

## Research Article

# Particle Swarm Optimization-Based Support Vector Regression for Tourist Arrivals Forecasting

Hsiou-Hsiang Liu,<sup>1</sup> Lung-Cheng Chang,<sup>1</sup> Chien-Wei Li ,<sup>2</sup> and Cheng-Hong Yang <sup>2,3</sup>

<sup>1</sup>Department of Tourism Management, National Kaohsiung University of Science and Technology, No. 415, Jiangong Rd., Sanmin District, Kaohsiung 807, Taiwan

<sup>2</sup>Department of Electronics Engineering, National Kaohsiung University of Science and Technology, No. 415, Jiangong Rd., Sanmin District, Kaohsiung 807, Taiwan

<sup>3</sup>Graduate Institute of Clinical Medicine, Kaohsiung Medical University, Kaohsiung 80708, Taiwan

Correspondence should be addressed to Cheng-Hong Yang; [chyang@cc.kuas.edu.tw](mailto:chyang@cc.kuas.edu.tw)

Received 14 May 2018; Revised 13 July 2018; Accepted 2 August 2018; Published 19 September 2018

Academic Editor: Mario Versaci

Copyright © 2018 Hsiou-Hsiang Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The tourism industry has become one of the most important economic sectors for governments worldwide. Accurately forecasting tourism demand is crucial because it provides useful information to related industries and governments, enabling stakeholders to adjust plans and policies. To develop a forecasting tool for the tourism industry, this study proposes a method that combines feature selection (FS) and support vector regression (SVR) with particle swarm optimization (PSO), named FS-PSOSVR. To ensure high forecast accuracy, FS and a PSO algorithm are employed to, respectively, select reliable input variables and to identify the optimal initial parameters of SVR. The proposed method was tested using a data set of monthly tourist arrivals to Taiwan from January 2006 to December 2016. The results reveal that the errors obtained using FS-PSOSVR are comparatively smaller than those obtained using other methods, indicating that FS-PSOSVR is an effective method for forecasting tourism demand.

## 1. Introduction

The tourism industry is one of the fastest promising economic sectors worldwide. Statistics from the World Tourism Organization [1] indicate that the number of inbound tourists increased by approximately 250 million between 2000 and 2010; moreover, this number is predicted to increase to 1.8 billion by 2030. Statistics from the World Travel and Tourism Council (WTTC) [2] show that the tourism industry accounts for 3.4% of world gross domestic product (GDP), and the tourism industry accounts for 2.8% of total employment globally. These statistics indicate how influential the tourism industry is in the global economic environment. The WTTC results related to Taiwan reveal that tourism generated 4.3% of its total GDP in 2017, and, furthermore, workers in this industry accounted for 5.2% of total employment [3]. These statistics indicate that the tourism industry has also become a critical component of the Taiwanese economy. Accurate forecasting of tourist volume

plays a major role in tourism planning because it enables destinations to predict requirements for infrastructural development in order to meet demand. In addition, accurately forecasting tourist arrivals and studying tourist arrival patterns are essential measures for tourism-related industries seeking to formulate efficient and effective strategies for maintaining and boosting the tourism sector.

Brida, Cortes-Jimenez, and Pulina [4] mentioned that, in several cases, a long-run bidirectional Granger causality exists between tourism and GDP; therefore, when testing tourism-led growth hypotheses, researchers should aggregate origin countries that exhibit similar features to avoid biased results. Misleading results may emerge in both the short and long term, because various source market segments may possess diverse characteristics. Thus, it may not be valid to state that expansion of the tourism sector contributes to long-term growth in a country that has negligible tourism sector in comparison with other economic sectors. The effect of uncertainty on growth is evident [5], and the

temporal relationship between GDP and international tourism is relatively isolated [6]. Therefore, countries (or groups of countries) where tourism is a bigger proportion of GDP were not considered as a threshold variable in the present study, which examines a hybrid artificial intelligence (AI) model to forecast tourist arrivals to Taiwan from its top four markets.

Many researchers have proposed approaches to forecasting demand in the tourism industry. The most common time series methods include the autoregressive integrated moving average (ARIMA) model [7] and exponential smoothing (ETS) [8]. These methods usually employ historical datasets to forecast future tourist flow through a univariate or multivariate mathematical function that is highly dependent on linear assumptions. Although widely recognized, such methods are limited by their poor nonlinear fitting capabilities [9]. This indicates the usefulness of nonlinear methods of demand forecasting, namely, the use of artificial neural networks (ANNs) [10] and support vector regression (SVR) [11]. Nevertheless, the benefits of AI methods depend on using appropriate parameter settings. Various techniques have been proposed to determine an adequate set of parameter values; however, a lack of thorough guidelines remains a problem [12]. Moreover, researchers have widely applied several consecutive lagged variables as input features in forecasting problems. Nevertheless, some input features may be redundant or even irrelevant to a specific demand forecasting problem, which reduces the accuracy of forecasting models [13]. Therefore, in this study, the feature selection (FS) method is applied to identify essential data and improve the forecasting effectiveness of the input variables.

To develop a reliable forecasting tool for the tourism industry, this study proposes a hybrid algorithm called FS-PSOSVR, which is a combination of FS techniques, the SVR model, and the particle swarm optimization (PSO) algorithm. FS is used to determine the most relevant input variables in the time series data. The PSO algorithm is employed to determine a set of optimal parameters for SVR, which are then used to construct the SVR model. To account for regional differences, tourism demand forecasting models are constructed for different regions. Monthly data of tourist arrivals in Taiwan from January 2006 to December 2016 are used as an example. The experimental results demonstrated that the proposed algorithm outperformed other forecasting methods such as ETS, ARIMA, seasonal ARIMA (SARIMA), grid search SVR (GRIDSVR), and PSOSVR.

## 2. Literature Review

Coshall and Charlesworth suggested that the methods used by studies on forecasting tourism demand can be categorized into causal econometric models and time series models [14]. Causal econometric models attempt to establish relationships between variables such as tourism demand (as measured by the number of tourist arrivals at a destination) or tourist expenditure and a set of hypothesized explanatory

factors. The most commonly used causal econometric models in the literature are cointegration and error correction models [15], vector autoregressive models [16], and linear almost-ideal system models [17]. These methods have also been combined [18].

In past decades, researchers have widely applied time series models, including ARIMA and ETS, to analyze issues concerning tourism demand forecasting. Lim and McAleer used ARIMA and SARIMA models to forecast the demand for tourism in Australia by analyzing data on tourists from Hong Kong, Malaysia, and Singapore [19]. Their results demonstrated that ARIMA is the most suitable model for predicting tourism demand for visiting Australia by visitors from Hong Kong and Malaysia. However, for Singaporean visitors, the SARIMA model generated superior results.

Chu used nine time series models—including two naive models [20], ARIMA-type models (ARIMA, SARIMA, and ARFIMA) and regression-based models—to forecast the volume of monthly tourist arrivals in Singapore. Chu reported that the ARFIMA model exhibited the highest forecasting accuracy both in the short and long term; however, for the medium term, the SARIMA model had superior performance [20].

More recently, Wan et al. used the SARIMA model and compared it with a seasonal moving average model and the Holt-Winter model [21]. Their findings indicated that the SARIMA model performed most favorably under all three h-step-ahead forecasting horizons. In addition, Baldigara and Mamula employed the SARIMA model to predict the number of German visitors to Croatia [22]. In their analysis, the predictive power of the SARIMA model was superior to that of the other methodologies mentioned in their study.

