to forecast tourism demand in mainland China, and the result indicated that SVR is the best forecasting model with the lowest error rate.

The accuracy of tourism demand forecasting is very important, and according to previous scholars' time-series models, MARS, ANN, and SVR all provide good forecast results. Hence, this paper adopts econometric models ANN, SVR, and MARS as the forecast models to assess which model is the best forecast model, that is, which one has the lowest error rate in tourism demand's forecast result.

3 Research Method and Data

3.1 Econometric Models

Concerning the variables of the econometric models for tourism demand forecasting, the dependent variable is the number of tourists and the explanatory variables include the population and income of the country of origin of tourists; tourists' travel expenses, accommodation fee every night and price level in the tourist destination; and the expenses of travelling to the tourist destination. Moreover, "the tourist destination's supply" variable, which includes the number of hotel rooms and the capacity of vehicle are taken into consideration, as the quantity of supply cannot increase shortly (Witt & Witt, 1995). This paper mainly adopts the total number of tourists in tourist destination as the dependent variable, so the selection of explanatory variables does not consider the variables of individual countries of origin but only that of the destination. The explanatory variables selected in this paper are the tourist destination's GDP, tourist consumption price and accommodation fee every night, the expenses of travelling to the destination, the number of hotel rooms in the destination, and the capacity of vehicle to the destination.

Among the various data of the tourist destination, the number of tourists is clearly available. The data of tourist consumption price are not easily available. Tourist consumption includes the expenses associated with purchasing food, goods, and services in the destination and it relates to the destination's income and price level. However, foreign tourists' consumption must consider the factor of exchange rate. The actual data of discretionary income are not easily available; thus, proxy variable shall be used in place of discretionary income (Song & Witt, 2006). The most frequently used proxy variable of income is per capita GDP (Lim & McAleer, 2001; Moutinho, et al. 2008; Song et al., 2003;). Many studies adopted CPI as a proxy variable of tourist price (Law et al., 2004; Lim & McAleer, 2001; Song et al., 2003); hence, this paper adopts CPI as the proxy variable for tourist consumption's price level. At present, USD is the strongest currency in the world, and international trade is mostly valuated in USD; therefore, the exchange rate of NTD against USD is adopted as the proxy variable of exchange rate.

The average per-night accommodation price of all tourist hotels in the destination is adopted as the proxy variable for tourists' per-night accommodation fee. Regarding the expenses of travelling to the destination, this paper studies Taiwan's tourism demand forecast. Since Taiwan is an island country, tourists to Taiwan take planes as the main means of transport, with some taking cruise ships. The airfare and cruise ship costs of tourists traveling to Taiwan from different countries are not equal; thus, we must select a variable as the proxy variable for traveling expenses. When evaluating transport costs of different means of transportation, the international oil price can be adopted as a proxy variable (Garin-Munoz, 2006). The data of the number of hotel rooms in the destination adopt the number of rooms of tourist hotels, and the data of the capacity of vehicle to the destination adopt the passenger capacity of Taiwan's international flights. The data for tourism demand forecast are usually collected yearly, quarterly, and monthly (Song & Li, 2008). This paper adopts monthly data for analysis. When using monthly data, the seasonality of monthly data cannot be ignored (Song, & Li, 2008). Thus, a seasonal variable is added into the tourism demand forecast model in this paper. According to the description above, this paper sets up a tourism demand forecast model as follows:

$$Q = f(Y, P, EX, HP, HR, T, AS, SE) \tag{1}$$

The variables of the forecast model are the number of passengers arrived (Q), GDP, proxy variable GDP (Y), tourist consumption price for which CPI is proxy variable (P), exchange rate of NTD against USD (EX),average hotel accommodation price (HP), the number of hotel rooms (HR), transport costs and traveling expenses represented by international oil price as a proxy (T), capacity of vehicle with the passenger capacity of international flights as a proxy (AS), And monthly/seasonal index as "the monthly number of passengers arrived" divided by "the monthly average number of passengers arrived in the year" (SE).

$$SEij = \frac{Qmij}{Qaj} \tag{2}$$

Equation 2 comprises monthly/seasonal index (*SEij*), the number of passengers arrived monthly (*Qmij*), the annual average number of passengers arrived (*Qaj*), year (j), and month (i).

3.2 Data

This paper adopts the forecast of Taiwan tourism demand as the research model, utilizing monthly data collected between Jan. 2004 and Dec. 2012, i.e., nine years (108 months). The data for the preceding 84 months, i.e., from Jan.2004 to Dec. 2010, are adopted as training samples, and the data collected in the following 24 months, i.e., from Jan. 2011 to Dec. 2012, are adopted as test samples. The data source is the four sets of data comprising the number of passengers arriving, the average hotel accommodation price, the number of hotel rooms, and the passenger capacity of international flights obtained from the Ministry of Transportation and Communications. R.O.C; Taiwan's Gross Domestic Product (GDP) data are obtained from National Statistic, R.O.C. (Taiwan); GDP is quarterly data, which are (divided by 3) converted into monthly data; the data of international oil price are obtained from the Bureau of Energy, Ministry of Economic

Affairs, R.O.C., the data of the exchange rate of NTD against USD are obtained from the Central Bank of the Republic of China (Taiwan), and the CPI data are obtained from Accounting and Statistics, Executive Yuan, ROC.

