Available online at www.sciencedirect.com

ScienceDirect

CAAI Transactions on Intelligence Technology 1 (2016) 30–42

http://www.journals.elsevier.com/caai-transactions-on-intelligence-technology/

Original article

Forecasting tourism demand by extracting fuzzy Takagi–Sugeno rules from trained SVMs

Xin Xu [a,*], Rob Law [b], Wei Chen [b], Lin Tang [c]

[a] College of Mechatronics and Automation, National University of Defense Technology, Changsha, 410073, China

[b] School of Hotel and Tourism Management, Hong Kong Polytechnic University, Hong Kong Special Administrative Region, China

[c] College of Science, Hunan Agriculture University, Changsha, 410003, China

Available online 4 June 2016

Abstract

Tourism demand forecasting has attracted substantial interest because of the significant economic contributions of the fast-growing tourism industry. Although various quantitative forecasting techniques have been widely studied, highly accurate and understandable forecasting models have not been developed. The present paper proposes a novel tourism demand forecasting method that extracts fuzzy Takagi–Sugeno (T–S) rules from trained SVMs. Unlike previous approaches, this study uses fuzzy T–S models extracted from the outputs of trained SVMs on tourism data. Owing to the symbolic fuzzy rules and the generalization ability of SVMs, the extracted fuzzy T–S rules exhibit high forecasting accuracy and include understandable pre-condition parts for practitioners. Based on the tourism demand forecasting problem in Hong Kong SAR, China as a case study, empirical findings on tourist arrivals from nine overseas origins reveal that the proposed approach performs comparably with SVMs and can achieve better prediction accuracy than other forecasting techniques for most origins. The findings demonstrated that decision makers can easily interpret fuzzy T–S rules extracted from SVMs. Thus, the approach is highly beneficial to tourism market management. This finding demonstrates the excellent scientific and practical values of the proposed approach in tourism demand forecasting.

Keywords: Fuzzy modeling; Rule extraction; Support vector machines; Tourism demand forecasting

1. Introduction

Forecasting is an essential requirement for decision making and policy planning. Forecasting is widely conducted in various fields, including the tourism industry. Over the past decade, an increasing number of studies have focused on forecasting techniques for tourism demand [48,22,32]. The importance of accurate forecasting is basically attributed to the perishable nature of products and services in the industry. For instance, vacant airline seats, unoccupied hotel rooms, and unsold event tickets cannot be stockpiled for future use. Thus, accurate short-term and long-term forecasts of future demand are crucial [20,35,46]. Such forecasts are necessary for Hong Kong SAR, China, a key travel destination in Asia with an economically significant tourism industry. This major industry underwent substantial changes in market segments for inbound tourists. These changes demonstrate the urgent need to develop accurate methods for forecasting international demand for travel to Hong Kong SAR, China, which can be quantified by the number of tourist arrivals. Thus, tourism researchers continue to develop various techniques to predict the future demand for tourism.

In the tourism industry, accuracy and good comprehensibility of forecasting are required from policy makers and practitioners. As far as industrial applications are concerned, tourism practitioners can check the predicted values of tourist arrivals from different origins and plan for a change in demand from specific market segments by obtaining highly accurate estimates of such demand. By interpreting forecasting models, policy makers can analyze the key factors that contribute to

* Corresponding author.

E-mail address: xinxu@nudt.edu.cn (X. Xu).

Peer review under responsibility of Chongqing University of Technology.

http://dx.doi.org/10.1016/j.trit.2016.03.004

the increase or decrease in tourism demand from various regions. These practitioners can understand the underlying regularities according to the comprehensibility of tourism forecasting models. Policy makers can also plan tourism projects and related infrastructure development activities accurately and reasonably.

Previous studies on tourism demand analysis and forecasting produced two types of results in terms of forecasting quality. One type focused on error magnitude accuracy and the other aimed to improve understandability based on symbolic rules or knowledge bases. Despite the consensus on the accuracy of forecasting and the clear understanding of the advantages of accurate forecasts, few tourism forecasting methods have been developed that outperform other methods both in terms of accuracy and comprehensibility. According to Walle [57], qualitative models may exhibit good interpretability, but these approaches usually lack generalization ability, which substantially limits their applications. Traditional statistical regression models, such as auto-regressive integrated moving-average (ARIMA) approaches, also have a certain degree of comprehensibility. However, the prediction accuracy of these techniques may be unsatisfactory when nonlinearity and noise exist in tourism demand data. Artificial neural networks (ANNs) and support vector machines (SVMs) have recently gained significant interest because of their generalization ability and forecasting accuracy. SVMs have been demonstrated to perform better than ANNs and ARIMA. For example, Chen and Wang [11] combined a genetic algorithm with a support vector regression (SVR) to model tourist arrivals in China from 1985 to 2001. Their study shows that this approach outperforms ANNs and ARIMA models based on the normalized mean square error and mean absolute percentage error (MAPE). SVMs are a class of machine-learning algorithms based on the structural risk minimization principle [11]. The generalizability of SVMs can be optimized by controlling structural complexity, which makes SVMs superior to other machine-learning and data-mining algorithms. Moreover, the SVM training procedure is a convex quadratic programming process through which a global optimal solution can be obtained. Thus, compared with previous approaches, SVMs provide a more accurate and flexible forecasting technique for tourism forecasting.

Although SVMs are computationally accurate and exhibit satisfactory performance in tourism studies, both SVMs and ANNs are basically "black-box" techniques with poor explanatory capability and comprehensibility. The knowledge gleaned from such techniques is difficult to understand. Thus, the rule extraction from SVMs or ANNs was recently studied [6,10] and applied to various domains, such as credit scoring and fraud detection [16]. By extracting the rules from SVMs and ANNs, the comprehensibility of these black-box models can be enhanced, and a compromise between forecasting accuracy and interpretability can be achieved. Although forecasting accuracy and good comprehensibility are essential for policy makers and practitioners in the tourism industry, existing studies do not incorporate symbolic rule extraction and SVMs into tourism forecasting.

To fill this research gap, a novel tourism demand forecasting method is proposed in this paper, which is based on support vector machines with rule extraction (SVMRE). This method can extract fuzzy T–S rules from SVMs trained on tourism demand data. The aim of the present study is to incorporate extracted fuzzy rules from highly accurate SVMs into tourism demand forecasting. The fuzzy T–S rules generated from the outputs of SVMs can verify the information encoded in these models. Thus, the fuzzy rules for tourism demand extracted from SVM models exhibit high forecasting accuracy and easy comprehensibility to industry practitioners. The tourism demand forecasting problem in Hong Kong SAR, China was studied as an application case. It is demonstrated that that the proposed approach performs comparably with SVMs and can achieve better prediction accuracy than other forecasting techniques in most cases. In addition, decision makers can easily interpret the fuzzy T–S rules extracted from SVMs. Thus, the approach is highly beneficial to tourism market management. This finding shows the excellent scientific and practical values of the proposed approach in tourism demand forecasting.

The rest of this paper is organized as follows. In Section 2, some research background is introduced. In Section 3, SVMRE approach is presented for Tourism Demand Forecasting. In Section 4, the tourism demand forecasting problem in Hong Kong SAR is used as a case study and the performance of the SVMRE is evaluated and compared with other popular techniques for tourism demand forecasting.

2. Research background

2.1. Tourism forecasting techniques

Accurate forecasts are crucial because of the unique nature of the tourism industry [19,29,31,34]. Tourism demand forecasting employs qualitative and quantitative approaches [20,28,48]. Qualitative approaches depend on substantial information and human experiences. Walle [57] criticized these techniques for their lack of generalizability. As a result, tourism researchers do not primarily use qualitative forecasting methods. Formal scientific techniques that unambiguously represent the relationship between demand for travel and its underlying factors are more useful than qualitative forecasting methods in helping tourism decision makers understand the travel demand for a given destination.

Quantitative tourism demand forecasting models adopt mathematical functions to form the relationships of certain phenomena using numeric data [15,38]. These models are used to estimate future values based on past performance. Quantitative tourism forecasting approaches include causal relationship (regression) and time series techniques [1,26,30,50,58]. Although these approaches have achieved a certain degree of success, one fundamental problem is their inability to predict changes associated with other determining factors.