Assaf et al. used a comprehensive and accurate systematic approach to analyze tourism demand that was based on a Bayesian global vector autoregressive model [23]. Huang et al. used the PSO algorithm combined with a back-propagation neural network to establish a demand estimation model [24]. Akin et al. considered the SARIMA,  $\nu$ -SVR, and multilayer perceptron-type neural networks and optimized network parameters using numerous approaches for evaluating performances on monthly tourist arrival data to Turkey from various countries [25]. Cang proposed a nonlinear combination method using multilayer perceptron neural networks to map the nonlinear relationship between inputs and outputs [26]. Huarng et al. proposed an innovative forecasting model to detect regime switching properly and used a fuzzy time series model for forecasting [27].

SVR overcomes classification problems, nonlinear function problems, and forecasting problems by using loss functions. Because tourism data usually display nonlinear characteristics, SVR is also widely applied in research on tourism demand forecasting. Chen and Wang combined SVR with a genetic algorithm (GA) to forecast the number of inbound visitors to China [28]. Their results illustrated that the forecast generated through GA-SVR was more persuasive than the forecasts of back-propagation neural network and ARIMA models. Hong et al. combined a chaotic

GA with an SVR model to obtain forecasting information of visitor arrivals to Barbados [29]. This combined method achieved a more accurate estimation than the other models.

Most recently, Tsaur and Chan suggested that gray SVR can be employed to forecast the number of visitors from China to Taiwan [30]. The results obtained using this proposed method were superior to those obtained using other methods such as regression analysis, simple ETS, and a gray model.

### 3. Methodology

**3.1. SVR.** Support vector machines (SVMs) were initially introduced to address classification problems [31]. SVR is a version of an SVM and was proposed by Drucker et al. [32]. The basic functions of the SVR model are designed to provide a nonlinear mapping function that maps the training data to a high-dimensional feature space. The training dataset is denoted by  $\{(x_i, y_i); i = 1, 2, \dots, N; x_i \in R^n; y_i \in R\}$ , where  $x_i$  is the  $i$ th input in the  $n$ th dimension,  $y_i$  is the actual output, and  $N$  is the data set size. The SVR function is then

$$y = f(x_i) = w^T \varphi(x_i) + b, \quad (1)$$

where  $f(x)$  denotes the forecast values,  $\varphi(x)$  is the feature function of the inputs, and  $w$  and  $b$  are adjustable coefficients. Through employing a penalty function to estimate the values of coefficients  $w$  and  $b$ , the penalty function  $R(C)$  becomes

$$R(C) = \frac{1}{2} \|w\|^2 + C \cdot \frac{1}{n} \sum_{i=1}^n |y_i - f(x)|_\varepsilon, \quad (2)$$

$$|y - f(x)|_\varepsilon = \begin{cases} 0, & |y - f(x)| \leq \varepsilon, \\ |y - f(x)| - \varepsilon, & \text{otherwise,} \end{cases} \quad (3)$$

where  $C$  is the penalty coefficient and  $\varepsilon$  is the maximum value of tolerable error [33]. Two slack variables  $\xi_i$  and  $\xi_i^*$  are introduced to cope with the infeasible constraints of the optimization problem, which becomes

$$\begin{aligned} & \min_{\omega, b, \xi^{(*)}} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*), \\ & \text{subject to} \begin{cases} -y_i + w^T \varphi(x_i) + b \leq \varepsilon + \xi_i, & (i = 1, \dots, m), \\ y_i - w^T \varphi(x_i) - b \leq \varepsilon + \xi_i^*, & (i = 1, \dots, m), \\ \xi_i^{(*)} \geq 0, & (i = 1, \dots, m), \end{cases} \end{aligned} \quad (4)$$

where  $\xi^{(*)}$  ensures that the constraint is satisfied,  $C$  controls the balance between model complexity and training error rate, and  $\varepsilon$  is a constant for controlling the tube size; if  $\varepsilon$  is too small, overfitting may occur, and the opposite situation may cause underfitting. Using the Lagrange equation, the dual optimization problem is obtained:

$$\begin{aligned} & \min_{\alpha_i, \alpha_i^*} \frac{1}{2} \sum_{i,j=1}^n y_i (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) k(x_i, x_j) \\ & + \sum_{i=1}^n ((\varepsilon - y_i) \alpha_i + (\varepsilon + y_i) \alpha_i^*), \\ & \text{subject to} \begin{cases} \sum_{i=1}^N (\alpha_i - \alpha_i^*) = 0, \\ 0 \leq \alpha_i^{(*)} \leq C, & (i = 1, \dots, m). \end{cases} \end{aligned} \quad (5)$$

To solve (5), the SVR function can be obtained as follows:

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) k(x_i, x) + b, \quad (6)$$

where  $\alpha_i$  and  $\alpha_i^*$  are the Lagrange multipliers and  $k(x_i, x)$  is a kernel function. The kernel function constructs a nonlinear decision hypersurface in the SVR input space. The most widely used kernel, the Gaussian radial basis function (RBF) kernel, not only performs nonlinear mapping between the input space and a high-dimensional space but is also easy to implement, and thus, it is suitable for solving nonlinear problems. Therefore, the Gaussian RBF kernel was employed in this study:

$$k(x_i, x) = \exp\left(-\sigma \|x - x_i\|^2\right), \quad (7)$$

where  $\sigma$  represents the scaling factor of the Gaussian RBF kernel.

**3.2. Particle Swarm Optimization.** PSO is a population-based iterative optimization algorithm inspired by the social behavior of bird flocking that was developed by Eberhart and Kennedy [34]. PSO has successfully been applied in numerous researches [35–37]. The optimization process starts with a randomly initialized population of solutions, which are called particles. The swarm consists of  $n$  particles, and each particle has a position vector  $x_i = (x_{i,1}, \dots, x_{i,d})$  and velocity vector  $v_i = (v_{i,1}, \dots, v_{i,d})$ , where  $i = 1, 2, \dots, n$  and  $d$  are the number of dimensions in the vector. Each particle is a potential solution to the problem in the  $D$ -dimensional search space. The particles share information with each other; thus, each particle can be influenced to adjust their search direction toward a promising search region. Each particle has its own optimal experience, represented as the best known position of particle  $i$  ( $pbest_i$ ) in the feature space, and the optimal experience derived from the population is represented as the best known position within the population ( $gbest$ ). During each generation, each particle is accelerated toward  $pbest_i$  and  $gbest$ . The value of experience is evaluated using the fitness function  $f(x)$  according to the problem definition. Both position and velocity must be limited to between the rational lower boundary  $b_{low}$  and upper boundary  $b_{up}$  in the feature space. The updated velocity and position can be obtained using the following equations:

$$v_{i,d}^{\text{new}} = wv_{i,d}^{\text{old}} + c_1r_1(pbest_{i,d} - x_{i,d}) + c_2r_2(gbest_d - x_{i,d}), \quad (8)$$

$$x_i^{\text{new}} = x_i^{\text{old}} + v_i^{\text{new}}, \quad (9)$$

where  $w$  is the inertia weight,  $c_1$  and  $c_2$  are acceleration constants, and  $r_1$  and  $r_2$  are uniformly random values between 0 and 1.

The inertia weight  $w$  controls the current velocity. A larger inertia weight facilitates global exploration, whereas a small value promotes local exploration. To balance global and local exploration capabilities; this study employs the commonly used linearly decreasing inertia weight (LDW), because an inertia weight that decreases with time from 0.9 to 0.4 is more favorable than a fixed inertia weight. The LDW formulation is as follows:

$$w = w_{\min} - (w_{\max} - w_{\min}) \times \frac{\text{iter}}{\text{max\_iter}}, \quad (10)$$

where  $w_{\max}$  is set at 0.9,  $w_{\min}$  is set at 0.4, and  $\text{max\_iter}$  and  $\text{iter}$  are the maximum iteration and current iteration, respectively.

All operations in PSO are repeated until the termination condition is reached. The termination condition corresponds to the maximum number of iterations. The PSO algorithm is described in Algorithm 1.