3.3 Multivariate Adaptive Regression Splines (MARS)

The statistician and physicist Friedman (1991) introduced Multivariate Adaptive Regression Splines. It is a new way of dealing with diverse information and issues. The basic idea is to add up sections of spline's basis function (BF) to form a flexible MARS prediction model, to determine the value of the function of the basic equations by referring to the cross-validation among the parameters, and to assess its loss of fit (LOS) by the judging criteria in order to get the best and the most suitable variable sets, knots, and the interaction to solve various high-dimensional data problems. It is a flexible regression analysis and can automatically create an accurate model for speculating the continuous and discrete response variables (Friedman, 1991).

$$\hat{f}(x) = a_0 + \sum_{m=1}^{M} a_m \prod_{k=1}^{K_m} [s_{km} \cdot (x_{v(k,m)} - t_{km})]_+ \tag{3}$$

Equation (3) is a universal MARS model in which BF is the part of back-end multiplication, and as shown in Equation (4), it changes mainly according to demand.

$$B_m(x) = \prod_{k=1}^{K_m} H[s_{km} \cdot (x_{v(k,m)} - t_{km})]_+ \tag{4}$$

Therein, a_0 and a_m are both parameter values the function of which is similar to the regression coefficient in a linear regression model; m is the number of BF, which is determined by evaluation criterion; Km the number of knots cut; the value of Skm is +1or-1 and shows the direction; v(k,m) is the mark of a variable; and tkm is the demarcation point (value) of each knot. In a set made up of a given target variable and a selectable forecast variable, MARS can automatize all models' building and scheduling, which includes separating significant variables from less appropriate variables, determining the interaction between explanatory variables, adopting a new variable-clustering technology to deal with missing value problems, and adopting a lot of self-testing to avoid over-fitting (Steinberg *et al.*, 1999).

We can regard BF as an explanatory equation, which belongs to each section of rule. Each BF is the number of effect variables, which are included in the criteria for evaluating its LOF. Meanwhile, forward algorithm and backward algorithm are adopted to search for appropriate knots and interaction and to solve various problems of high-dimensional data. It is a considerably elastic regression processing procedure that can build accurate models rapidly and automatically to infer its continuous or binary response variables (Friedman, 1991).

The backward stepwise procedure algorithm determines the number of BF according to LOF and examines whether each BF contributes to the outcome after it is added to a main model by removing the non-contributing BFs and retaining those contributing to the main model. LOF is adopted as the judgment criterion, and Craven and Wahba's (1979) GCV (generalized cross-validation criterion) is adopted to delete the BFs that contribute less to the model and finally produce the optimal MARS model, reducing the model's complexity within an acceptable range and accelerating data processing and judgment. The processing method of GCV is shown below:

$$LOF(f_M) = GCV(M) = \frac{1}{N}\sum_{i=1}^{n}\left[y_i - \hat{f}_M(x_i)\right]^2 / \left[1 - \frac{C(M)}{N}\right]^2 \tag{5}$$

Therein, *yi* is the actual value; $\hat{f}_M(x_i)$ is the forecast value; *C* (*M*) is the cost of adopting *M* (number of) BFs. Its main concept comes from the equation below.

$$\Delta\left[f(x), \hat{f}(x)\right] = \left[f(x) - \hat{f}(x)\right]^2 \tag{6}$$

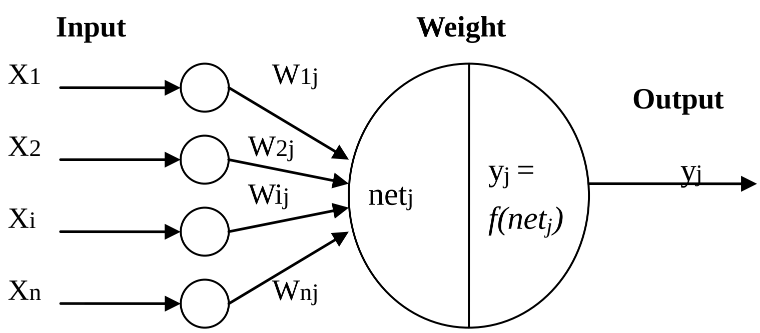
Thus, the optimal MARS model is determined by two processing procedures. First, forward algorithm is adopted, where BF increases (all the main effect, knots or interaction) until a complete model is found. Subsequently, backward algorithm is adopted, where less-contributing basis functions are removed until the optimal balance point between error and the number of variables is found, thereby completing the optimal model of MARS.

A variable's relative importance can be judged by observing the decrease in the degree of GCV value after deleting a certain BF. When the GCV value decreases after a certain BF is deleted from the model, it would imply that this BF is an important key factor in the model.

3.4 Artificial Neural Network (ANN)

It is a science to use a computer to recreate artificial neural network to simulate a human brain's nerve cells. Freeman and Skapura (1992) thought artificial neural network was an information processing system that used plenty of artificial neurons to simulate biological neural network, enabling the computer to simulate human nervous system structure to carry out data processing. Its conceptual method is shown in Fig. 1.