Causal relationship techniques establish the relationship among multiple variables via statistical analysis [2,13,49]. These techniques have the advantage of explicitly representing the relationships that are evident in reality, assisting decision makers in assessing alternative plans, and accommodating a wide range of relationships. Multivariate regression forecasting models generally exhibit high degrees of explanatory power and prediction accuracy. However, these models also have limitations, including a large amount of time and financial resources involved as well as substantial skills required to establish correct relationships.

Researchers have also developed other tourism demand forecasting techniques based on multivariate regression analyses, such as gravity models that measure the degree of interaction between two geographical areas. The success of computer systems that simulate the human nervous system has drawn the attention of tourism researchers; initial research was conducted to investigate the feasibility of incorporating computerized neural systems in tourism demand analyses [27,54].

2.2. Rule extraction from ANNs and SVMs in tourism demand analysis

SVMs represent one of the main advancements in statistical learning theory for data forecasting and classification [18,40,60]. The SVM training procedure is a convex quadratic programming process that eliminates the local minima problem in ANNs [14,37,42]. Although preliminary work was conducted to use computational intelligence techniques such as ANNs, SVMs, and Gaussian processes in tourism demand analysis [59], prior studies employed black-box modeling methods, which have weak explanatory capability. This weakness in comprehensibility is the primary obstacle in applying neural networks and SVM-based models in tourism demand analysis.

The main motivation for rule extraction from ANNs or SVMs is the improved representation of a well-trained intermediate model, which is achieved by filtering out noise in samples [5,56]. Symbolic rule extraction approaches, which can be classified into decomposition-based and learning-based approaches, are applied mostly to ANNs [4]. Decomposition-based approaches disassemble the architecture of the trained neural network and achieve rule extraction using searching techniques. Examples include similar weight clustering by Towell and Shavlik [53], separation of the activation values of relevant hidden units by Setiono and Leow [45], and two-stage procedures by Ishikawa [24]. By contrast, learning-based approaches consider the network as a whole and attempt to extract rules that can explain their functions. For example, Schmitz et al. [44] constructed binary decision trees from trained neural networks by adopting an attribute selection criterion. Similarly, Tanaka et al. [52] extracted linguistic rules from a trained network.

Some studies were conducted to determine the symbolic rules for SVM models. For instance, Núñez et al. [41] introduced a rule extraction approach for SVMs known as the SVM + prototype method. However, the approach is unsuitable for larger input spaces because complex rules that lack interpretability would be extracted. To facilitate the applicability of SVM rule extraction methods to large-scale problems, Barakat and Diederich [6] proposed a learning-based approach using two different datasets. The present study focuses on fuzzy rule extraction from SVM models for tourism demand analysis. In this paper, we provide innovations on new fuzzy rule extraction models for SVMs in tourism demand forecasting. These models exhibit better generalization ability than other tourism demand analysis models and have the advantage of comprehensibility, which addresses the drawbacks of existing black-box models.

2.3. Typical tourism demand analysis models

In this study, five time-series forecasting models and four causal relationship models were tested to forecast tourist arrivals from nine major tourist-generating origins for comparison. The former models include naive, moving average, single exponential smoothing, double exponential smoothing, and ARIMA models, which are some of the most commonly used models in tourism demand forecasting [48]. The latter models include multiple regression, ANN, SVM, and SVMRE. The operations of the selected models are briefly delineated as follows.

A simple model assigns a value of a at time t, which is the same as the value of a' at time $t+1$ where a' and a represent the forecasting value and the actual arrival numbers, respectively; that is,

$$a'_{t+1} = a_t. \tag{1}$$

Similarly, the moving average (3) model employed in this study calculates the value of a at time period $t+1$ by

$$a'_{t+1} = (a_t + a_{t-1} + a_{t-2})/3. \tag{2}$$

A single exponential smoothing (0.3) model predicts the value of arrival a at time $t+1$ by

$$a'_{t+1} = 0.3a'_t + 0.7a_t. \tag{3}$$

A double exponential smoothing model (also known as Brown's one-parameter adaptive method) was applied to forecast arrivals with the following formulas:

$$L_t = \alpha A_t + (1-\alpha)(L_{t-1} + b_{t-1}) \tag{4}$$

$$b_t = \alpha(L_t - L_{t-1}) + (1-\alpha)b_{t-1}, \tag{5}$$

$$a'_{t+h} = L_t + hb_t \tag{6}$$

where α is the smoothing constant between 0 and 1 and h is the number of time periods to be forecasted. In our case, $h = 1$.

According to Athiyaman and Robertson [3], the four time-series-based models are relatively easy to implement and can generate relatively accurate predictions for international tourism demand. An ARIMA model is a complex time-series forecasting technique that includes five different phases, namely, preparation, identification, estimation, diagnostic checking, and forecasting. An ARIMA(p,d,q) process is obtained by integrating an ARMA(p,q) process, where p, d, and q are positive integers or zeros. Given a time series of data X_t, an ARMA(p,q) model has the following form:

$$\left(1-\sum_{i=1}^{p}\phi_i L^i\right)X_t=\left(1+\sum_{i=1}^{q}\theta_i L^i\right)\varepsilon_t \tag{7}$$

where θ_i, ϕ_i are the parameters of the moving average part and the autoregressive part, respectively. L is the lag operator and ε_t denotes the error terms.

Thus, an ARIMA model can be formulated as follows:

$$\left(1-\sum_{i=1}^{p}\phi_i L^i\right)(1-L)^d X_t=\left(1+\sum_{i=1}^{q}\theta_i L^i\right)\varepsilon_t \tag{8}$$

where d is a positive integer.

We also applied these models in our experiments because of the good performance of ARIMA models in time-series forecasting models. The ARIMA function in the SPSS trend module was used to forecast the arrival values in the destination. The expert modeler in this module determines the best parameters in ARIMA models (p,d,q) to obtain the best forecasting results.

A multiple linear regression model takes the following form:

$$Y_i=\beta_0+\beta_1 X_{i1}+\beta_2 X_{i2}+\ldots+\beta_p X_{ip}+\varepsilon_i,\quad i=1,2,\ldots,n \tag{9}$$

where (Y_i,X_{i1},X_{i2},…,X_{ip}) are random samples.

In the present study, a neural network-based model contained six input nodes, which were used to predict the tourist arrivals to the output node. Three sigmoid nodes existed in the hidden layer. The output nodes were non-threshold linear units because of the numeric nature of the class. The learning rate was set to 0.01, and the maximum number of iterations was set to 50,000 based on different trials. The momentum applied to the weights during update was set to 0.2, which was the default value. The preceding parameters are manually selected to improve the performance of the method.

In recent years, SVMs exhibited improved prediction accuracy and generalization performance compared with that of traditional methods not only in terms of classification but also in regression tasks. In the present study, SVMs were applied in regression problems. A general SVM framework with extracted rules or SVMRE is also presented for tourism demand forecasting. On the basis of this framework, the SVM regression method was used to construct prediction models with good generalization ability where the structure selection and weight learning of the SVM can be implemented via global optimization techniques. By adopting fuzzy rule extraction, the SVM model can be transformed into symbolic fuzzy rules to achieve improved comprehensibility.

3. SVMRE for tourism demand forecasting

To combine the advantages of SVMs and fuzzy rules, the proposed framework for fuzzy rule extraction from SVMs includes three main stages, as depicted in Algorithm 1. Since the fuzzy Takagi–Sugeno (T–S) model has been shown to have advantages in fast training speed and simple rule structures, we use the data generated from trained SVMs to construct a fuzzy T–S rule set for improved comprehensibility.