**3.3. Selecting the SVR Parameters Using PSO.** In SVR modeling, the parameter settings affect the performance of a forecast time series, as mentioned in the preceding discussion of PSO. The crucial parameters are the regularization parameter ( $C$ ), bandwidth of the kernel function ( $\sigma$ ), and tube size of the  $\epsilon$ -insensitive loss function ( $\epsilon$ ). Improper choice of parameter values can result in either overfitting or underfitting [38]. Consequently, selecting the optimal parameters is crucial when employing SVR to forecast a time series. In this study, the PSO algorithm is utilized to select approximations of the three parameters of the SVR model. Figure 1 presents a flowchart of PSOSVR.

The procedures of the PSOSVR model are as follows:

*Step 1: initialization.* First, the initialization values of the parameters are set; then, the particles are generated in the feature space. Each particle  $i$  is represented by  $x_i = \{C, \sigma, \epsilon\}$ .

*Step 2: fitness evaluation.* After the encoding procedure is completed, the three values of parameters  $C$ ,  $\sigma$ , and  $\epsilon$  are inserted into the SVR model to forecast the problem; a  $k$ -fold cross-validation (CV) is employed in the training phase to avoid overfitting, and the validation error is calculated. Figure 2 illustrates the concept of  $k$ -fold CV. The PSOSVR model uses a rolling-based procedure to forecast data. Figure 3 illustrates the rolling-based mechanism. The previous 12 lagged observation data points are selected as input variables and the current data as output variables. First, the top 12 tourist datasets are fed into the proposed model. Following this, a one-step-ahead forecasting value is

obtained. The next rolling 12 data points are fed into the proposed model again, and the second forecasting value is obtained. The process is repeated until all the forecasts in the training set are obtained, following which the validation error is calculated. In this study, the mean absolute percentage error (MAPE) is adopted as the fitness function. The MAPE is calculated using the following:

$$\text{MAPE}_{\text{CV}} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - f_i}{y_i} \right| \times 100\%, \quad (11)$$

where  $y_i$  is the actual value,  $f_i$  is the forecast value, and  $N$  is the sample size.

*Step 3: update of pbest.* If the fitness value of particle  $i$  in the current iteration exceeds that of  $pbest_i$ , then  $pbest_i$  is replaced by  $x_i$ .

*Step 4: update of gbest.* If the fitness value of  $pbest_i$  in the current iteration exceeds that of  $gbest$ , then  $gbest$  is replaced by  $pbest_i$ .

*Step 5: update of velocity.* The velocity of each particle is calculated according to (8).

*Step 6: update of position.* The position of each particle is calculated according to (9).

*Step 7: stop criteria.* The processes are repeated in the aforementioned order until the maximum iteration is reached.

**3.4. Random Forest.** Random forest (RF) is an ensemble learning method for both classification and regression problems [39]. The principle of RF is to combine a set of binary decision trees, each of which is constructed using a bootstrap sample obtained from the learning sample and a subset of features (input variables or predictors) randomly chosen at each node. The prediction is made using a majority vote of the trees (in classification) or by averaging their outputs (in regression). In addition to classification and regression, RF provides an internal measure of variable importance through computing importance scores. Similarly, it can be used to select crucial features. During the construction of an RF, each node of a decision tree is split into two children, whereas a splitting criterion is used to reduce the impurity of a node, which is measured through Gini importance [39]. In the process of node splitting,  $i$  is the impurity of the node, and the node's Gini importance is defined as follows:

$$i = 1 - \sum_j p^2(j), \quad (12)$$

where  $p(j)$  is the proportion of samples that are labeled  $j$  in this node. After splitting, the impurity of the node is described as follows:

$$\Delta i = i_{\text{parent}} - (p_{\text{left}} \times i_{\text{left}} + p_{\text{right}} \times i_{\text{right}}), \quad (13)$$

where  $p_{\text{left}}$  and  $p_{\text{right}}$  are the sample proportions of the left and right child nodes, respectively, and  $i_{\text{parent}}$ ,  $i_{\text{left}}$ , and  $i_{\text{right}}$

```

(1) Define: let  $f()$  be the fitness function,  $N$  is the number of particles,  $D$  is the number of dimensions,  $x_i$  and  $v_i$  are the position and velocity of each particle, respectively,  $pbest_i$  is the best known position of particle  $i$ , and  $gbest$  is the best known position of the entire swarm.
(2) Output:  $gbest$  // the optimal solution
(3) // initialize the swarm
(4) for  $i \leftarrow 1$  to  $N$  do // each particle
(5)   for  $d \leftarrow 1$  to  $D$  do // each dimension
(6)      $x_{i,d} \leftarrow \text{rand}(lb, ub)$  //  $lb$  and  $ub$  are the lower and upper boundaries of the search space
(7)      $v_{i,d} \leftarrow \text{rand}(-|ub - lb|, |ub - lb|)$ 
(8)   end
(9)    $pbest_i \leftarrow x_i$ 
(10)  if  $f(pbest_i) < f(gbest)$  then
(11)     $gbest \leftarrow pbest_i$ 
(12)  end
(13) end
(14) while  $iter < \text{max\_iter}$  do // iterate until termination criterion met
(15)  for  $i \leftarrow 1$  to  $N$  do
(16)    for  $d \leftarrow 1$  to  $D$  do
(17)       $r_1 \leftarrow \text{rand}(0, 1)$ ;  $r_2 \leftarrow \text{rand}(0, 1)$ 
(18)       $v_{i,d} \leftarrow wv_{i,d} + c_1r_1(pbest_{i,d} - x_{i,d}) + c_2r_2(gbest_d - x_{i,d})$  // update the particle's velocity
(19)    End
(20)     $x_i \leftarrow x_i + v_i$  // update the particle's position
(21)    if  $f(x_i) < f(pbest_i)$  then // compare particle's best position
(22)       $pbest_i \leftarrow x_i$ 
(23)    if  $f(pbest_i) < f(gbest)$  then // compare swarm's best position to current particle's position
(24)       $gbest \leftarrow pbest_i$ 
(25)    End
(26)  End
(27) End
(28)  $w \leftarrow w_{\text{max}} - (w_{\text{max}} - w_{\text{min}}) \times \text{iter}/\text{max\_iter}$  // decrease inertia weight
(29)  $iter \leftarrow iter + 1$ 
(30) end

```

ALGORITHM 1: Particle swarm optimization algorithm.

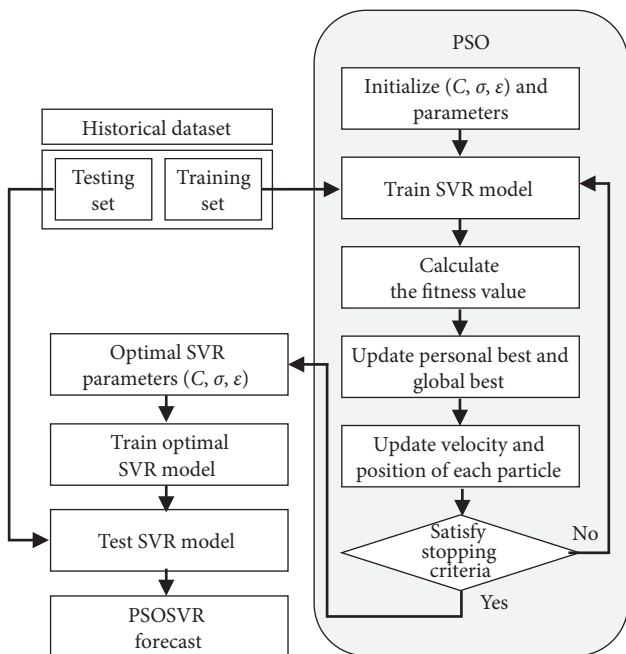


FIGURE 1: Flowchart of PSOSVR.

are the Gini importances of the parent, left child node, and right child node, respectively. For any one feature  $X_i$ , the sum of its impurity decrement in all decision trees is the Gini importance of  $X_i$ :

$$\alpha \Delta I = \sum_k \Delta i_k. \quad (14)$$

This equation indicates the importance of each feature, and a greater value indicates that the feature is more important.