Since the neural network has segmentation and identification ability (Zhang et al., 1998), it is widely used in various commercial and financial aspects, e.g., credit card fraud judgment, stock prices, exchange rates, interest rates, and bankruptcy prediction. Financial analysis has been used in various scientific applications, e.g., weather forecasting, medical images judgment, and fingerprint recognition system, among others (Berry & Linoff, 1997; Chiu *et al.*, 2003; Fish *et al.*, 1995; Lee & Chiu, 2002; Lee *et al.*, 2002; Leung *et al.*, 2000; Vellido *et al.*, 1999).



$$net_j = \sum_{i=1}^{N} W_{ij} X_i + \theta_i$$

Figure 1: An artificial neuron

Many sophisticated models have been proposed to illustrate the development of neural network. They can be divided into 3 structural networks of learning strategies: supervised learning, non-supervised learning, and associative learning. Among all the network models, the back-propagation network (BPN) of supervised learning is the most representative and the most widely used. According to the research of Vellido et al. (1999), 78% of the researchers used the BPN type of artificial neutral network in the commercial aspects between 1992 and 1998. This is quite a high proportion. They chose BPN because it has the advantages of a high learning accuracy and quick retrospect speeds; hence, BPN is also used as an analytical tool in this study

The structure of back-propagation neural network (BPN) consists of three layers: input layer, hidden layer and output layer. The number of neurons of the input layer serves as the number of variables of forecasted output values. The hidden layer is the transfer layer in the neural network, and it indicates the interaction between input layer neurons. The output layer neurons' results are the final output values. This paper adopts BPN model, as shown in Fig. 2.

3.5. Support Vector Regression (SVR)

3.5.1. Support vector regression

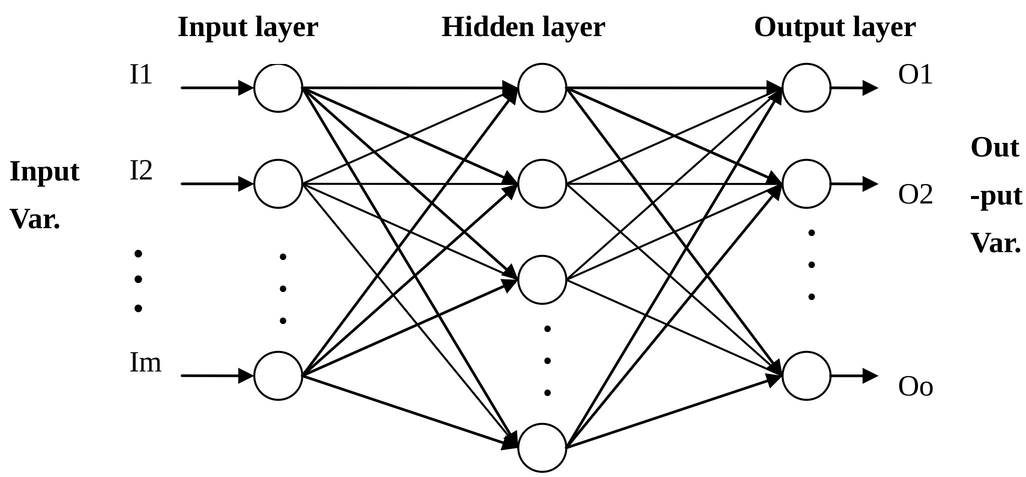


Figure 2: back-propagation network

The main concept of support vector regression is to use the principle of structural risk minimization to convert a non-linear problem in low-dimensional input space to a linear regression problem in high-dimensional feature space. SVR through training data $G = \{(x_i, d_i)\}_i^n$, in which x_i is input vector, d_i is forecast target value, and n is the number of training samples. Equation (7) is an estimating function of SVR expressed by mathematical model:

$$f(x) = w\psi(x) + b \tag{7}$$

Therein, w is weight vector, b is bias, and ψ (x) is a transfer function, which is used to convert the input data into high-dimensional feature space in a non-linear manner. According to the radial basis function suggested by Hsu *et al*. (2003), where w and b go through adjusted risk minimization equation, the equations are shown as follows:

$$R_{SVR}(C) = C\frac{1}{n}\sum_{i=1}^{n} L_{\varepsilon}(d_i, y_i) + \frac{1}{2}\|w\|^2 \tag{8}$$

$$L_{\varepsilon} = \begin{cases} |d-y| - \varepsilon \\ 0 \end{cases} if \begin{cases} |d-y| \geq \varepsilon \\ otherwise \end{cases} \tag{9}$$

Equation (8) is Vapnik's ε-insensitive loss function (Vapnik *et al*. 1997), ε represents a parameter to define a tube zone that surrounds regression equation f (x) (as shown in Fig. 3),

In Fig. 3, the tube zones are the so-called ε-insensitivity zones. When the forecast value falls in the tube zone, its loss is zero, i.e., when the forecast error is less than ε, its loss is zero. Conversely, when the forecast value falls outside the tube zone, its loss is the difference between forecast value and zone boundary that can be indicated by slack variables ξ and ξ^*.