Algorithm 1 The SVMRE algorithm

1. input: Normalized training data set ($\mathbf{x}_i$, y_i), (i=1,2,…,n).
2. SVM Training on the training data set:
   2.1) Construct the objective function

$$\min_{\varpi}\frac{1}{2}\|\omega\|^2+C\frac{1}{n}\sum_{i=1}^{n}L(y_i,f(\mathbf{x}_i,\omega))$$

   2.2) Solve the optimization problem in a kernel-induced dual space and obtain the following estimated function:

$$\hat{f}(\mathbf{x})=\sum_{i=1}^{n_{SV}}(\alpha_i-\alpha_i^*)K(\mathbf{x}_i,\mathbf{x})\quad \text{s.t. } 0\le\alpha_i^*\le C,\ 0\le\alpha_i\le C$$

3. Data regeneration based on the trained SVM model:
   3.1) Generate SVM outputs z_i (i = 1,2,…,m) on the inputs in the training data set ($\mathbf{x}_i$, y_i) (i = 1,2,…,m);
   3.2) Generate SVM outputs z_i (i = 1,2,…,n) on the inputs in the testing data set ($\mathbf{s}_i$, y_i) (i = 1,2,…,n);
4. Fuzzy T-S model learning:
   4.1) Combine the preceding two subsets as a new training set;
   4.2) Train the fuzzy T-S model on the newly generated training set.
5. output:
   the trained fuzzy T-S model;

3.1. Support vector regression for tourism demand forecasting

The first stage in the SVMRE framework involves the SVM regression on the training data. At this stage, the training samples can be employed to tune the SVM hyper-parameters, such as the kernel parameters.

To formulate tourism demand forecasting as a regression problem, the following statistical model for data generation is considered:

$$y = f(\mathbf{x}) + \delta, \tag{10}$$

where $f(\mathbf{x})$ is an unknown continuous-valued function with m-dimensional input $\mathbf{x}$, and δ is an additive zero mean noise with noise variance σ^2. The task of the regression is to estimate unknown function $f(\mathbf{x})$ by collecting a finite number of noisy training data samples $(\mathbf{x}_i, y_i)$, $(i = 1,2,\ldots,n)$.

In traditional regression-based forecasting methods, which includes ANNs, the estimated function is usually constructed by minimizing an empirical error function on the training data. The formula is given as

$$E\left[\widehat{f}(\mathbf{x})\right] = \sum_i h\left(y_i - \widehat{f}(\mathbf{x}_i)\right), \tag{11}$$

where $\widehat{f}(\mathbf{x})$ is the estimated function and $h: R \to R$ is the error function.

After estimated function $\widehat{f}(\mathbf{x})$ is constructed, the forecasted value of a new input, $\mathbf{x}_{i+1}$ and the corresponding prediction error e_{i+1} can be obtained as follows:

$$\widehat{y}_{i+1} = \widehat{f}(\mathbf{x}_{i+1}) \quad \text{and} \tag{12}$$

$$e_{i+1} = y_{i+1} - \widehat{f}(\mathbf{x}_{i+1}). \tag{13}$$

The forecasting accuracy of the estimated function is closely related to the generalization ability of the regression method, which was widely studied in statistical learning. According to the research results of statistical learning theory, the generalization ability of a learning machine is determined by minimizing the empirical errors and the structural complexity of the machine. Thus, improving the generalization ability or forecasting accuracy of traditional regression methods is difficult when the empirical error function defined in (10) is only minimized.

SVM regression or SVR differs from traditional regression methods, such as ANNs, because structural and empirical risks are minimized [14]. Using the kernel trick [55], the objective function for an SVR takes the following form:

$$\frac{1}{2}\|\omega\|^2 + C\frac{1}{n}\sum_{i=1}^{n} L(y_i, f(\mathbf{x}_i,\omega)), \tag{14}$$

where $L(y, f(\mathbf{x},\omega))$ is the loss function on training data. In SVR, the following ε-insensitive loss function proposed by Vapnik [55] is commonly adopted:

$$L_\varepsilon(y, f(\mathbf{x},\omega)) = \begin{cases} 0 & \textit{if}\, |y - f(\mathbf{x},\omega)| \le \varepsilon \\ |y - f(\mathbf{x},\omega)| - \varepsilon & \textit{otherwise} \end{cases}, \tag{15}$$

where ε is the size of the insensitivity margin.

Cristianini and Schölkopf [14] demonstrated that the optimization problem can be solved dually. The solution is given by

$$\widehat{f}(\mathbf{x}) = \sum_{i=1}^{n_{SV}} \left(\alpha_i - \alpha_i^*\right) K(\mathbf{x}_i, \mathbf{x}) \quad \text{s.t. } 0 \le \alpha_i^* \le C,\ 0 \le \alpha_i \le C, \tag{16}$$

where n_{SV} is the number of support vectors (SVs), α and α^* are Lagrange multipliers, and $K(\mathbf{x},\mathbf{x}_i)$ is the kernel function. Smola and Schölkopf [47] delineated the details on the optimization process for computing solutions of the SVR.

3.2. Data regeneration from trained SVMs

In the second stage, the SVM model is adopted to generate new data samples for the training of fuzzy rules. The idea of data regeneration and symbolic rule learning based on trained ANNs was used by Zhou and Jiang [62], who proposed the NeC4.5 algorithm. In their algorithm, a neural network ensemble was initially trained on the original training set. The final output of the NeC4.5 is a C4.5 decision tree trained on the new training set. The comprehensibility of the NeC4.5 is better than that of the ANN ensemble because of the symbolic rules expressed by the decision tree.

In this study, the SVM model trained in the first stage of the SVMRE approach was used to generate new training samples, which allows the generalization ability of the SVM to be employed to train fuzzy rules. For tourism demand forecasting, three subsets of new training samples can be generated and combined because the input variables of both the training and testing samples are available.

Let $(\mathbf{x}_i, y_i)$ $(i = 1,2,\ldots,m)$ denote the original training data set and $(\mathbf{s}_i, y_i)$ $(i = 1,2,\ldots,n)$ denote the original testing data set. The first subset is denoted as $(\mathbf{x}_i, z_i)$ $(i = 1,2,\ldots,m)$. The outputs are generated by considering input vectors $\mathbf{x}_i$ in the original training samples as the inputs to the SVM model as follows:

$$z_i = \widehat{f}(\mathbf{x}_i,\omega) = \sum_{j=1}^{n_{SV}} \omega_j g_j(\mathbf{x}_i) + b. \tag{17}$$

Similarly, the second subset $(\mathbf{s}_i, z_i)$ $(i = 1,2,\ldots,n)$ is generated by replacing the outputs y_i $(i = 1,2,\ldots,n)$ in the testing samples with the predicted outputs of the SVM model. The generation of the second subset aims to use the generalization ability of trained SVMs to obtain predicted values on testing inputs. Then, the first and second subsets can be combined as an expanded training set to construct the fuzzy T–S model so that the prediction ability of the fuzzy T–S model can be improved. To further improve the generalization ability of the fuzzy rules, an additional subset can also be constructed by computing the predicted outputs of the SVM model on some randomly generated or selected input vectors.

3.3. Fuzzy rule extraction from trained SVMs

After the new training data samples were generated, a fuzzy rule extraction method using the T–S fuzzy model [51] was employed to extract symbolic fuzzy rules from the trained SVMs. In principle, a T–S fuzzy inference system has three main mechanisms, namely, the rule base with inference machines, membership functions of input variables, and consequence parts of each rule. The consequence part of a T–S inference system is a linear equation or constant coefficient that corresponds to a first-order or zero-order Sugeno inference system, respectively.

A first-order Sugeno inference system is considered in the present study. We denote l as the number of inputs. A typical fuzzy rule can be expressed as follows:

$$\text{Rule } i\text{: IF } x_1 \text{ is } A_{1i} \text{ and } x_2 \text{ is } A_{2i}, \ldots, x_l \text{ is } A_{li}, \text{ THEN } f_i = \sum_{j=1}^{l} p_{ij} x_j + r_i, \tag{18}$$

where x_i $(i = 1,2,\ldots,l)$ is the crisp input to node i, A_{ji} $(j = 1,2,\ldots,l)$ is the linguistic label described by the membership functions, N is the total number of fuzzy rules, and p_{ij} $(i = 1,2,\ldots,N, j = 1,2,\ldots,l)$ is the consequence parameters.