Recursive feature elimination (RFE) is a recursive process based on feature ranking techniques [40]. According to a certain feature ranking standard, RFE starts from a complete set and then eliminates the least relevant feature one at a time to select the most important features. This study uses an FS method that combines RFE and RF, named, RF-RFE. The process is presented in Algorithm 2.

**3.5. FS-PSOSVR.** The FS-PSOSVR model is proposed to determine the most effective feature subset and improve the forecasting performance of PSOSVR. Figure 4 illustrates the algorithm, and the detailed steps are as follows:

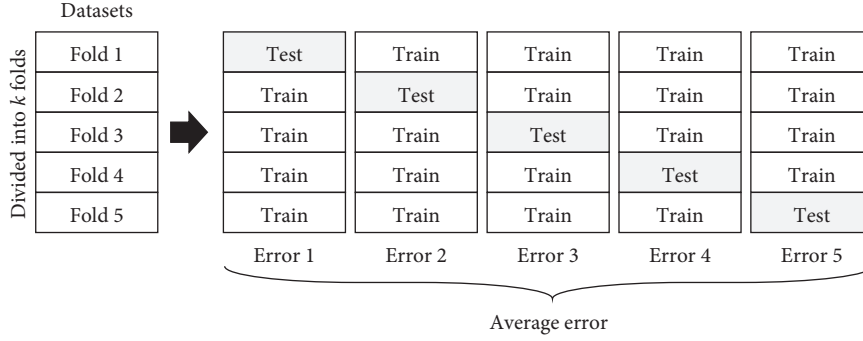
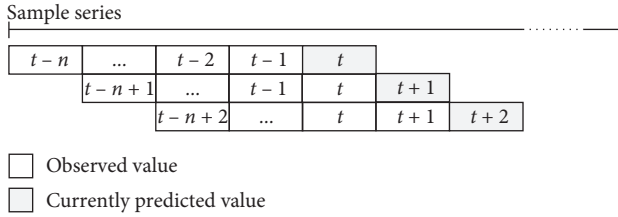
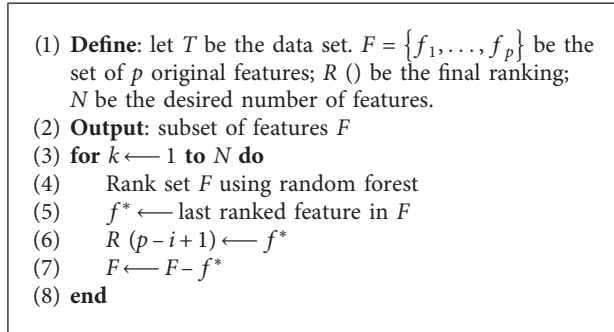
FIGURE 2: Concept of  $k$ -fold CV ( $k=5$ ).

FIGURE 3: Rolling-based forecasting mechanism.



ALGORITHM 2: Random forest—Recursive feature elimination (RF-RFE).

*Step 1.* the data set is divided into a training set and a test set. The training set is used as the original subset  $F$ .

*Step 2.* the RF model is trained using subset  $F$ , and the variable importance scores of each feature in the subset are calculated.

*Step 3.* the least important feature is eliminated from  $F$ , and Step 2 is repeated until the desired number of features is obtained.

*Step 4.* after allowing the new training set to be the feature subset  $F$  obtained by RF-RFE, the PSOSVR process is initiated.

**3.6. Performance Criteria.** Two common statistical metrics, root-mean-square error (RMSE) and MAPE, are used to evaluate the performance of the forecasting models (Table 1) by comparing the deviation between the real and forecast values. Lower RMSE and MAPE values represent higher

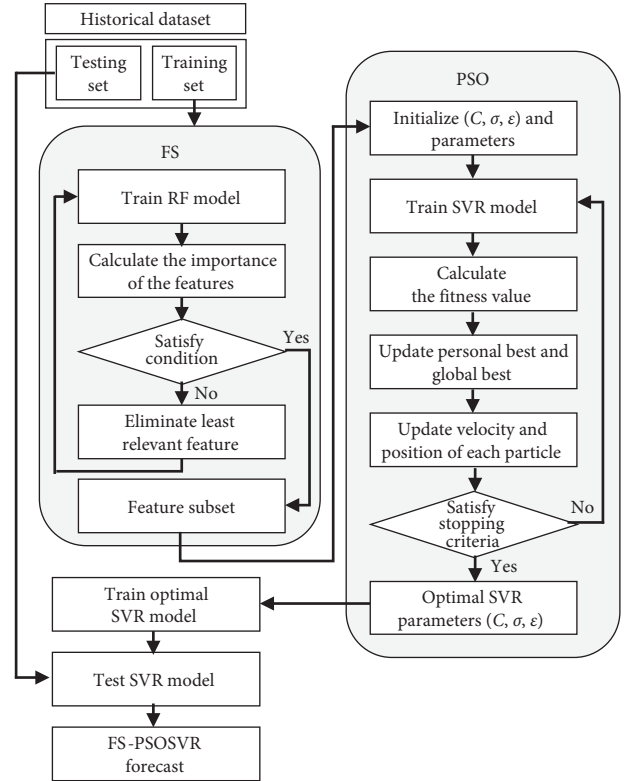


FIGURE 4: Flowchart of FS-PSOSVR.

TABLE 1: Performance metrics.

Metrics	Calculation
RMSE	$\sqrt{(1/N) \sum_{i=1}^N (y_i - f_i)^2}$
MAPE	$(1/N) \sum_{i=1}^N  (y_i - f_i)/y_i  \times 100\%$

$y_i$  is the actual value,  $f_i$  is the forecast values, and  $N$  is the sample size.

accuracy, thus indicating that the forecast values are reliable. Lewis [41] developed a table (Table 2) containing typical MAPE values for analyzing and interpreting industrial and business data.

**3.7. Parameter Settings.** The size of the population is set at 50; acceleration factors  $c_1$  and  $c_2$  are both set at 2.0; and the

TABLE 2: Interpretation of MAPE values.

MAPE	Interpretation
<10%	Highly accurate forecasting
10–20%	Good forecasting
20–50%	Reasonable forecasting
>50%	Inaccurate forecasting

maximum number of iterations (`max_iter`) is set at 100. These parameters are selected according to Bratton and Kennedy [42]. In this study, the search scopes of the SVR parameters are set at  $C = [2^0, 2^{10}]$ ,  $\sigma = [2^{-8}, 2^0]$ , and  $\varepsilon = [2^{-8}, 2^0]$ . The traditional SVR model uses the grid search method (GRIDSVR) to determine the optimal parameters. GRIDSVR increments the parameters exponentially; thus, the search spaces of GRIDSVR are set as  $C = [2^0, 2^1, 2^2, \dots, 2^{10}]$ ,  $\sigma = [2^{-8}, 2^{-7}, 2^{-6}, \dots, 2^0]$ , and  $\varepsilon = [2^{-8}, 2^{-7}, 2^{-6}, \dots, 2^0]$ .

## 4. Results and Discussion

The `auto.arima` and `ets` functions of the *R* forecast package [43] were used to identify (S)ARIMA and ETS models. The *Python* module `sklearn.svm`, which is an interface of the LIBSVM library [44], was used to train the SVR-based models.