In Equation (8), $C\frac{1}{n}\sum_{i=1}^{n} L_{\varepsilon}(d_i, y_i)$ is the empirical risk measured through Vapnik's ε -insensitive loss function, which enables sparse data to estimate the forecast equation through this model. Additionally, $\frac{1}{2}\|w\|^2$ measures the

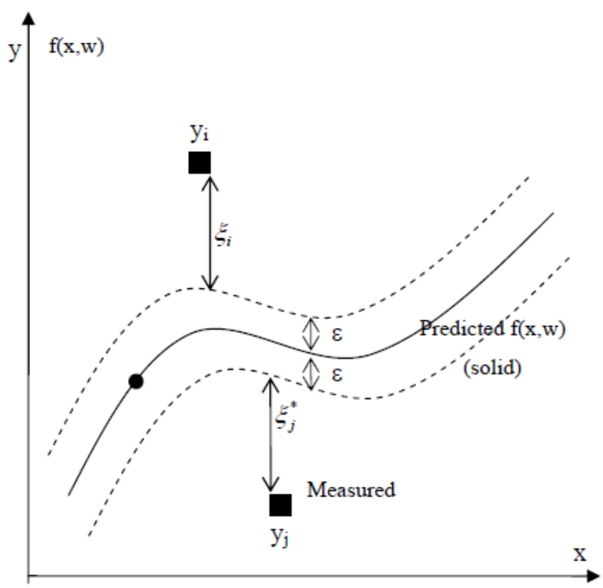


Fig. 3 A schematic representation of the SVR using ε-insensitive loss function. (form Lu et al., 2009)

Greater C value suggests that increased importance is placed on empirical risk. In order to solve the optimization problem in Equation (8), Cortes and Vapnik (1995) proposed to add a non-negative slack variable ξ in the optimization problem, thus SVR's optimization problem is shown in Equation (10):

Minimize:

$$R_{SVR}\left(w,\xi^{(*)}\right)=\frac{1}{2}\|w\|^{2}+C\sum_{i=1}^{n}\left(\xi_i+\xi_i^{*}\right) \tag{10}$$

Subjected to:

$$\begin{aligned} &d_i - w\phi(x_i) - b_i \le \varepsilon + \xi_i \\ &w\phi(x_i) + b_i - d_i \le \varepsilon + \xi_i^{*} \end{aligned} \qquad \xi_i, \xi_i^{(*)} \ge 0$$

Therein, ξ is the value of the training error greater than tube zone boundary and ξ* is the value of the training error lower than the zone boundary. They are both positive numbers. In order to satisfy the optimization problem, Lagrange multipliers can be introduced to finally obtain support vector regression's forecast model equation as follows:

$$f(x,\alpha_i,\alpha_i^{*})=\sum_{i=1}^{i}\left(\alpha_i-\alpha_i^{*}\right)\phi(x_i)\phi\left(x_j\right)+b=\sum_{i=1}^{i}\left(\alpha_i-\alpha_i^{*}\right)K\left(x_i,x_j\right)+b \tag{11}$$

Therein, K (x_i, x_j) is kernel function.

3.5.2 Lagrange multipliers

In Equation (11), α_i and α_i^{*} are "Lagrange multipliers", which are introduced into Equation (10) to obtain the maximization dual equation (12), as is shown below:

Maximize

$$R\left(\alpha_i,\alpha_i^{*}\right)=\sum_{i=1}^{n}\left(\alpha_i-\alpha_i^{n}\right)-\varepsilon\sum_{i}^{n}\left(\alpha_i+\alpha_i^{*}\right)-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\alpha_i-\alpha_i^{*}\right)\left(\alpha_j-\alpha_j^{*}\right)K\left(x_i,x_j\right) \tag{12}$$

Subjected to $\sum_{i=1}^{n}\left(\alpha_i-\alpha_i^{*}\right)=0$ $\quad \begin{aligned} &0\le\alpha_i\le C \quad i=1,2,\cdots,n \\ &0\le\alpha_I^{*}\le C \quad i=1,2,\cdots,n \end{aligned}$

The above equation solves the quadratic programming problem through KKT (Karush–Kuhn–Tucker), and only a part of $\left(\alpha_i-\alpha_i^{*}\right)$ in Equation (11) is assumed as non-zero values. The evaluated error of the data points of such non-zero values is expected to be ≥ ε, and one that falls on or outside the tube space of the determining equation is called "support vector". From Equation (11), the so-called support vectors refer to the data other than those for which $\left(\alpha_i-\alpha_i^{*}\right)$ coefficient is 0 in the process of being used in the determining equation. These values have a decisive influence on the optimal solution. Normally, greater ε indicates fewer support vectors that are of greater importance for the optimal solution. However, greater ε will reduce the training data's accuracy in estimating the optimal solution.

In Equation (11), K (x_i, x_j) is defined as a kernel function, which is able to convert input data into high-dimensional feature space in a non-linear manner. Its value is the inner product of the two-feature vectors x_i and x_j, i.e., $K\left(x_i,x_j\right)=\phi(x_i)\times\phi\left(x_j\right)$.