The final output of the Sugeno fuzzy system for a given input vector $\mathbf{x}$ is expressed as follows:

$$\widehat{h}(\mathbf{x}) = \frac{\sum_{i=1}^{N} z_i f_i}{\sum_{i=1}^{N} z_i}, \tag{19}$$

where z_i is the firing strength of rule i, which can be computed as

$$z_i = \min\left\{\mu_{Aji}(x_i),\ j = 1,2,\ldots,l\right\}. \tag{20}$$

The training of a T–S fuzzy system includes three steps. The first step is determining the membership functions of the input variables using a clustering algorithm to generate the centers of the fuzzy membership functions. The second step is constructing fuzzy rules. The third step is identifying the consequence parameters. Once the centers of the fuzzy membership functions are computed, the T–S fuzzy rules can be directly constructed by designing a fuzzy rule according to each clustering center. The consequence parameters in the T–S fuzzy rules can be estimated based on the work of Jang [25]. In the hybrid-learning algorithm, the gradient descent method is adopted to assign the nonlinear parameters of the input membership functions. The linear output parameters (p_{ij}, r_i) are computed via the least-squares method.

Clustering methods are used to determine the cluster centers and the number of fuzzy rules. Thus, these approaches are critical to the performance of T–S fuzzy systems. Two popular clustering methods, namely, subtractive clustering [12,61] and fuzzy C-means (FCM), can be used to generate the cluster centers and the number of rules. The FCM algorithm was proposed by Dunn [17] and improved by Bezdek [7]; this algorithm has been widely used in the literature. In our experiments, the subtractive clustering algorithm was employed to determine appropriate clusters for the input and output data.

4. Experimental results

In this study, the data for analysis were collected based on availability and validity. The data encompass nine origins in North America, Europe, and Asia. These origins included the US, Australia, Canada, France, Germany, the UK, Japan, Korea, and Taiwan, China. The trends in tourist arrivals in Hong Kong SAR from these origins are shown in Fig. 1.

Previous studies [20,21,33,35,36,48] indicated that the independent variables in forecasting models for international tourism demand mainly include the following:

- Population of the origin
- Real disposable personal income of the origin
- Marketing expenditure on promotional activities by the destination
- Cost of living in the destination
- Foreign exchange rate between the origin and the destination
- Relative price of tourism services in the destination

Thus, the demand for travel to the destination which is measured by the number of tourist arrivals from these origins, can be represented as

- $$Arrival_i = f(SP_i, FER_i, Pop_i, Mkt, GDE_i, AHR), \tag{21}$$

where *Arrival* is the number of tourist arrivals from *origin*$_i$. *SP* is the relative price of services relative to *origin*$_i$ (measured as the ratio of the consumer price index [CPI] to the CPI in *origin*$_i$). *FER* is the foreign exchange rate (measured in *Currency*$_i$/US$). *Pop* is the population in *origin*$_i$ (measured in numbers). *Mkt* is the promotional and marketing expenditure by the destination's tourism industry (measured in US$). *GDE* is the real gross domestic expenditure per capita in *origin*$_i$ (measured in US$). *AHR* is the average hotel room rate in the destination (measured in US$).

SP for year *i* was calculated similar to the method of Carey [9] and Morley [39].

$$SP_t = \frac{CPI_t(Destination)/CPI_{2000}(Destination)}{CPI_t(Origin_i)/CPI_{2000}(Origin_i)}. \tag{22}$$

The CPI figures in our dataset were chained values relative to the base year 2000. Thus, the formula is

$$SP_t = \frac{CPI_t(Destination)}{CPI_t(Origin_i)}. \tag{23}$$

Relevant data for forecasting tourism demand for travel to Hong Kong SAR from the aforementioned origins from 1967 to 2007 were taken from a number of sources including the World Bank Group, Hong Kong SAR Tourist Association, Hong Kong SAR Tourism Board, and Statistics Department of the Hong Kong SAR Government. To eliminate the effects of the

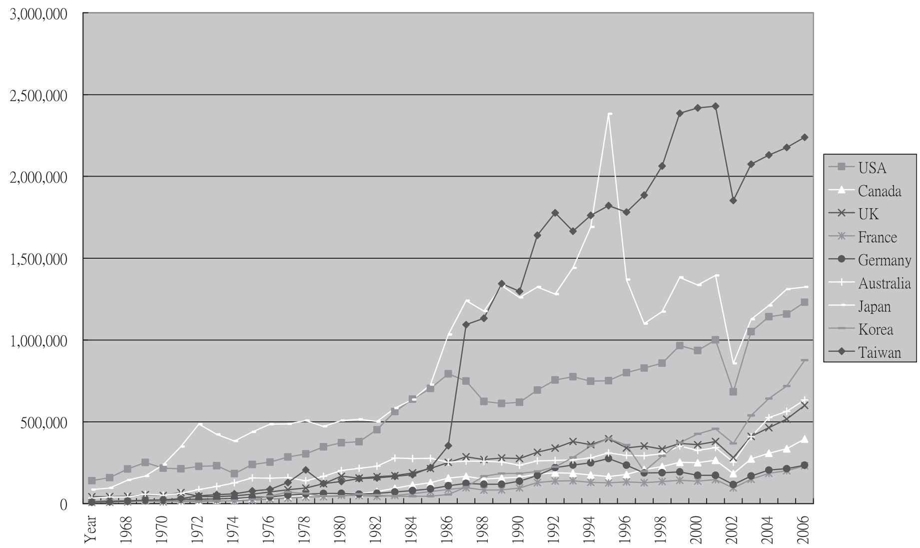


Fig. 1. Tourist arrivals from nine origins to Hong Kong SAR.

dimensions of the input variables and to improve the learning effectiveness of the ANN, SVM, and SVMRE models, we standardized or normalized the input and target variables before data analysis [43]. Scaling for original data points helps improve forecasting accuracy. The adopted standardization process is expressed as follows:

$$S_t = \frac{X_t - X_{\min}}{X_{\max} - X_{\min}}. \tag{24}$$

In (24), X_t is the input or target variable at time t, X_{max} is the maximum of X_t during the period of our dataset, and X_{min} is the minimum of X_t. We rescaled the data used for the experiments within the range [0, 1] by applying these formulas to the input and output variables.

In all of the forecasting experiments using SVMs, Gaussian or polynomial kernel functions were selected. The four parameters selected for the SVM algorithm were insensitive parameter ε in the loss function, bound C on the Lagrangian multipliers, width parameter of the Gaussian kernels or degree for the polynomial kernels, and conditioning parameter λ for quadratic programming methods. In our implementation, λ and ε were set to positive numbers. The radial basis function kernel width and the bound on the Lagrangian multipliers were selected manually based on performance evaluations.

When the T−S fuzzy modeling approach was used to extract fuzzy rules from the trained SVMs, a Sugeno-type fuzzy inference system was generated using subtractive clustering. The radius parameter of subtractive clustering was set to 0.5 for all experiments. The premise and consequence parameters of the T−S fuzzy model were identified based on the training data generated by the SVM forecasting model. The rule extraction method first uses the subtractive algorithm to determine the rule number and the antecedent membership functions. The output of the T−S fuzzy model is linear with respect to the consequence parameters. Thus, the least-squares estimation method can be used to determine the consequent parameters of each rule. The fuzzy modeling approach was discussed further by Chiu [12].

Since the objective of forecasting models based on machine learning is to realize prediction abilities on new data, the data set is divided into two subsets. One is for training the forecasting model and the other is for performance testing. The design of machine learning algorithms needs to improve the approximation precision both in the training set and in the testing set. In the experiments, to test the accuracy of these forecasting models, we divided the datasets into training sets (80%) and testing sets (20%). The training sets consisted of the arrival data from 1967 to 1999. The testing sets consisted of the arrival data from 2000 to 2007. The former was used to build the forecasting models, and the latter was intended to evaluate the accuracy of the various models. The forecasting quality of the nine forecasting approaches was measured in terms of the MAPE and mean absolute scaled error (MASE), which were calculated using the following formula:

$$MAPE = \frac{\sum_{i=1}^{n} |a_i - p_i| / a_i}{n} \cdot 100\%, \tag{25}$$

where a_i is the actual tourist arrivals and p_i stands for the predicted value of tourist arrivals.

MASE is a scaled error proposed by Hyndman and Koehler [23]. MASE scaled the errors based on the in-sample mean absolute error (MAE) from the naive (random walk) forecasting method. A scaled error is defined as:

$$q_t = \frac{e_t}{\frac{1}{n-1}\sum_{i=2}^{n} |Y_i - Y_{i-1}|}. \tag{26}$$

where Y_i is the real output, e_t is the prediction error defined as $e_t = Y_t - F_t$, and F_t is the predicted output.