**4.1. Data Sets and Preprocessing.** This paper presents a hybrid AI model to forecast tourist arrivals to Taiwan from the top four markets. To evaluate the proposed approach, we applied it to data on tourist arrivals to Taiwan that have been used in several papers [6, 45–47] as a case study. Specifically, data for monthly tourist arrivals to Taiwan from 2006 to 2016, collected from the Tourism Statistics Database [48], were used. Japan, Hong Kong and Macao, South Korea, and the United States (Figure 5) were selected as the four groups that contributed the most visitors to Taiwan. Visitors from China were excluded because drastic fluctuations often occur as a result of the cross-strait relationship and political uncertainty. Each data set was divided into two subsets:

*Training set:* used for training the model; it consisted of the monthly data for 2006–2015.

*Test set:* used for testing the forecast accuracy; it consisted of the monthly data for 2016.

Hsu et al. [49] noted that scaling original data helps improve forecast performance and predictive accuracy by not allowing attributes with greater numerical ranges to dominate those with smaller numerical ranges and avoiding potential numerical problems. In this study, each data set was normalized within the range [0.15, 0.85] using the following:

$$\widetilde{X}_t = \frac{X_t - X_{\min}}{X_{\max} - X_{\min}} \times 0.7 + 0.15, \quad (15)$$

where  $X_t$  is the number of tourist arrivals at time  $t$  and  $X_{\max}$  and  $X_{\min}$  are the maximum and minimum numbers of tourist arrivals in the period of the data set, respectively.

**4.2. Comparison of Time Series Models and SVR-Based Models.** Figure 6 depicts the differences between the actual data and forecast results. These figures reveal that the results

obtained using the proposed FS–PSOSVR method more accurately reflect the actual data compared with the results obtained using the other methods. The MAPE and RMSE for each method in Table 3 were also employed to compare forecast performances. To test and verify the forecasting ability of the AI model for the time series model, the ETS, ARIMA, and SARIMA models were set as the objects of comparison. Compared with the time series model, instead of confirming whether the data belong to a stationary state and considering whether other statistical tests should be applied, the AI model learns from characteristics of the training data. Table 4 lists the average MAPE values of the forecasts obtained using ETS, ARIMA, SARIMA, GRIDSVR, PSOSVR, and FS–PSOSVR. For the time series forecasting problem, the AI model demonstrated a similar ability to the time series model. Moreover, the FS–PSOSVR method was superior in solving the forecasting problems.

### 4.3. Comparison of GRIDSVR, PSOSVR, and FS–PSOSVR.

This study measured the ability to obtain the optimal parameters of GRIDSVR and PSOSVR to prove that the effectiveness of the SVR method depends on the parameters selected. The method that GRIDSVR employs to reach the most suitable parametric combination is a grid search, which involves calculating the fitness value of each grid. This may indicate that the most suitable parameter combinations do not exist in the grid; therefore, PSO was used to improve analysis of the problem. Table 5 presents the optimal values of the three SVR parameters for each SVR-based model. The results obtained using PSOSVR were more accurate than those obtained using GRIDSVR (Figure 6). This observation indicates that PSO achieved more favorable forecasting results.

To increase forecasting accuracy, the FS method was employed. RF–RFE was used to identify the reliable lagged variables. To determine the appropriate number of features, this study tested four to eight features to determine which number was optimal. The lagged variables are presented in Table 6, and  $y_{t-i}$  indicates the number of visitors  $i$  months ago. The results indicated that after applying FS, the prediction ability was superior to that of PSOSVR without FS. Additionally, by removing the input variable with less influencing power, a more suitable result was obtained.

### 4.4. Analysis for Individual Data

**4.4.1. Japan.** Increasing the accuracy of forecasting is especially helpful for Japanese tourists because the tourism market from Japan has remained steady. The MAPE value of each forecasting model was less than 10; according to the interpretation of MAPE values by Lewis [41], all MAPE results can be categorized as denoting high accuracy. For the FS–PSOSVR method, the greatest gap between the experimental results and real data occurred in the data from August 2016. Based on the data from Taiwan’s Tourism Bureau, compared with the growth rate in 2015—which remained constant—the growth rate for the entire year of 2016 increased significantly to 16.5%. Among the data for each month, the growth rate of

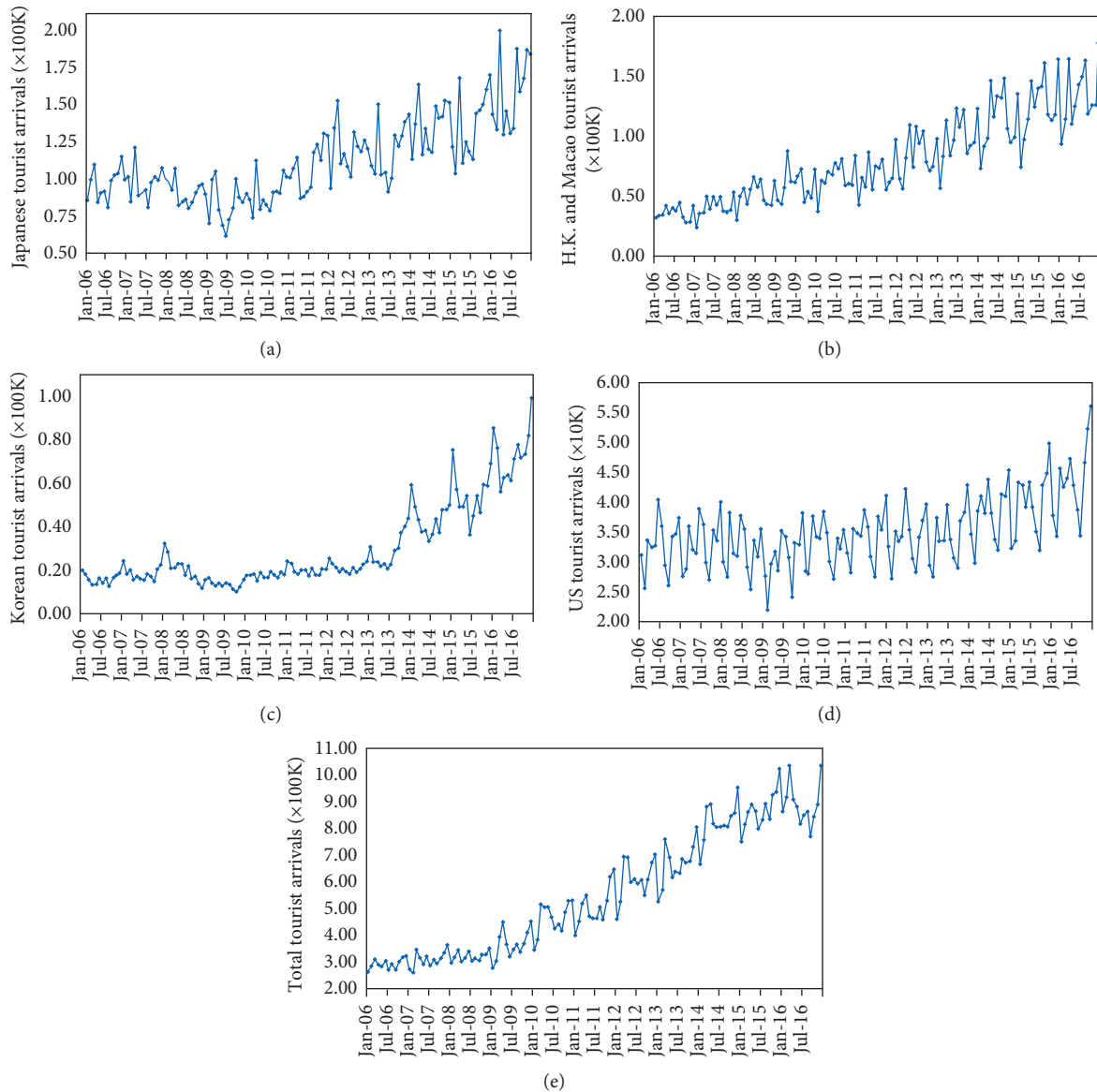


FIGURE 5: Monthly tourist arrivals from January 2006 to December 2016 from (a) Japan, (b) Hong Kong and Macao, (c) South Korea, (d) the United States, and (e) the total number.