The common kernel functions are as follows:

(a) linear kernel function: $K(x_i, x_j) = x_i \times x_j$ (13)

(b) polynomial kernel function : $K(x_i, x_j) = (x_i \times x_j + 1)^d$ (14)

(c) Gaussian Radial basis function (RBF): $K(x_i, x_j) = \exp\left(-\|x_i - x_j\|^2 / 2\sigma^2\right)$ (15)

σ^2 is the bandwidth of Gaussian radial basis function.

3.5.3 Parameter setting of support vector regression

Three parameter set values, C, ε, and γ, are included in a SVR model to solve the problem. C value is the trade-off relation between adjusted empirical risk and structural risk, γ is the parameter of RBF kernel function, and ε is episso loss function. In the setting of parameters C, ε and γ, the parameter set values of best-fitting parameters often differ according to different data. Thus, parameter setting relies mostly on empirical law, which in practice often adopts try and error to search for the optimal parameter combination. Hsu et al.'s (2003) suggested that RBF as a kernel function has already been able to solve most problems. To shorten the time to search by trial and error, exponential growth sequence (e.g., $2^{-15}, 2^{-14}, 2^{-13}, \ldots, 2^{14}, 2^{15}$) can be adopted to search for the optimal parameter combination.

4 Empirical Results

4.1 MARS

To analyze tourism demand forecasting based on MARS model, the first 84 months of data are adopted as training samples and the last 24 months of data as test samples. The number of passengers arriving to a destination is adopted as a dependent variable while the remaining 8 variables are used as explanatory variables. The adjusted R^2 of the MARS regression model is 0.969. It comprises six important variables: HR (the number of hotel rooms), SE (monthly/seasonal index), P (tourist consumption price, CPI), Y (GDP), AS (capacity of vehicle), and EX (exchange rate), as shown in Table 1.

Table 1: Table of Important Variables

Variable	Importance	Variable	Importance
HR	100	Y(GDP)	20.926
SE	74.72	AS	13.291
P(CPI)	55.347	EX	7.018

10 basis functions for MARS are obtained, as shown in Table 2

Table 2: Basis Functions of MARS

NO	BF	NO	BF
1	BF2 = max(0, 21570.000 - HR);	6	BF9 = max(0, GDP - 1301.000);
2	BF3 = max(0, SE - 89.896);	7	BF11 = max(0, HR - 21898.000);
3	BF4 = max(0, 89.896 - SE);	8	BF13 = max(0, AS - 2960806.250);
4	BF6 = max(0, 97.420 - CPI);	9	BF14 = max(0, 2960806.250 - AS);
5	BF7 = max(0, HR - 23012.000);	10	BF15 = max(0, EX - 30.350)

MARS regression model is made up of 10 basic functions, as follows:

$$
\begin{aligned}
Y = {} & 311639.188 + 30.690 * BF2 + 3339.143 * BF3 - 2067.524 * BF4 \\
& - 12093.573 * BF6 - 48.327 * BF7 - 224.273 * BF9 + 152.346 * BF11 \\
& + 0.065 * BF13 + 0.053 * BF14 - 5778.533 * BF15
\end{aligned} \tag{16}
$$

4.2 ANN

As mentioned above, to analyze tourism demand forecast based on BPN model, the first 84 months of data are adopted as training samples and the last 24 months of data as test samples. In the training samples, GDP, CPI (tourist consumption price), EX (exchange rate), HP (average hotel accommodation price), HR (the number of hotel rooms), TC (transport costs), AS (the passenger capacity of international flights) and SE (monthly/seasonal index) are adopted as the input layer data consisting of 8 neurons. The hidden layer consists of 14 to 18 neurons and the output layer consists of one neuron. The number of passengers arriving (Q) is adopted as output layer data. The learning rate is set for 0.001, 0.003, 0.005, 0.007, 0.009, 0.01, 0.03, 0.05, 0.07, and 0.09, and the number of iterations is set from 1,000 to 120,000 to carry out the analysis of the optimal forecasting model. The results indicated that the number of neurons of the hidden layer is 18 and learning rate is 0.03. When the number of iterations is 5,000, test samples' RMSE of 0.0723 is the least value, indicating the optimal training model. See Table 3.

Table 3: The Results of BPN Forecast model

Hidden layer neurons	Learning rate	iterations	Training samples RMSE	Test samples RMSE
14	0.009	16000	0.032934	0.075178
15	0.009	16000	0.032953	0.073822
16	0.005	30000	0.032748	0.074817
17	0.005	25000	0.032891	0.073569
18	***0.03***	***5000***	***0.03291***	***0.0723***

4.3 SVR

The SVR model adopts the first 84 months of data as training samples and the last 24 months of data as test samples. Simple validation with try and error are adopted to search for the optimal combination of input parameters and variables. The search scope for the three parameters C,γ,ε are: 2^{-31},2^{-30},2^{-29},2^{-28},….,2^{31}. The results obtained are shown in Table 3. When C is 2^{29}, γ is 2^{-29}, ε is 2^{-1}, and test samples' mean square error 796.2980 is the least value. Thus, this model is optimal0. See Table 4:

Table 4: Forecasting result of SVR in Different Parameter Combinations.