The MASE is defined as

$$\mathrm{MASE} = \mathrm{mean}(|\mathrm{q_t}|). \tag{27}$$

Tables 1 and 2 compare the MAPE and MASE values as the results of the different forecasting models on the data from the nine origins. The MAPE indicates the deviation between the predicted and actual tourist arrivals. Thus, a smaller MAPE

Table 1
Comparison of MAPE for nine origins (The bold numbers denote the best performance among different forecasting methods).

Origin	SVMRE	ANN	Multiple regression	ARIMA	Double exponential smoothing (0.3)	Single exponential smoothing (0.3)	Moving average (3)	Naïve
France	**15.84%**	22.04%	17.59%	17.40%	20.63%	19.74%	21.17%	19.27%
Canada	**12.57%**	25.75%	**12.54%**	16.31%	17.59%	17.99%	20.57%	16.88%
Australia	**19.85%**	24.28%	29.76%	23.60%	20.64%	20.35%	22.88%	**19.63%**
Germany	**13.65%**	22.70%	46.29%	18.83%	32.23%	17.83%	20.4 2%	16.58%
Japan	**13.69%**	17.33%	15.63%	17.83%	18.35%	15.49%	17.44%	17.07%
UK	**9.31%**	16.59%	25.73%	17.48%	18.00%	16.84%	16.84%	16.28%
Taiwan	**6.91%**	8.31%	9.11%	14.27%	9.37%	8.49%	11.17%	8.12%
Korea	**17.58%**	31.62%	20.22%	36.74%	17.71%	21.34%	25.70%	18.10%
US	**10.82%**	28.04%	11.26%	15.19%	16.72%	17.62%	19.64%	17.54%

Table 2
Comparison of MASE for nine origins (The bold numbers denote the best performance among different forecasting methods).

Origin	SVMRE	ANN	Multiple Regression	ARIMA	Double Exponential Smoothing (0.3)	Single Exponential Smoothing (0.3)	Moving Average (3)	Naïve
Australia	**1.04**	1.46	1.53	1.41	1.12	1.06	1.27	**0.96**
Canada	**0.70**	1.56	0.76	1.01	1.02	1.05	1.24	0.95
France	**0.84**	1.17	0.97	0.87	1.07	0.98	1.11	0.90
Germany	**0.74**	1.27	3.07	0.98	1.92	0.98	1.17	0.89
Japan	**0.89**	1.11	1.01	1.14	1.20	0.94	1.11	1.04
UK	**0.52**	1.03	1.44	1.14	1.14	1.02	1.19	0.94
Taiwan	**1.08**	1.34	1.49	2.12	1.32	1.24	1.67	1.16
Korea	1.10	1.99	1.25	2.31	**1.07**	1.19	1.48	**0.99**
US	**0.63**	1.84	**0.63**	0.88	0.94	1.01	1.16	0.97

value implies increased forecasting accuracy. MASE is defined as the error scaled by the MAE from naive forecast method. Thus, the proposed forecasting methods produce smaller errors than that of the naive method when the value of MASE is less than 1. The results indicate that the forecasting output of the SVMRE model is accurate with a relatively small degree of error. Except for the result for Australia, the SVMRE models outperformed the other models in forecasting accuracy.

According to the findings in Tables 1 and 2, the SVMRE models in almost all the datasets attained the lowest MAPEs among the nine forecasting models. To predict the results for Australian arrivals, the MAPE values of the SVMRE models are only slightly higher than the corresponding numbers of the naive method, but lower than those of other methods. Based on the comparison of MASE, the same conclusion could be drawn. MASE is less sensitive to outliers and less variable on small samples than MdASE [23]. We used the results of MASE in this study because of the relatively small sample size of the testing sets. The MAPE values of the SVMRE were the same in all cases because of the approximation ability of T–S fuzzy models. These results confirm the good generalization ability of SVMs with structural risk minimization based on statistical learning theory. Statistical significance tests (the Wilcoxon signed test) have been conducted to compare the statistical difference between the forecasting results of two different methods. We compared the predicted values for each model with actual tourist arrivals. The predicted values of SVMRE were also compared with those of other models. The results of the significance test are presented in Tables 3 and 4. These tables indicate that the SVMRE exhibited significant difference in forecasting performance compared with that of other methods in many cases.

These tables also demonstrate that symbolic fuzzy rules can be extracted from trained SVR models without compromising prediction precision by using the T–S fuzzy modeling approach in the SVMRE. These rules are more readily understandable than that of SVMs, which allows policy makers to understand fuzzy prediction models more easily and facilitate planning and decision making. For example, nine T–S fuzzy rules were extracted from the trained SVM model to predict the UK arrivals. We denote A_{ki} ($k = 1,2,\ldots,6$) as the linguistic variables or fuzzy membership functions of input x_i. Based on the fuzzy prediction model for these arrivals, the T–S fuzzy rules take the following form:

Rule i:
If (x_1 is A_{1i}) and (x_2 is A_{2i}) and (x_3 is A_{3i}) and (x_4 is A_{4i}) and (x_5 is A_{5i}) and (x_6 is A_{6i}),
then $y_i = p_1x_1 + p_2x_2 + p_3x_3 + p_4x_4 + p_5x_5 + p_6x_6 + r_i.$

Once the T–S fuzzy rules were obtained, designing symbolic rules for the tourism demand forecasting problem was easy. The fuzzy membership functions were constructed for the input variables. Thus, symbolic variables can be defined for these functions. During our implementation, the fuzzy membership functions of each input

Table 3
Wilcoxon signed-rank test between predicted arrivals (using different methods) and actual arrivals.

Origin	SVMRE vs. arrival	ANN vs. arrival	Regression vs. arrival	ARIMA vs. arrival	Double exponential smoothing vs. arrival	Single exponential smoothing vs. arrival	Moving average vs. arrival	Naive vs. arrival
Australia	−0.560	−1.680	−0.845	−1.820	−1.755	−1.752	−1.960	−1.400
	0.575	0.093**	0.398	0.069**	0.079**	0.080**	0.050**	0.161
Canada	−0.420	−1.820	−0.700	−1.540	−1.404	−1.402	−1.680	−1.402
	0.674	0.069	0.484	0.123	0.160	0.161	0.093**	0.161
France	−0.980	−0.700	−1.521	−0.420	−1.051	−1.334	−1.825	−1.400
	0.327	0.484	0.128	**0.674**	0.293	0.182	0.068**	0.161
Germany	−0.280	−2.380	−2.380	−0.840	−0.211	−0.560	−0.491	−0.631
	0.779	0.017***	0.017***	0.401	**0.833**	0.575	0.623	0.528
Japan	−0.280	−0.141	−0.070	−1.260	−0.420	−1.185	−1.263	−1.122
	0.779	**0.888**	0.944	0.208	0.674	0.236	0.206	0.262
UK	−0.280	−0.280	−2.527	−1.680	−1.120	−1.521	−1.521	−1.400
	0.779	**0.779**	0.012***	0.093**	0.263	0.128	0.128	0.161
Taiwan	−0.840	−1.442	−1.829	−1.687	−0.771	−1.409	−0.771	−1.192
	0.401	0.149	0.067**	0.092	**0.441**	0.159	0.441	0.233
Korea	−2.380	−2.521	−1.400	−2.521	−1.960	−2.240	−2.380	−1.893
	0.017***	0.012***	**0.161**	0.012***	0.050**	0.025***	0.017***	0.058**
US	−0.771	−1.960	−0.421	−1.400	−1.402	−1.402	−1.400	−1.260
	0.441	0.050**	**0.674**	0.161	0.161	0.161	0.161	0.208

The first row of each origin is the Z value of the Wilcoxon signed-rank test, and the second row is the asymp. sig. (two-tailed) of the test.
*** = significant at the 0.05 level; ** = significant at the 0.1 level (The bold numbers denote the best performance among different forecasting methods).

variable were defined based on the center positions of the Gaussian membership functions by the following five symbolic variables:

Very large, *Large*, *Medium*, *Small*, *and Very small*

To predict the UK arrivals, the following T−S fuzzy symbolic rules were constructed:

Rule 1:
If (*SP* is very large) and *(FER* is very small) and (*Pop* is very small) and (*Mkt* is very small) and (*GDE* is very small) and (*AHR* is very small),
then $Arrival = 0.04221*SP + 0.113*FER + 0.189*Pop + 0.2502*Mkt + 0.3763*GDE + 0.303*AHR - 0.1329$.