Japanese visitors traveling to Taiwan increased to approximately 30%. This dramatic growth is the main reason that the experimental outcomes do not match the forecast [50]. Although the SARIMA model obtained results that were more similar to the real numbers in March and August 2016, causing their RMSE values to decrease, the holistic results were still slightly higher than the actual results. This suggests that the SARIMA model does not accurately reflect the actual tourist trend.

**4.4.2. Hong Kong and Macao.** As with the results for Japan, all MAPE results for Hong Kong and Macao indicated high accuracy. The number of visitors from Hong Kong and Macao in March and April is influenced by the Easter vacation. Every model used by this study suggested that the number of tourists increases in April but decreases in May. In 2016, the

Easter vacation, occurring in March, was the external factor. This external factor—the variable dates of the Easter vacation—is not accounted for in the univariate analysis. In contrast to August and September of the previous year (2015), the growth rate in 2016 slightly increased (1.75% for August and 0.35% for September). This reveals that most of the models used in this study overestimated visitors from Hong Kong and Macao, and they are therefore unsuitable.

**4.4.3. South Korea.** Only the SARIMA and FS-PSOSVR models reached the standard of high accuracy for South Korea. The other models achieved adequate results, but substantial differences between each model's results were observed. Differences in tourism levels can be inferred from changing government policies. In the past 3 years, the Taiwanese government has commissioned more advertisements to promote





FIGURE 6: Forecast results for different datasets: (a) Japan; (b) Hong Kong and Macao; (c) South Korea; (d) the United States; and (e) total tourist numbers.

TABLE 3: Forecast tourist arrivals obtained using ETS, ARIMA, SARIMA, GRIDSVR, PSOSVR, and FS-PSOSVR.

Case		ETS	ARIMA	SARIMA	GRIDSVR	PSOSVR	FS-PSOSVR
Japan	MAPE (%)	8.54	12.87	7.25	7.22	6.95	<b>5.23</b>
	RMSE	18091.23	24665.02	<b>11878.9</b>	14670.95	13616.38	13308.91
Hong Kong and Macao	MAPE (%)	11.07	16.28	12.10	12.39	12.24	<b>10.65</b>
	RMSE	21045.26	23224.72	20043.2	21168.17	21013.52	<b>19738.13</b>
South Korea	MAPE (%)	13.48	13.58	9.91	11.45	11.14	<b>7.66</b>
	RMSE	13482.26	13515.24	8681.24	11439.02	11137.59	<b>6807.39</b>
The United States	MAPE (%)	4.79	10.03	3.95	5.45	4.62	<b>3.84</b>
	RMSE	3626.19	6508.72	2786.49	3027.27	2883.04	<b>2218.12</b>
Total	MAPE (%)	10.64	12.08	11.24	11.21	11.14	<b>9.76</b>
	RMSE	100954.40	116687.03	111400.6	108457.15	107765.04	<b>95910.99</b>

Bold: the superior values.

TABLE 4: Average MAPE of ETS, SARIMA, GRIDSVR, PSOSVR, and FS-PSOSVR.

	ETS	ARIMA	SARIMA	GRIDSVR	PSOSVR	FS-PSOSVR
MAPE (%)	9.70	12.97	8.89	9.54	9.22	<b>7.43</b>

Bold: the lowest average MAPE.

TABLE 5: Training results of GRIDSVR, PSOSVR, and FS-PSOSVR.

Model	Data set	$C$	$\epsilon$	$\sigma$
GRIDSVR	Japan	1024.00	0.0313	0.0156
	Hong Kong and Macao	128.00	0.0039	0.0313
	South Korea	1024.00	0.0156	0.0039
	The United States	128.00	0.2500	0.0625
	Total	256.00	0.0039	0.0039
PSOSVR	Japan	805.03	0.0423	0.0130
	Hong Kong and Macao	209.96	0.0039	0.0367
	South Korea	776.87	0.0159	0.0039
	The United States	415.75	0.0467	0.0367
	Total	216.66	0.0039	0.0039
FS-PSOSVR	Japan	564.45	0.1371	0.0241
	Hong Kong and Macao	389.08	0.0039	0.0039
	South Korea	673.88	0.0222	0.0094
	The United States	531.61	0.0268	0.0405
	Total	1016.71	0.0087	0.0060

TABLE 6: Lagged variables of FS-PSOSVR.

Data set	Lagged variables
Japan	$y_{t-12}, y_{t-3}, y_{t-2}, y_{t-1}$
Hong Kong and Macao	$y_{t-12}, y_{t-4}, y_{t-3}, y_{t-1}$
South Korea	$y_{t-12}, y_{t-11}, y_{t-10}, y_{t-3}, y_{t-2}, y_{t-1}$
The United States	$y_{t-12}, y_{t-6}, y_{t-3}, y_{t-1}$
Total	$y_{t-12}, y_{t-8}, y_{t-4}, y_{t-3}, y_{t-2}, y_{t-1}$

tourism. Compared with the statistics from 3 years prior to that, the number of South Korean tourists rose considerably. This noticeable growth produces deviations that influenced the final forecasting results. Because of the SVR model's ability to learn from historical data, the differences that GRIDSVR and PSOSVR obtained were markedly larger than those obtained using other models from August, September, and December of 2016. Nonetheless, in contrast to the forecasting results obtained using the other two SVR models, FS-PSOSVR—which selects the features with higher influence authority in advance—obtained an accurate result.

*4.4.4. United States.* All the MAPE results indicated high accuracy; this is because the tourism market of the United States has remained steady over the last decade. Visible fluctuations occurred in April, November, and December of 2016. In the past two years, the number of American visitors increased dramatically in April—by 22.03% in 2015 and 4.95% in 2014—even though the total number of visitors in 2016 decreased by 0.84%. For November and December, the real data indicated rises of 16.71% and 12.66%, respectively, which were much higher than those observed in 2015 (8.99% and 9.87%) and 2014 (7.36% and 5.55%). Consequently, a larger discrepancy occurred between the real and predicted numbers.

*4.4.5. Total.* From July 2016, the total number of visitors increased slowly, mainly because the number of visitors from China decreased dramatically. Notably, the statistics for September suggested much fewer visits than the previous year. This suggests that policy changes and political events have a dramatic effect on the willingness and ability of tourists from China to travel to Taiwan. The influence of policies is an external factor that cannot be predicted in the univariate analysis. For all tourists, as for those from Hong Kong, the Easter vacation had a strong influence, with the most obvious discrepancy occurring in March and April. This demonstrates that the model cannot identify a steady mode during the training process, leading to apparent differences.

Because tourists from different countries may perceive a destination differently because of unique motivations and expectations [51], effective policy interventions can be implemented to solve this problem. Furthermore, policy interventions can guidance for tourism planning and development, thereby creating a favorable tourism environment [52]. Also crucial for effective tourism planning is maintaining the quality of tourists' experiences and shaping their perceptions of places and lifestyles [52, 53]. To increase the number of visitors to Taiwan, the government and major

tour operators must tailor their marketing efforts to each individual territory and strive to construct and maintain Taiwan's positive image using various methods.