NO	C	γ	ε	MSE	NO	C	γ	ε	MSE
1	2^{31}	2^{-19}	2^{5}	11594.22	6	2^{26}	2^{-26}	2^{3}	1848.33
2	2^{30}	2^{-26}	2^{6}	2007.28	7	2^{25}	2^{-26}	2^{-1}	2004.48
3	$\boldsymbol{2^{29}}$	$\boldsymbol{2^{-29}}$	$\boldsymbol{2^{-1}}$	***796.30***	8	2^{24}	2^{-24}	2^{-30}	2302.92
4	2^{28}	2^{-29}	2^{1}	2287.01	9	2^{23}	2^{-23}	2^{1}	2387.28
5	2^{27}	2^{-26}	2^{-23}	1194.00	10	2^{22}	2^{-21}	2^{-2}	3122.48

4.4 The Results of Measuring the Forecast Errors

4.4.1 Formulas for measuring the forecast errors

It is very important to evaluate the results of the correctness of tourism demand forecasting model. Many different methods can be used to measure the forecast error, and the selection of an appropriate method to measure forecast error affects the evaluation of tourism demand forecasting model and the selection a good tourism demand forecasting model. Li et al. (2005) studied the formulas for measuring the forecast errors and found that the most frequently used is MAPE and the less frequently used is RMSE. MAPE, RMSE, and MAD are meant to measure the deviation of the forecast value from the actual value, such that the smaller the value of the three indexes, the closer the forecast value is to the actual value. Therefore, this paper adopts MAPE, RMSE, and MAD to evaluate the accuracy of the forecast error of each model. The formulas are shown as follows:

A_i be the Actual value in period i; F_i the forecast value in period i; and n the number of periods used in the calculation.

Mean absolute percentage error,

$$\text{MAPE}=\frac{1}{n}\sum_{i}^{n}\left|(A_i - F_i)/A_i\right| \tag{17}$$

root-mean-squared error,

$$\text{RMSE}=\sqrt{\frac{1}{n}(\sum_{i}^{n}(A_i - F_i)^2/F_i)} \tag{18}$$

Mean absolute deviation,

$$\text{MAD}=\frac{1}{n}\sum_{i}^{n}\left|A_i - F_i\right| \tag{19}$$

4.4.2 The results of measuring the forecast errors

The three kinds of forecast models compare the errors of the forecasted number of tourists and the actual number of tourists in 2011 and 2012. The results obtained show that the

SVR model has the lowest error rates with respect to RMSE, MAD, and MAPE; thus, it has the best performance; ANN model had the second lowest error rates while MARS model had the highest error rates; thus, it had the worst.

Table 5: Comparison of the Error Values of the Three Models

	RMSE	MAD	MAPE
MARS	82105.88	67724.26	11.2634
BPN	50585.58	40994.47	7.0809
SVR	28219.99	20729.55	3.6122

5 Discussion

This paper adopted MARS, ANN, and SVR to develop tourism demand econometric forecasting models based on the monthly data of tourists to Taiwan and compared the forecasted results. The results showed that SVR model's forecasted values are RMSE=28219.99, MAD=20729.55, and MAPE=3.61%, which is the best outcome; ANN model's forecasted values are RMSE=50585.58, MAD=40994.47, and MAPE=7.08%, which is the second best outcome; and MARS model's forecasted values are RMSE=82105.88, MAD=67724.26, and MAPE=11.26%, which is the worst outcome.

This paper found that SVR model provides a better forecasted result compared to ANN model, which is consistent with the findings of Chen and Wang (2007) on tourism demand in mainland China based on time-series model. Another finding of this paper is that ANN model provides a better forecasted result than does MARS model, which is consistent with the findings of Lin *et al.* (2011) based on time-series model; however, it is not consistent with the findings of De Veaux *et al.* (1993b) who found that in most cases, MARS model provided a better forecasted result compared to ANN model. Therefore, regarding the forecasted result of time-series model and econometric model for tourism demand forecasting, SVR model provided the best outcome with a MAPE of 3.61%, followed by ANN model with a MAPE of 7.08%, which is also low, and finally MARS model with a MAPE of 11.26%.

This paper adopted the total number of tourists in the tourist destination as a dependent variable. The explanatory variables did not consider the variables of individual countries of origin but only those of the destination. Hence, the explanatory variables included the tourist destination's consumption price and accommodation fee every night, the expenses associated with traveling to the destination, the destination's the exchange rate, the number of hotel rooms, the passenger capacity of international flights, as well as the monthly/seasonal index of the number of tourists.

The important variables in the MARS analysis were GDP (Y), tourist consumption price (CPI) (P), exchange rate (EX), the number of hotel rooms (HR), the passenger capacity of international flights (AS), and monthly/seasonal index (SE). It can be seen that among these 6 variables, GDP, CPI, and EX relate to tourist consumption price and all have a negative effect on tourism demand. HR and AS relate to tourism supply and have a positive effect on tourism demand. Thus, the main factors affecting tourism demand are tourist consumption price, the amount of tourism supply, and seasonal changes.