Rule 2:
If (*SP* is large) and (*FER* is large) and (*Pop* is medium) and (*Mkt* is small) and (*GDE* is medium) and (*AHR* is large),
then $Arrival = 0.0928*SP + 0.04937*FER + 0.2057*Pop + 0.3252*Mkt + 0.1376*GDE + 0.2706*AHR - 0.00365$.

Rule 3:
If (*SP* is very small) and (*FER* is small) and (*Pop* is small) and (*Mkt* is very small) and (*GDE* is very small) and (*AHR* is small),
then $Arrival = 0.02424*SP + 0.05774*FER - 0.004358*Pop - 0.02861*Mkt + 0.1443*GDE + 0.1694*AHR + 0.02203$.

Rule 4:
If (*SP* is medium) and (*FER* is large) and (*Pop* is medium) and (*Mkt* is medium) and (*GDE* is large) and (*AHR* is very small),
then $Arrival = -0.05025*SP + 0.05392*FER + 0.03842*Pop + 0.1951*Mkt + 0.1276*GDE + 0.372*AHR + 0.2732$.

Rule 5:
If (*SP* is very small) and (*FER* is medium) and (*Pop* is small) and (*Mkt* is small) and (*GDE* is medium) and (*AHR* is large),
then $Arrival = 0.2317*SP - 0.01644*FER + 0.2598*Pop + 0.05078*Mkt + 0.3558*GDE + 0.07281*AHR + 0.06992$.

Rule 6:
If (*SP* is small) and (*FER* is medium) and (*Pop* is very large) and (*Mkt* is large) and (*GDE* is very large) and (*AHR* is medium),
then $Arrival = 0.04923*SP + 0.1982*FER + 0.1509*Pop + 0.01057*Mkt + 0.2491*GDE + 0.1917*AHR + 0.1464$.

Rule 7:
If (*SP* is small) and (*FER* is very large) and (*Pop* is small) and (*Mkt* is small) and (*GDE* is small) and (*AHR* is medium),
then $Arrival = 0.1642*SP + 0.03257*FER + 0.03902*Pop + 0.3137*Mkt + 0.1733*GDE + 0.166*AHR - 0.01604$.

Rule 8:
If (*SP* is small) and (*FER* is medium) and (*Pop* is small) and (*Mkt* is very small) and (*GDE* is very small) and (*AHR* is very small),
then $Arrival = -0.03721*SP + 0.04701*FER - 0.01845*Pop + 0.1066*Mkt - 0.05008*GDE - 0.01654*AHR + 0.02875$.

Rule 9:
If (*SP* is medium) and (*FER* is medium) and (*Pop* is medium) and (*Mkt* is medium) and (*GDE* is medium) and (*AHR* is medium),
then $Arrival = 0.09033*SP + 0.07145*FER + 0.07509*Pop + 0.1147*Mkt + 0.09799*GDE + 0.1684*AHR + 0.1389$.

Table 4
Wilcoxon signed-rank test between predicted arrivals of SVMRE and those of other models.

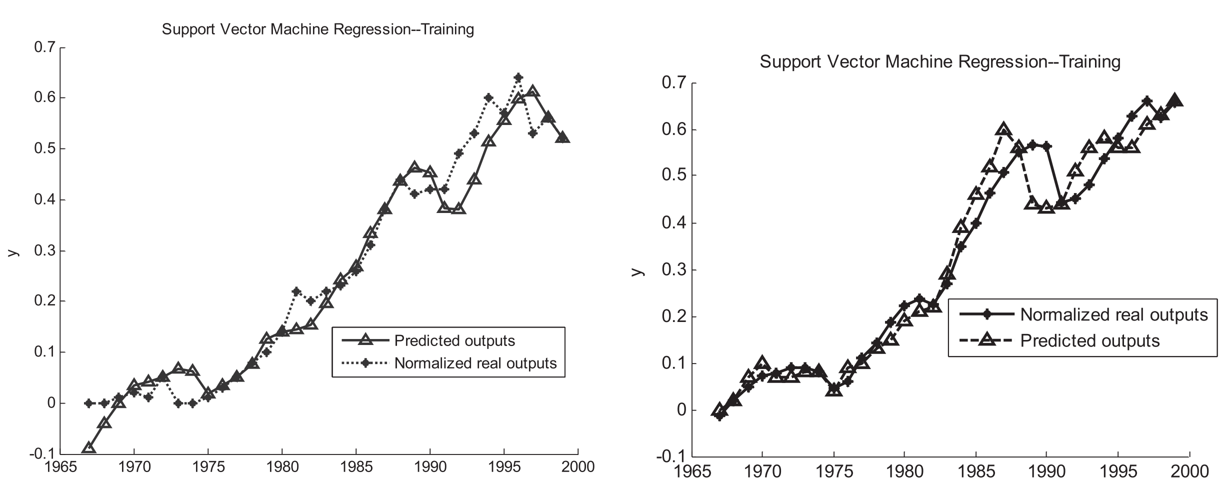
Origin	SVMRE vs. ANN	SVMRE vs. regression	SVMRE vs. ARIMA	SVMRE vs. double exponential smoothing	SVMRE vs. single exponential smoothing	SVMRE vs. moving average	SVMRE vs. naive
Australia	−2.521	−0.980	−2.521	−1.680	−1.260	−1.680	−0.700
	0.012***	0.327	**0.012*****	**0.093****	0.208	**0.093****	0.484
Canada	−2.521	−0.560	−2.521	−2.380	−1.820	−2.380	−1.120
	0.012***	0.575	**0.012*****	**0.017*****	**0.069****	**0.017*****	0.263
France	−0.840	−1.540	−1.120	−0.420	−0.140	−1.120	−0.560
	0.401	0.123	0.263	0.674	0.889	0.263	0.575
Germany	−2.521	−2.521	−2.240	0.000	−0.280	−0.560	−0.280
	0.012***	**0.012*****	**0.025*****	1.000	0.779	0.575	0.779
Japan	−0.280	−0.420	−2.380	−0.840	−0.840	−1.120	−0.420
	0.779	0.674	**0.017*****	0.401	0.401	0.263	0.674
UK	−1.120	−2.521	−2.521	−1.540	−2.380	−2.380	−2.100
	0.263	**0.012*****	**0.012*****	0.123	**0.017*****	**0.017*****	**0.036*****
Taiwan	−2.240	−2.521	−2.521	−2.380	−0.420	−0.140	−0.420
	0.025***	**0.012*****	**0.012*****	**0.017*****	0.674	0.889	0.674
Korea	−2.521	−0.420	−2.521	−1.260	−0.420	−1.820	−0.980
	0.012***	0.674	**0.012*****	0.208	0.674	**0.069****	0.327
USA	−1.120	−2.524	−0.700	−0.423	−0.140	−0.981	−0.423
	0.263	**0.012*****	0.484	0.672	0.889	0.326	0.672

The first row of each origin is the Z value of the Wilcoxon signed-rank test and the second row is the asymp. sig. (two-tailed) of the test.
*** = significant at the 0.05 level; ** = significant at the 0.1 level (The bold numbers denote the best performance among different forecasting methods).

These symbolic fuzzy rules significantly improve the understandability of the prediction model, which is beneficial to decision making and policy planning in the tourism industry. Furthermore, the prediction accuracy of the extracted fuzzy rules is higher than that of other methods, such as ANNs and ARIMA. This outcome is attributed to the generalization ability of the SVMs and the approximation ability of the T−S fuzzy modeling method. Fig. 2 presents the predictive results of the SVMs for the training data of UK arrivals as well as that of US arrivals. Fig. 3 shows that the prediction outputs of the foregoing extracted T−S fuzzy results on the test data of the UK and US arrivals. In the proposed SVMRE approach, SVMs are used only on the training data so the prediction results are plotted in Fig. 2. The final output model of the SVMRE approach is evaluated on the testing data to show its prediction ability on new data that are not used for building the model. In the experiments, the proposed SVMRE method can achieve good prediction accuracy for almost the entire dataset by extracting a small number of fuzzy symbolic rules from the SVM model. The results indicate that the proposed method for symbolic rule extraction from SVMs not only exhibits a high level of prediction accuracy but also improves the comprehensibility of the SVMs. In the following section, the managerial implications of the extracted fuzzy rules from SVMs are analyzed and discussed.