Lin and Kuo [54] asserted that the concept of nationality implies notions of cultural values, social norms, and economic development. Given the dynamic effect of nationality, travel agencies could improve their inclusive group packages to generate diversified images. Moreover, a culture of co-operation between different and heterogeneous actors in the business value chain should be fostered [52]. Another necessary approach for Taiwan's destination managers to expand their market reach is to explore potential products that may lure newly emerging markets to Taiwan, such as the younger generation Y segment, while continuing to attract the older Baby Boomer generation [55].

Taiwan's policy makers and destination managers must also acknowledge market diversity to devise more specific strategies for different market segments and design desirable destination experiences to attract more tourists throughout the year [56]. Moreover, to provide a deeper level of engagement with destinations, service providers must offer a wide range of services and resources to tourists in order to enable them to fully enjoy their stay, thereby positively affecting their evaluation of Taiwan as a destination [53]. Destination managers in Taiwan must develop their capacity and capability to deliver new and innovative products and services to leverage the full potential of Taiwan as a tourist destination [57].

**4.5. Advantages and Limitations.** The results show that FS-PSOSVR performs superior forecasting of tourist arrivals. FS-PSOSVR retains the advantages of FS, PSO, and SVR. There are no additional parameters in the PSO and no limits to the number of constraints. During the numerous iterations, the most optimal particle transmits the information onto other particles; thus, FS-PSOSVR has fast re-searching speed [58]. Furthermore, PSO has no overlapping and mutation calculations; thus, the search can be conducted at the particle speed. FS-PSOSVR also has advantages in high dimensionality space because SVR does not depend on the dimensionality of the input space. The advantage of feature space representation in tourist arrival forecasting is the mean squared error as well as that the loss function is also the mean squared error [32]. The FS in FS-PSOSVR can reduce noise to improve forecasting accuracy; furthermore, more interpretable features can facilitate understanding of the importance of features. However, FS-PSOSVR has some limitations in terms of convergence speed because PSO easily suffers from partial optimism, which causes less accuracy in terms of regulation of its speed and direction. The SVR model is a well-known forecasting approach that has been applied to solve time series problems [59]. Our results also demonstrated that FS-PSOSVR was superior to SARIMA in terms of forecasting tourist arrivals. However, FS-PSOSVR does not address how to handle seasonal time series datasets. Thus, we suggest applying decomposition techniques to obtain decomposed seasonal time series data when a data set belongs to seasonal time series data [60].

## 5. Conclusion

Forecasting of tourist arrivals is critical to accurately predict requirements for infrastructure development. In this study, we proposed an FS-PSOSVR algorithm for the forecast of tourism demand. In FS-PSOSVR, FS is used to identify essential data and improve the SVR forecasting effectiveness of input variables. We applied PSO to tune the suitable parameters for SVR and more effectively forecast tourism demand. The predictive power of the method was compared with that of five forecasting models: ETS, ARIMA, SARIMA, GRIDSVR, and PSOSVR. The parameters acquired by FS-PSOSVR were more accurate than the parameters derived from GRIDSVR and PSOSVR, indicating that FS-PSOSVR is more effective at optimizing the parameters of SVR than GRIDSVR and PSOSVR are. Moreover, FS-PSOSVR achieved greater forecasting accuracy than other methods such as ARIMA, SARIMA, and ETS, indicating that FS-PSOSVR is a relatively more effective means of forecasting tourism demand.

## Data Availability

The data for monthly tourist arrivals to Taiwan from 2006 to 2016 used to support the findings of this study have been deposited in the Tourism Statistics Database.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This study was partly supported by the Ministry of Science and Technology of Taiwan for Grant nos. 103-2221-E-151-029-MY3 and 105-2221-E-151-053-MY2.

## References

- [1] UNWTO, *UNWTO Tourism Highlights*, UNWTO, Madrid, Spain, 2016.
- [2] WTTC, *Travel & Tourism Economic Impact 2016*, WTTC, London, UK, 2016.
- [3] WTTC, *Travel & Tourism Economic Impact 2018: Taiwan*, WTTC, London, UK, 2018.
- [4] J. G. Brida, I. Cortes-Jimenez, and M. Pulina, "Has the tourism-led growth hypothesis been validated? A literature review," *Current Issues in Tourism*, vol. 19, no. 5, pp. 394–430, 2016.
- [5] C.-F. Chen and S. Z. Chiou-Wei, "Tourism expansion, tourism uncertainty and economic growth: new evidence from Taiwan and Korea," *Tourism Management*, vol. 30, no. 6, pp. 812–818, 2009.
- [6] C.-H. Wang, "Predicting tourism demand using fuzzy time series and hybrid grey theory," *Tourism Management*, vol. 25, no. 3, pp. 367–374, 2004.
- [7] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003.
- [8] R. J. Hyndman, A. B. Koehler, R. D. Snyder, and S. Grose, "A state space framework for automatic forecasting using

