A Wavelet Support Vector Machine Combination Model for Singapore Tourist Arrival to Malaysia

A. Rafidah¹, Ani Shabri², A. Nurulhuda³, Y. Suhaila³

¹Technical Foundation Department, Universiti Kuala Lumpur (UniKL), Malaysian Institute of Industrial Technology, Persiaran Sinaran Ilmu, Bandar Seri Alam, 81750, Johor, Malaysia
²Departments of Mathematics, Science Faculty, University of Technology Malaysia, Skudai, Johor Malaysia
³Technical Foundation Department, Universiti Kuala Lumpur (UniKL), Malaysian Institute of Industrial Technology, Persiaran Sinaran Ilmu, Bandar Seri Alam, 81750, Johor, Malaysia

Corresponding author: rafidahali@unikl.edu.my

Abstract. In this study, wavelet support vector machine model (WSVM) is proposed and applied for monthly data Singapore tourist time series prediction. The WSVM model is combination between wavelet analysis and support vector machine (SVM). In this study, we have two parts, first part we compare between the kernel function and second part we compare between the developed models with single model, SVM. The result showed that kernel function linear better than RBF while WSVM outperform with single model SVM to forecast monthly Singapore tourist arrival to Malaysia.

1. Introduction

The time series analysis is importantly suggested on the many application including the control of physical systems, process of engineering, biochemistry, environmental economic system, company management and economy country.[1] The forecasting of Singapore tourist arrivals has become a more pressing issue in recent years for governments at all levels, in order to reliably estimate tourism growth and the economic benefits generated by expanding tourism activities. Therefore, increasing the accuracy of forecasts is an essential requirement to improve the managerial, operational, and tactical decision-making process especially in the private sector.

This paper focuses on the application of new approach to solve problem in tourism demand. The wavelet and support vector machine are proposed to forecast Singapore tourist arrival to Malaysia. The objectives of this paper are to compare of different kernel functions for SVM and WSVM, to determine applicability of WSVM in modelling the prediction of Singapore tourist time series and to compare the prediction results by WSVM model with single SVM model. The research will be focused on the forecasting Singapore tourist by using WSVM and SVM model. In this paper, the data gather from monthly Singapore tourist arrival to Malaysia. The data was collected from January 1999 to December 2015. The monthly Singapore tourists were selected for this study. This paper is organized as follows. Section 2 presents model input and describes various kinds of forecasting
models in detail. In Section 3, these forecasting models are compared and evaluated. Finally, Section 4 presents the main conclusion of the paper.

2. Forecasting models

2.1. SVM Model

The SVM is a new technique for regression. The basic concept of the SVM is to map nonlinearly the original data $x$ into higher dimensional feature space. The SVM predictor is trained using a set of time series history values as inputs and a single output as the target value.\[7\]

Consider a given training set of \( n \) data points \( \{x_i, y_i\}_{i=1}^{n} \) with input data \( x_i \in \mathbb{R}^n \), \( p \) is the total number of data patterns and output \( y_i \in \mathbb{R} \). SVM approximate the function in the following form:

\[
y(x) = w^T \phi(x) + b
\]

(2.1)

where \( \phi(x) \) represent the higher dimensional feature space, which is nonlinearly mapped the input space \( x \).

In SVM for function estimation, the estimation by minimizing regularized risk function:

\[
\frac{1}{2} \|w\|^2 + C \sum_{i=1}^{m} L(y_i)
\]

(2.2)

is an arbitrary penalty parameter called the regularization constant.