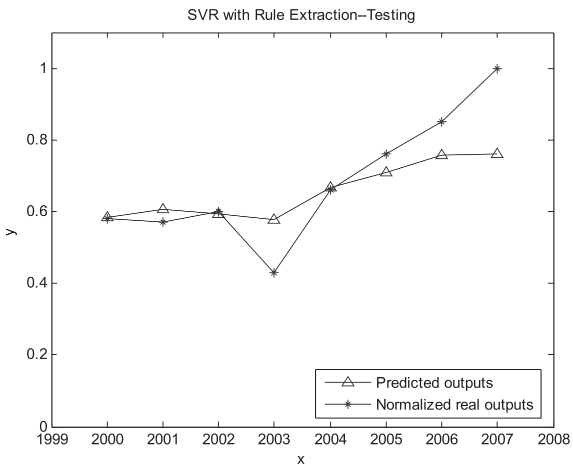
5. Analysis and discussions

Aside from the high forecasting accuracy of the fuzzy rules extracted from trained SVMs, the interpretability of the fuzzy

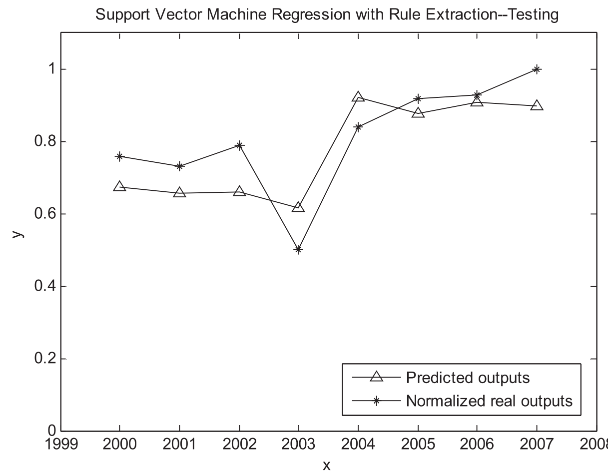


(a) Prediction of UK arrivals (training) (b) Prediction of US arrivals (training)

Fig. 2. Predictive results of SVMs for training data. x = year; y = normalized or standardized tourist arrivals.



(a) Prediction of UK arrivals (testing)



(b) Prediction of US arrivals (testing)

Fig. 3. Predictive results of SVMRE for testing data. x = year; y = normalized tourist arrivals.

symbolic rules model can benefit policy makers in the tourism industry. Based on the fuzzy rules in forecasting tourist arrivals, the main influential factors for the decision variables under different conditions can be determined. Policy makers can receive suggestions to observe or adjust the influential factors to effectively predict or control tourist arrivals. We adopt the fuzzy forecasting model for Canadian arrivals to illustrate the managerial implications of the proposed method. The understandability of the SVMRE model for other countries is a straightforward extension of the following case.

Based on the eight rules for predicting the Canadian arrivals, the main factors that influence tourist arrivals under eight different conditions can be determined. Several rules may be activated for a certain condition determined by the input variables. However, only a local rule is usually activated with the maximum firing strength. Thus, this rule can be employed as a symbolic rule for data analysis and policy making. Rule 1 is used for demonstration as follows:

If (*SP* is very small) and *(FER* is very small) and (*Pop* is very small) and (*Mkt* is very small) and (*GDE* is very small) and (*AHR* is small),

then $Arrival = 0.05182^* SP + 0.0166^* FER + 0.1566^* Pop + 0.2463^* Mkt - 0.0337^* GDE - 0.08318^* AHR - 0.006227.$

This rule implies that when the observation variables are very small, tourist arrivals from Canada are mainly influenced by the two factors of the population in $origin_i$ (Pop) and the promotional and marketing expenditure by the destination's tourism industry (*Mkt*). These observation variables include the relative price of services in the destination relative to $origin_i$, foreign exchange rate, population in $origin_i$, promotional and marketing expenditure by the destination's tourism industry, and real gross domestic expenditure per capita in $origin_i$, as well as the average hotel room rate in the destination. This step was adopted because the coefficients for the two factors or variables mentioned earlier were 0.1566 and 0.2463, which were significantly higher than those for other factors.

When Rule 6 is taken as another example, the following form is obtained:

Rule 6: If (*SP* is medium) and (*FER* is small) and (*Pop* is very large) and (*Mkt* is large) and (*GDE* is very large) and (*AHR* is medium),

then $Arrival = 0^* SP + 0^* FER + 0^* Pop + 2.477^* Mkt + 0^* GDE + 1.828^* AHR -2.393.$

Under the conditions described by the premise, i.e., if (*SP* is small) and (*FER* is medium) and (*Pop* is very large) and (*Mkt* is large) and (*GDE* is very large) and (*AHR* is medium), tourist arrivals would be mainly determined by the two factors of *Mkt* and *AHR*.

Considering the preceding examples, we can determine that the fuzzy rules extracted from trained SVMs are easy to understand. The form of conditional rules is more straightforward than numeric parameters in traditional forecasting techniques. Thus, the interpretability of the fuzzy rules extracted from trained SVMs is appropriate and valuable for the tourism industry. Managers and policy makers could also choose the rules that are directly applicable to them because several fuzzy rules are generated from one SVM model. In other words, the failure of one rule does not necessarily lead to poor performance of the other rules, which is more meaningful than either success or failure results of traditional forecasting methods. Given the dynamic global environment at present, this feature of SVMRE could help decision makers improve their business plans.

6. Conclusions

The major contribution of the present study is the SVMRE approach to increase the understandability and accuracy of tourism demand forecasting. Thus far, no prior research on tourism demand forecasting using SVMs with fuzzy rule extraction has been reported. We also present the empirical results and managerial implications of SVMRE, which demonstrates that SVMRE can obtain the best forecasting precision with good understandability for most of the datasets from different regions. The functions and performance of traditional forecasting approaches have existed for a long time and have been broadly exploited to the extent that these approaches may have reached their peak. Thus, even minor improvements in tourism demand forecasting require a substantial amount of investment using traditional approaches. This condition implies the need for evolutionary, if not revolutionary, scientific modeling techniques. In a review of tourism demand forecasting articles published since 2000, Song and Li [48] proposed the development of innovative approaches to improve forecasting accuracy. The introduction of SVMRE in this study contributes to this need.

Although the SVMRE method demonstrated excellent forecasting accuracy for the datasets of tourist arrivals included in this study, further work is needed to generalize the research findings. Comprehensive future testing of SVMRE with datasets in different travel destinations can provide additional insights to help international research and professional communities enhance their understanding of the relationships among SVMRE, traditional forecasting methods, and tourism demand.