- exponential smoothing methods,” *International Journal of Forecasting*, vol. 18, no. 3, pp. 439–454, 2002.
- [9] J. Wang, S. Zhu, W. Zhang, and H. Lu, “Combined modeling for electric load forecasting with adaptive particle swarm optimization,” *Energy*, vol. 35, no. 4, pp. 1671–1678, 2010.
- [10] G. Zhang, B. E. Patuwo, and M. Y. Hu, “Forecasting with artificial neural networks: the state of the art,” *International Journal of Forecasting*, vol. 14, no. 1, pp. 35–62, 1998.
- [11] F. E. Tay and L. Cao, “Application of support vector machines in financial time series forecasting,” *Omega*, vol. 29, no. 4, pp. 309–317, 2001.
- [12] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, Cambridge University Press, New York, NY, USA, 2000.
- [13] Z. Hu, Y. Bao, T. Xiong, and R. Chiong, “Hybrid filter-wrapper feature selection for short-term load forecasting,” *Engineering Applications of Artificial Intelligence*, vol. 40, pp. 17–27, 2015.
- [14] J. T. Coshall and R. Charlesworth, “A management orientated approach to combination forecasting of tourism demand,” *Tourism Management*, vol. 32, no. 4, pp. 759–769, 2011.
- [15] B. Algieri, “An econometric estimation of the demand for tourism: the case of Russia,” *Tourism Economics*, vol. 12, no. 1, pp. 5–20, 2006.
- [16] H. Song and S. F. Witt, “Forecasting international tourist flows to Macau,” *Tourism Management*, vol. 27, no. 2, pp. 214–224, 2006.
- [17] Z. Han, R. Durberry, and M. T. Sinclair, “Modelling US tourism demand for European destinations,” *Tourism Management*, vol. 27, no. 1, pp. 1–10, 2006.
- [18] G. Li, K. K. Wong, H. Song, and S. F. Witt, “Tourism demand forecasting: a time varying parameter error correction model,” *Journal of Travel Research*, vol. 45, no. 2, pp. 175–185, 2006.
- [19] C. Lim and M. McAleer, “Time series forecasts of international travel demand for Australia,” *Tourism Management*, vol. 23, no. 4, pp. 389–396, 2002.
- [20] F.-L. Chu, “A fractionally integrated autoregressive moving average approach to forecasting tourism demand,” *Tourism Management*, vol. 29, no. 1, pp. 79–88, 2008.
- [21] S. Wan, S. Wang, and C. Woo, “Aggregate vs. disaggregate forecast: case of Hong Kong,” *Annals of Tourism Research*, vol. 42, pp. 434–438, 2013.
- [22] T. Baldigara and M. Mamula, “Modelling international tourism demand using seasonal ARIMA models,” *Tourism and Hospitality Management*, vol. 21, pp. 19–31, 2015.
- [23] A. G. Assaf, G. Li, H. Song, and M. G. Tsionas, “Modeling and forecasting regional tourism demand using the Bayesian global vector autoregressive (BGVAR) model,” *Journal of Travel Research*, 2018.
- [24] H.-C. Huang and C.-C. Ho, “Back-propagation neural network combined with a particle swarm optimization algorithm for travel package demand forecasting,” *International Journal of Digital Content Technology and its Applications*, vol. 6, no. 17, pp. 194–203, 2012.
- [25] M. Akin, “A novel approach to model selection in tourism demand modeling,” *Tourism Management*, vol. 48, pp. 64–72, 2015.
- [26] S. Cang, “A non-linear tourism demand forecast combination model,” *Tourism Economics*, vol. 17, no. 1, pp. 5–20, 2011.
- [27] K.-H. Huarng, T. H.-K. Yu, and F. Sole Parellada, “An innovative regime switching model to forecast Taiwan tourism demand,” *The Service Industries Journal*, vol. 31, no. 10, pp. 1603–1612, 2011.
- [28] K.-Y. Chen and C.-H. Wang, “Support vector regression with genetic algorithms in forecasting tourism demand,” *Tourism Management*, vol. 28, no. 1, pp. 215–226, 2007.
- [29] W.-C. Hong, Y. Dong, L.-Y. Chen, and S.-Y. Wei, “SVR with hybrid chaotic genetic algorithms for tourism demand forecasting,” *Applied Soft Computing*, vol. 11, no. 2, pp. 1881–1890, 2011.
- [30] R.-C. Tsaur and S.-F. Chan, “Grey support vector regression model with applications to China tourists forecasting in Taiwan,” *International Journal of Information and Management Sciences*, vol. 25, pp. 121–138, 2014.
- [31] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer Science & Business Media, New York, NY, USA, 2013.
- [32] H. Drucker, C. J. Burges, L. Kaufman, A. J. Smola, and V. Vapnik, “Support vector regression machines,” in *Proceedings of Advances in Neural Information Processing Systems*, pp. 155–161, Denver, CO, USA, May 1997.
- [33] B. Schölkopf, A. J. Smola, R. C. Williamson, and P. L. Bartlett, “New support vector algorithms,” *Neural Computation*, vol. 12, no. 5, pp. 1207–1245, 2000.
- [34] R. Eberhart and J. Kennedy, “A new optimizer using particle swarm theory,” in *Proceedings of Sixth International Symposium on Micro Machine and Human Science*, pp. 39–43, Piscataway, NJ, USA, 1995.
- [35] C.-H. Yang, Y.-D. Lin, L.-Y. Chuang, and H.-W. Chang, “SNP barcodes generated using particle swarm optimization to detect susceptibility to breast cancer,” *Natural Science*, vol. 5, no. 3, pp. 359–367, 2013.
- [36] C.-H. Yang, Y.-D. Lin, Y.-C. Chiang, and L.-Y. Chuang, “A hybrid approach for CpG island detection in the human genome,” *PLoS One*, vol. 11, no. 1, Article ID e0144748, 2016.
- [37] L.-Y. Chuang, Y.-D. Lin, and C.-H. Yang, “Data clustering using chaotic particle swarm optimization,” *IAENG International Journal of Computer Science*, vol. 39, pp. 208–213, 2012.
- [38] S. S. Keerthi, “Efficient tuning of SVM hyperparameters using radius/margin bound and iterative algorithms,” *IEEE Transactions on Neural Networks*, vol. 13, no. 5, pp. 1225–1229, 2002.
- [39] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [40] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, “Gene selection for cancer classification using support vector machines,” *Machine Learning*, vol. 46, no. 1–3, pp. 389–422, 2002.
- [41] C. D. Lewis, *Industrial and Business Forecasting Methods: A Practical Guide to Exponential Smoothing and Curve Fitting*, Butterworth-Heinemann, Oxford, UK, 1982.
- [42] D. Bratton and J. Kennedy, “Defining a standard for particle swarm optimization,” in *Proceedings of IEEE Swarm Intelligence Symposium*, pp. 120–127, Indianapolis, IN, USA, April 2007.
- [43] R. J. Hyndman and Y. Khandakar, *Automatic Time Series for Forecasting: the Forecast Package for R*, Department of Econometrics and Business Statistics, Monash University, Melbourne, Australia, 2007.
- [44] C.-C. Chang and C.-J. Lin, “LIBSVM: a library for support vector machines,” *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 3, pp. 1–27, 2011.
- [45] M.-S. Chen, L.-C. Ying, and M.-C. Pan, “Forecasting tourist arrivals by using the adaptive network-based fuzzy inference system,” *Expert Systems with Applications*, vol. 37, no. 2, pp. 1185–1191, 2010.
- [46] R.-C. Tsaur and T.-C. Kuo, “The adaptive fuzzy time series model with an application to Taiwan’s tourism demand,”

- Expert Systems with Applications*, vol. 38, no. 8, pp. 9164–9171, 2011.
- [47] P.-F. Pai, K.-C. Hung, and K.-P. Lin, “Tourism demand forecasting using novel hybrid system,” *Expert Systems with Applications*, vol. 41, no. 8, pp. 3691–3702, 2014.
- [48] Trinidad & Tobago Bureau of Standards, *Monthly Report on Tourist Hotel Operations in Taiwan*, Trinidad & Tobago Bureau of Standards, Macoya, Trinidad and Tobago, 2017.
- [49] C.-W. Hsu, C.-C. Chang, and C.-J. Lin, “A practical guide to support vector classification,” 2016, <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>.
- [50] Trinidad & Tobago Bureau of Standards, *Visitor Statistics*, Trinidad & Tobago Bureau of Standards, Macoya, Trinidad and Tobago, 2017.
- [51] M. A. Bonn, S. M. Joseph, and M. Dai, “International versus domestic visitors: an examination of destination image perceptions,” *Journal of Travel Research*, vol. 43, no. 3, pp. 294–301, 2005.
- [52] B. Algieri, A. Aquino, and M. Succurro, “International competitive advantages in tourism: an eclectic view,” *Tourism Management Perspectives*, vol. 25, pp. 41–52, 2018.
- [53] D. M. Frías-Jamilena, C. M. Sabiote-Ortiz, J. D. Martín-Santana, and A. Beerli-Palacio, “The effect of cultural intelligence on consumer-based destination brand equity,” *Annals of Tourism Research*, vol. 72, pp. 22–36, 2018.
- [54] C.-H. Lin and B. Z.-L. Kuo, “The moderating effects of travel arrangement types on tourists’ formation of Taiwan’s unique image,” *Tourism Management*, vol. 66, pp. 233–243, 2018.
- [55] K. J. Semrad and M. Rivera, “Advancing the 5E’s in festival experience for the Gen Y framework in the context of eWOM,” *Journal of Destination Marketing & Management*, vol. 7, pp. 58–67, 2018.
- [56] C. Mejia, N. Hua, W. Wei, X. Fu, and Y. Wang, “The promise of the 2016 USA-China tourism year: strategies to boost tourism exchange,” *Journal of Destination Marketing & Management*, vol. 8, pp. 423–425, 2018.
- [57] R. Croes, J. Ridderstaat, and M. van Niekerk, “Connecting quality of life, tourism specialization, and economic growth in small island destinations: the case of Malta,” *Tourism Management*, vol. 65, pp. 212–223, 2018.
- [58] Y. del Valle, G. K. Venayagamoorthy, S. Mohagheghi, J. C. Hernandez, and R. G. Harley, “Particle swarm optimization: basic concepts, variants and applications in power systems,” *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 2, pp. 171–195, 2008.
- [59] H. Song and G. Li, “Tourism demand modelling and forecasting—a review of recent research,” *Tourism Management*, vol. 29, no. 2, pp. 203–220, 2008.
- [60] P. F. Pai, K. P. Lin, C. S. Lin, and P. T. Chang, “Time series forecasting by a seasonal support vector regression model,” *Expert Systems with Applications*, vol. 37, no. 6, pp. 4261–4265, 2010.