6 Conclusions

The development of tourism has been receiving increasing attention from countries around the world, and in order to increase the number of tourists, a country must invest in the construction of tourism in all aspects, including traffic, tourist hotels, and tourist and recreational facilities. The correctness of such investment entails correct forecast of tourism demand. This paper adopted the destination country's tourist consumption price and accommodation fee per night, the expenses associated with traveling to the destination, the number of hotel rooms, the passenger capacity of international flights in the destination, as well as monthly/seasonal index as explanatory variables. The number of tourists was used as a dependent variable to make up the econometric model for tourism demand forecast. MARS, ANN, and SVR were adopted to develop the econometric model and a two-year forecast was made. The results indicated that SVR model is the optimal model with the mean error rate of 3.61% suggesting a good model fit.

The findings indicate that the main factors affecting tourism demand are tourist consumption price, the amount of tourism supply, and seasonal changes. Tourist consumption price variables, i.e., GDP, CPI, and exchange rate, had a negative effect on tourism demand, while tourism supply variables, i.e., the number of hotel rooms and the passenger capacity of international flights had a positive effect on tourism demand. That is to say, the increase in the destination country's GDP, CPI, and exchange rate will reduce tourism demand, and the increase in a SVR model to solve the problem the number of hotel rooms and the passenger capacity of international flights will increase tourism demand. Therefore, to increase tourism demand, a country needs to increase the number of hotel rooms and the passenger capacity of international flights. Regarding the tourism demand forecast model, SVR had the lowest error rate of 3.61%, making it the optimal forecast model. ANN turned out to be the sub-optimal forecast model while MARS turned out to be the worst forecast model.

The data in this paper are Taiwan's monthly data. Since GDP data were quarterly data, they were converted into monthly data for the analysis, which had some effect on the correctness of analysis results. Explanatory variables adopted in this paper are mostly proxy variables, which also have some effect on the correctness of forecasted results. Taiwan is an island country, and international tourists to Taiwan take planes as the main means of transport. Nevertheless, tourists from some inland countries, like Switzerland, may also take trains or buses; thus, their means of transport are different from that in this paper. Accordingly, the explanatory variables of such means of transport are also different.

Future researchers can consider screening out important explanatory variables through MARS and then analyzing them through ANN and SVR model. They can also consider comparing the forecasting accuracy of the forecast model of all explanatory variables with that of important explanatory variables. Furthermore, a potential future research would consider tourism demand forecasting through MARS, ANN and SVR based on an econometric model made up of the GDP of main countries of origin of tourists and the tourist consumption price, exchange rate, the number of hotel rooms, and the passenger capacity of international flights of the destination country.

References

[1] S. F. Witt, and C. A. Witt, Forecasting tourism demand: A review of empirical research, *International Journal of Forecasting*, **11**, (1995), 447-475.

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

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

[4] K.-H. Huarng, L. Moutinho, and T. H.-K. Yu, An Advanced Approach to Forecasting Tourism Demand in Taiwan, *Journal of Travel & Tourism Marketing,* **21(4)**, (2006), 15-24.

[5] C. J. Lin, H. F. Chen, and T. S. Lee, Forecasting Tourism Demand Using Time Series, Artificial Neural Networks and Multivariate Adaptive Regression Splines: Evidence from Taiwan, *International Journal of Business Administration,* **2 (2)**, (2011), 14-24.

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

[7] G. Li, H. Song, and S. F. Witt, Recent developments in econometric modeling and forecasting. *Journal of Travel Research,* **44**, (2005), 82–99.

[8] R. Law, Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting, *Tourism Management,* **21**, (2000), 331-340.

[9] L. Moutinho, K.-H. Huarng, T. H.-K. Yu, and C.-Y. Chen, Modeling and forecasting tourism demand: the case of flows from Mainland China to Taiwan, *Service Business,* **2,** (2008), 219-232

[10] H. A., Bicak, M. Altinay, and H. Jenkins, Forecasting tourism demand of North Cyprus. *Journal of Hospitality and Leisure Marketing,*. **12**, (2005), 87–99.

[11] K. N. Lee, Forecasting long-haul tourism demand for Hong Kong using error correction models, *Applied Economics,* **43**, (2011), 527–549.

[12] H. Song, and S. F. Witt, Forecasting international tourist flows to Macau. *Tourism Management*, **27,** (2006), 214–224.

[13] K. K. F. Wong, H. Song, and K. S. Chon, Bayesian models for tourism demand forecasting, *Tourism Management*, **27,** (2006), 773–780.

[14] R. Law, C. Goh, and R. Pine, Modeling tourism demand: A decision rules based approach, *Journal of Travel & Tourism Marketing,* **16**, (2004), 61–69.

[15] Y.-S. Wang, The impact of crisis events and macroeconomic activity on Taiwan's international inbound tourism demand, *Tourism Management*, **30,** (2009), 75–82.

[16] R. D. De Veaux, A. L. Gordon, J. C. Comiso, and N. E. Bacherer, Modeling of Topographic Effects on Antarctic Sea Ice Using Multivariate Adaptive Regression Splines, *Journal of Geophysical Research,* **98(C11),** (1993), 20307-20319.