References

[1] G. Athanasopoulos, R.A. Ahmad, R.J. Hyndman, Int. J. Forecast. 25 (1) (2008) 146–166.
[2] G. Athanasopoulos, R.J. Hyndman, Tour. Manag. 29 (1) (2008) 19–31.
[3] A. Athiyaman, R.W. Robertson, Int. J. Contemp. Hosp. Manag. 4 (4) (1992) 8–11.
[4] G. Atsalakis, E. Chnarogiannaki, C. Zopounidis, Tourism demand forecasting based on a neuro-fuzzy model, in: Hospitality, Travel, and Tourism: Concepts, Methodologies, Tools, and Applications, Business Science Reference, USA, 2014, pp. 106–115.
[5] B. Baesens, T. van Gestel, S. Viaene, M. Stepanova, J. Suykens, J. Vanthienen, J. Oper. Res. Soc. 54 (6) (2003) 627–635.
[6] N. Barakat, J. Diederich, Learning-based rule-extraction from support vector machines: performance on benchmark datasets, in: N. Kasabov, Z.S.H. Chan (Eds.), Proc. of the Conference on Neurocomputing and Evolving Intelligence, Knowledge Engineering and Discovery Research Institute (KEDRI), Auckland, New Zealand, 2004, pp. 247–252.
[7] J.C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, 1981.
[9] K. Carey, Atl. Econ. J. 19 (3) (1991) 32–40.
[10] J.L. Castroa, L.D. Flores-Hidalgob, C.J. Mantasa, J.M. Puchea, Fuzzy Sets Syst. 158 (18) (2007) 2057–2077.
[11] K.Y. Chen, C.H. Wang, Tour. Manag. 28 (1) (2007) 215–226.
[12] S. Chiu, J. Intell. Fuzzy Syst. 2 (3) (1994) 267–278.
[13] M.P. Clements, D.F. Hendry, Forecasting Economic Time Series, Cambridge University Press, Cambridge, 1998.
[14] N. Cristianini, B. Schölkopf, AI Mag. 23 (3) (2002) 31–42.
[15] K. Dalrymple, K. Greenidge, Ann. Tour. Res. 26 (1) (1999) 188–191.
[16] J. Diederich (Ed.), Rule Extraction from Support Vector Machines, Studies in Computational Intelligence, vol. 80, Springer, Heidelberg, Berlin, 2008.
[17] J.C. Dunn, Cybern. Syst. Int. J. 3 (3) (1973) 32–57.
[18] R. Fan, P. Chen, C. Lin, J. Mach. Learn. Res. 6 (2005) 1889–1918.
[19] D.C. Frechtling, Practical Tourism Forecasting, Butterworth-Heinemann, Jordan Hill, Oxford, 1996.
[20] D.C. Frechtling, Forecasting Tourism Demand: Methods and Strategies, Butterworth-Heinemann, Oxford, 2001.
[21] P. Gonzalez, P. Moral, Int. J. Contemp. Hosp. Manag. 10 (1) (1995) 41–43.
[22] E. Hadavandi, A. Ghanbari, K. Shahanaghi, et al., Tour. Manag. 32 (5) (2011) 1196–1203.
[23] R.J. Hyndman, A.B. Koehler, Int. J. Forecast. 22 (4) (2006) 679–688.
[24] M. Ishikawa, Neural Netw. 13 (10) (2000) 1171–1183.
[25] J.S.R. Jang, IEEE Trans. Syst. Man Cybern. 23 (3) (1993) 665–685.
[26] R. Law, J. Hosp. Leis. Mark. 6 (4) (2000a) 17–29.
[27] R. Law, Tour. Manag. 21 (3) (2000b) 331–340.
[28] R. Law, J. Travel Tour. Mark. 10 (2/3) (2001) 47–65.
[29] R. Law, N. Au, Tour. Manag. 20 (1) (1999) 89–97.
[30] R. Law, C. Goh, R. Pine, J. Travel Tour. Mark. 16 (2/3) (2004) 61–69.
[31] R. Law, H. Mok, C. Goh, Data mining in tourism demand analysis: a retrospective analysis, in: Advanced Data Mining and Applications, Lecture Notes in Computer Science, Springer Berlin, Heidelberg, 2007, pp. 508–515.
[32] T. Law, T. To, C. Goh, Int. J. Hosp. Manag. 27 (3) (2008) 346–354.
[33] G. Li, H. Song, S.F. Witt, J. Travel Res. 44 (1) (2005) 82–99.
[34] C. Lim, Tour. Econ. 3 (1) (1997a) 69–81.
[35] C. Lim, Ann. Tour. Res. 24 (4) (1997b) 835–849.
[36] C. Lim, J. Travel Res. 37 (3) (1999) 273–284.
[37] C.J. Lin, Neural Comput. 13 (2) (2001) 307–317.
[38] S. Makridakis, S.C. Wheelwright, R.J. Hyndman, Forecasting Methods and Applications, John Wiley & Sons, New York, 1998.
[39] C.L. Morley, J. Tour. Stud. 4 (1) (1993) 19–25.
[40] M. das C. Moura, E. Zio, I.D. Lins, et al., Reliab. Eng. Syst. Saf. 96 (11) (2011) 1527–1534.
[41] H. Núñez, C. Angulo, A. Catala, Rule-extraction from support vector machines, in: Proc. of European Symposium on Artificial Neural Networks, Bruges, 2002, 2002, pp. 107–112.
[42] P.F. Pai, W.C. Hong, Ann. Tour. Res. 32 (4) (2005) 1138–1141.
[43] S. Saarinen, R. Bramley, G. Cybenko, SIAM J. Sci. Comput. 14 (3) (1993) 693–714.
[44] G.P.J. Schmitz, C. Aldrich, F.S. Gouws, IEEE Trans. Neural Netw. 10 (6) (1999) 1392–1401.
[45] R. Setiono, W.K. Loew, Appl. Intell. 12 (1/2) (2000) 15–25.
[46] E. Smeral, A. Weber, Ann. Tour. Res. 27 (4) (2000) 982–1006.
[47] A.J. Smola, B. Schölkopf, Stat. Comput. 14 (3) (2004) 199–222.
[48] H. Song, G. Li, Tour. Manag. 29 (2) (2008) 203–220.
[49] H. Song, S.F. Witt, Tourism Demand Modelling and Forecasting: Modern Econometric Approaches, Elsevier Science, Kidlington, Oxford, 2000.
[50] H. Song, K.K.F. Wong, K.K.S. Chon, Int. J. Hosp. Manag. 22 (4) (2003) 435–451.
[51] T. Takagi, M. Sugeno, IEEE Trans. Syst. Man Cybern. 15 (1) (1985) 116–132.
[52] K. Tanaka, M. Nii, H. Ishibuchi, Learning from linguistic rules and rule extraction for function approximation by neural networks, in: Simulated Evolution and Learning, Lecture Notes in Computer Science, Springer Berlin, Heidelberg, 1999, pp. 317–324.
[53] G. Towell, J. Shavlik, Mach. Learn. 13 (1) (1993) 71–101.
[54] M. Uysal, M.S.E. Roubi, J. Travel Res. 38 (2) (1999) 111–118.
[55] V. Vapnik, Statistical Learning Theory, John Wiley & Sons, New York, 1998.
[56] S. Viaene, R. Derrig, B. Baesens, G. Dedene, J. Risk Insur. 69 (3) (2002) 433–443.
[57] A.H. Walle, Ann. Tour. Res. 24 (3) (1997) 524–536.
[58] S.F. Witt, C.A. Witt, Int. J. Forecast. 11 (3) (1995) 447–475.
[59] Q. Wu, R. Law, X. Xu, Expert Syst. Appl. 39 (5) (2012) 4769–4774.

[60] X. Xu, X. Wang, An adaptive network intrusion detection method based on PCA and support vector machines, in: Advanced Data Mining and Applications, Lecture Notes in Computer Science, Springer Berlin, Heidelberg, 2005, pp. 696−703.
[61] R. Yager, D. Filev, J. Intell. Fuzzy Syst. 2 (3) (1994) 209−219.
[62] Z.-H. Zhou, Y. Jiang, IEEE Trans. Knowl. Data Eng. 16 (6) (2004) 770−773.

Data sources

[63] Hong Kong Tourism Board, Annual Report of the Hong Kong Tourism Board, HKTB, Hong Kong, 2000-2007.
[64] Statistics Department, Commissioner of Inland Revenue Annual Review, Hong Kong SAR Government, 1967-1996.
[65] Hong Kong Tourist Association, Hong Kong Hotel Industry, HKTA, Hong Kong, 1988-2007.
[66] Hong Kong Tourist Association, Hong Kong Tourist Association Annual Report, HKTA, Hong Kong, 1951-2007.
[67] Hong Kong Tourist Association, Statistical Review of Tourism in Hong Kong, HKTA, Hong Kong, 1976-1997.
[68] Hong Kong Tourist Association, Visitor Arrival Statistics, HKTA, Hong Kong, 1980-2007.
[69] World Bank Group, World Development Indicators Online, World Bank Group, Washington, DC, 1967-2007. Retrieved from, http://ddp-ext.worldbank.org/ext/DDPQQ/member.do?method=getMembers&userid=1&queryId=6.