Basically, SVM penalize \( f(x_i) \) when it departures from \( y_i \) by means of an \( \varepsilon \)-insensitive loss function:

\[
L(y_i) = \begin{cases} 
0 & \text{if } |f(x_i) - y_i| < \varepsilon \\
|f(x_i) - y_i| - \varepsilon & \text{otherwise}
\end{cases}
\]

(2.3)

The minimization of expression (2.2) is implemented by introducing the slack variable \( \xi_i^- \) and \( \xi_i^+ \). Specifically, \( \varepsilon \)-Support Vector Regression (\( \varepsilon \)-SVR) solves the following quadratic programming problem [1]:

\[
\min_{\omega,b,\xi^-} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^{n} (\xi_i^- + \xi_i^+ )
\]

(2.4)

Subject to:

\[
y_i - (\omega^T \phi(x_i) + b) \leq \varepsilon + \xi_i^-
\]

\[
(\omega^T \phi(x_i) + b) - y_i \leq \varepsilon + \xi_i^+
\]

\[
\forall i, \xi_i^- \text{ and } \xi_i^+ \geq 0
\]

The solution to this minimization problem is of the form

\[
f(x) = \sum_{i=1}^{m} (\lambda_i - \lambda_i^*) K(x_i, x) + b
\]

(2.5)

Where \( \lambda_i \) and \( \lambda_i^* \) are the Lagrange multipliers associated with the constrains \( y_i - (\omega^T \phi(x_i) + b) \leq \varepsilon + \xi_i^- \) and \( (\omega^T \phi(x_i) + b) - y_i \leq \varepsilon + \xi_i^+ \) respectively.

The kernel function can be defined as:

\[
K(x_i, x_j) = \phi(x_i)^T \phi(x_j)
\]

(2.6)
The value of the kernel is equal to the inner product of two vectors \( x_i \) and \( x_j \) in the feature space \( \phi(x_i) \) and \( \phi(x_j) \).

Below are theKERNEL types:

- **Polynomial (homogeneous)**: \( k(x_i, x_j) = (x_i \cdot x_j)^d \)
- **Polynomial (inhomogeneous)**: \( k(x_i, x_j) = (x_i \cdot x_j + 1)^d \)
- **Gaussian Radial Basis Function**: \( k(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right) \), for \( \gamma > 0 \) (or \( \gamma = 1/2\sigma^2 \))
- **Hyperbolic Tangent**: \( k(x_i, x_j) = \tanh(\kappa x_i \cdot x_j + C) \) for some \( \kappa > 0 \) and \( C < 0 \)

The radial basis kernel is a popular choice in the SVM literature. Therefore our computations are based on such a kernel.[4,5]

### 2.2. Wavelet Analysis

The proposed WSVM is based on wavelet analysis. The principle of wavelet analysis is to express or approximate a signal (or function) by a family of functions generated by dilations and translations of a mother wavelet as follows [6]:

\[
k_{m,n}(z) = |m|^{-1/2} k\left(\frac{z-n}{m}\right)
\]

where \( m \) is a dilation factor; \( n \) is a translation factor; and \( k(z) \) is the mother wavelet, which satisfies the following condition [4, 5]:

\[
W_k = \int_{-\infty}^{\infty} \frac{|F(w)|^2}{|w|} dW(\infty)
\]

Where \( F(w) \) is the Fourier transform of \( k(z) \). The wavelet transform of a function \( g(z) \) can be expressed as

\[
W_{m,n}(g) = \langle g(z), k_{m,n}(z) \rangle
\]

where \( \langle \bullet \rangle \) denotes the dot product. The right-hand side of (2.9) means the decomposition of \( g(z) \) the function on a wavelet basis \( k_{m,n}(z) \), and \( W_{m,n}(g) \) are the coordinates of \( g(z) \) in the space spanned by \( k_{m,n}(z) \). Then the function \( g(z) \) can be reconstructed as follows [30]:

\[
g(z) = \frac{1}{W_k} \int_{-\infty}^{\infty} \alpha \int_{-\infty}^{\infty} W_{m,n}(g) k_{m,n}(z) dmdn
\]

Equation (2.10) can be approximated by taking the finite terms

\[
\tilde{g}(z) = \sum_{i=1}^{N} W_i \cdot k_{m,n}(z)
\]

where \( W_i \) are the reconstruction coefficients.