[17] R. D. De Veaux, D. C. Psichogios, and L. H. Ungar, A Comparison of Two Nonparametric Estimation Schemes: MARS and Neural Networks, *Computers Chemical Engineering,* **17**, (1993), 819–837.

[18] J. G. De Gooijer, B. K. Ray, and H. Krager, Forecasting Exchange Rates Using Time Series MARS, *Journal of International Money and Finance,* **17 (3),** (June 1998), 513-534.

[19] R. Law, and N. Au, A neural network model to forecast Japanese demand for travel to Hong Kong, *Tourism Management.* **20**, (1999), 89-97.

[20] A. Palmer, J. J. Montano, and A. Sese, Designing an artificial neural network for forecasting tourism time-series. *Tourism Management*, **27**, (2006), 781–790.

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

[22] H. Drucker, C. J. C. Burges, L. Kaufman, A. Smola, and V. N. Vapnik, Support Vector Regression Machines, *Advances in Neural Information Processing Systems,* **9,** (1997), 155-161.

[23] V. N. Vapnik, S. Golowich, and A. J. Smola, Support vector method for function approximation, regression estimation, and signal processing. *Advances in Neural Information Processing Systems,* **9**, (1997), 281-287.

[24] C.-J. Lu, T.-S. Lee, and C.-C. Chiu, Financial time series forecasting using independent component analysis and support vector regression. *Decision Support Systems,* **47**, (2009), 115–125.

[25] L. Wang, and J. Zhu, Financial market forecasting using a two-step kernel learning method for the support vector regression. *Annals of Operations Research,* **174(1)**, (Feb 2010), 103-120.

[26] A. Viacaba, M. H. Poursaeidi, and O. E. Kundakcioglu, Natural Gas Price Forecasting via Selective Support Vector Regression. *Proceedings of the 2012 Industrial and Systems Engineering Research Conference.* Orlando, Florida, (May 2012), 1-9.

[27] C. Lim, and M. McAleer, Cointegration analysis of quarterly tourism demand by Hong Kong and Singapore for Australia. *Applied Economics,* **33**, (2001), 1599–1619.

[28] H. Song, K. K. F. Wong, and K. K. S. Chon, Modelling and forecasting the demand for Hong Kong tourism. *International Journal of Hospitality Management,* **22,** (2003), 435–451.

[29] T. Garin-Munoz, Inbound international tourism to Canary Islands: a dynamic panel data approach, *Tourism Management,* **27,** (2006), 281–291.

[30] J. H. Friedman, Multivariate Adaptive Regression Splines (with discussion), *Annals of Statistics,* **19,** (1991), 1-141.

[31] D. Steinberg, B. Bernard, C. Phillip, and M. Kerry, *MARS User Guide,* Salford Systems, San Diego, CA., 1999.

[32] P. Craven, and G. Wahba, Smoothing Noisy Data with Spline Functions. Estimating the Correct Degree of Smoothing by the Method of Generalized Cross-Validation, *Numberische Mathematik,* **31**, (1979), 317-403.

[33] J. A. Freeman, and D. M. Skapura, *Neural Networks Algorithms, Applications, and Programming Techniques,* Addison-Wesley Publishing Company, 1992.

[34] G. Zhang, B. E. Patuwo, and M. Y. Hu, Forecasting with Artificial Neural Networks: TheState of the Art, *International Journal of Forecasting,* **14,** (1998), 35-62.

[35] M. J. A. Berry, and G. Linoff, *Data Mining Technique for Marketing, Sale, and Customer Support,* Wiley Computer, 1997.

[36] C. C. Chiu, Y. J. Shao, and T. S. Lee, Identification of Process Disturbance Using SPC/EPC and Neural Networks, *Journal of Intelligent Manufacturing,* **14(3)**, (2003), 379-388

[37] K. E. Fish, J. H. Barnes, and M. W. Aiken, Artificial Neural Networks: A New Methodology for Industrial Market Segmentation, *Industrial Marketing Management,* **24,** (1995), 431-438.

[38] T. S. Lee, and C. C. Chiu, Neural Network Forecasting of an Opening Cash Price Index, *International Journal of Systems Science,* **33(3)**, (2002), 229-237

[39] T. S. Lee, C. C. Chiu, C. J. Lu, and I. F. Chen, Credit Scoring Using the Hybrid Neural Discriminant Technique, *Expert Systems with Applications*, **23(3)**, (2002), 245-254.

[40] M. T. Leung, A. S. Chen, and H. Daouk, Forecasting Exchange Rates Using General Regression Neural Networks, *Computers and Operations Research*, **27,** (2000), 1093-1110.

[41] A. Vellido, P. J. G. Lisboa, and J. Vaughan, Neural Networks in Business: A Survey of Applications (1992-1998), *Expert Systems with Applications*, **17,** (1999), 51-70.

[42] C. W. Hsu, and C. J. Lin, and C. C. Chang, *A practical guide to support vector classification*, Technical Report, Department of Computer Science and Information Engineering, National Taiwan University, 2003.

[43] C. Cortes, and V. N. Vapnik, Support vector networks, *Machine Learning*, **20,** (1995), 273-297.