### 2.3. Model Evaluation

In this study, the performance of the models is evaluated by the index of the root mean squared error (RMSE) and mean absolute error (MAE). These indexes respectively define as follow.
3. Result and Discussion

For analysis, the minimum RMSE and MAE from the SVM and WSVM models are compared. The results have two parts. First part comparing the kernel functions for SVM and WSVM. (See Table 1). Second part, comparing between the results from Table 1. (See Table 2).

**Table 1.** The comparisons of RMSE and MAE between kernels function for SVM Model and WSVM model of Singapore tourist arrivals to Malaysia.

| Model | Kernel Function Linear | Kernel Function RBF |
|-------|------------------------|---------------------|
|       | RMSE | MAE | RMSE | MAE |
| SVM   | 0.07912 | 0.05865 | 0.1024 | 0.07926 |
| WSVM  | 0.04478 | 0.03138 | 0.1541 | 0.09963 |

From the Table 1, the value of RMSE and MAE of kernel function linear for SVM model and WSVM model are lower than those of kernel function RBF for SVM model and WSVM model. The results suggest that kernel function linear is superior for SVM and WSVM model. Figure 1 and 2 shows the testing and forecasting accuracy the kernel function linear for SVM and WSVM model. Figure 3 and 4 shows the testing and forecasting accuracy the kernel function RBF for SVM and WSVM model.

**Table 2.** The comparisons of RMSE and MAE between SVM Model and WSVM model of Singapore tourist arrivals to Malaysia.

| Model | RMSE | MAE |
|-------|------|-----|
| SVM   | 0.07912 | 0.05865 |
| WSVM  | 0.04478 | 0.03138 |

From the Table 2, the results shows that RMSE and MAE proposed model wavelet and support vector machine (WSVM) had the lower value than single model support vector machine (SVM). It means that the WSVM is the best model for predict monthly Singapore tourist arrival to Malaysia.
Figure 1. Testing accuracy and SVM forecasting accuracy for kernel function linear.

Figure 2. Testing accuracy and SVM forecasting accuracy for kernel function RBF.

Figure 3. Testing accuracy and WSVM forecasting accuracy for kernel function linear.

Figure 4. Testing accuracy and WSVM forecasting accuracy for kernel function RBF.

4. Conclusion
Overall, it is found that the best method to forecast the Singapore tourist arrivals to Malaysia by using WSVM with kernel function linear. The forecasts provided a rich body of information for ASEAN tourism and related sectors, to be used in strategic planning, decision making and support for sustainable development of the in Malaysia tourism. The study will contribute to the further development and understanding of tourism forecasting in general. Besides that, the study will also encourage more organizations and individuals to engage in the forecasting process.

References
[1] Kim G and Ri S Y, 2012 Hybrid Forecasting of Exchange Rate by Using Chaos Wavelet SVM-Markov Model and Grey Relation Degree arXiv: 1207.1547.
[2] Menon A K. Large-Scale Support Vector Machines: Algorithms and Theory
[3] Muhammad A and Shotar M M The Application Of Time Series Modelling To Some Major Economic Variables. Ph.D Qatar University
[4] Zhang Q H and Benveniste A 1992 Wavelet networks IEEE Trans. Neural Netw., vol. 3, no. 6, pp. 889–898
[5] Daubechies I 1990 The wavelet transform, time-frequency localization and signal analysis IEEE Trans. Inf. Theory, vol. 36, no. 5, pp. 961–1005
[6] Zhang L, Zhou W, and Jiao L 2004 Wavelet support vector machine *IEEE Trans. Syst., Man, Cybern.*, vol. 34, no. 1, pp. 34–39
[7] Rafidah A and Suhaila Y 2013 Modelling River Stream Flow Using Support Vector Machine *Applied Mechanics and Materials* Vol. 315, pp 602-605