Journal of Optimization in Industrial Engineering
Vol. 16, Issue 2, Summer & Autumn 2023, 157-165
DOI: 10.22094/JOIE.2023.1986767.2071

Research Article

A Comparative Simulation Study of Nonlinear Time Series Model for Forecasting Tourism Data

Ali Rafidah [a],*

[a] *Department of Technical Foundation, Universiti Kuala Lumpur (UniKL), Malaysian Institute of Industrial Technology, Persiaran Sinaran Ilmu, Bandar Seri Alam, 81750, Johor Bahru, Malaysia*

Received 22 May 2023; Revised 15 August 2023; Accepted 27 August 2023

Abstract

Simulation is a tool to evaluate the performance existing and proposed under configured conditions of the simulation data. A simulation process can be useful to test theories and understand behavior of the statistical methods. This study aimed to compare SVM, WSVM and EMDWSVM model in order to identify the best model for forecasting time series data based on 10 replicates on 2040 generated data of the SARIMA (3,1,3) (3,1,1) [12] model of Brunei data set. This SARIMA model comes from the lowest error between SARIMA models. The simulations were performed with three criteria namely root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE). The results of the study show the lowest error value for the EMDWSVM time series model and the performance of all measurements is small then other models. The results also proved that combination of three method EMDWSVM is the advanced forecasting techniques in all the considered situation in providing better forecasting accuracy, the application of an EMD-based combined model particularly with wavelet method reduction approach for tourist arrivals forecasting due to better prediction results and stability than those achieved from single and current hybrid models. Therefore, the modified the existing hybrid model WSVM combined with the empirical mode decomposition (EMD) to decrease the complexity of dataset to improve its prediction accuracy.

Keywords: EMDWSVM; WSVM; SVM; EMD

1. Introduction

The tourism industry is becoming one of the major sectors that contribute to the world's economic development. Over the last decade, the Asia Pacific region, which includes North-East Asia, South-East Asia and Oceania, has been the fastest growing tourism regions in the world. Tourism is said to be the world's largest industry in the World Travel and Tourism Report (WTTC) as the industry is not only able to contribute to the employment sector in creating job opportunities, but its contribution is also significant in generating wealth of a country. Therefore, tourism research in various fields like promoting, marketing, forecasting, and planning are indeed imperative and important. With the ability to forecast tourist arrivals, it will aid the government and other tourism related organizations in their future planning.

Tourism has become one of the largest and fastest growing industries in the modern world (Werthner and Ricci, 2004). Global economic development is primarily linked to the tourism industry particularly in the number of tourist arrivals (Song and Li, 2008 and Song et al., 2011). Malaysia is one of the most popular tourist destinations. Strong exchange rate and political stability have made Malaysia an affordable destination in Southeast Asia. Since the first launching of "Visit Malaysia Year" in 1990, Malaysian government has launched several tourism promotions programs to attract foreign tourists especially from Asia, Middle East, and Europe. (Nanthakumar, Han and Kogid, 2013). The tourism industry is an increasingly important industry for Malaysia; therefore, the policymakers and industry players have paid a close attention to the development of the tourism industry (Liang, 2014).

Frequently, in real world scenarios, due to the complexity of the system under investigation, it may not be possible to evaluate the system's behavior by applying analytical methods (Robert Davies and Tim Coole, 2014). Under such conditions, an alternative approach to model such a system is through creating a simulation. Briefly, simulation methods provide an alternative approach to studying system behaviors by creating an artificial replication or simulation of the real-world system (Stewart Robinson, 2005).

According to Haji, Sadik and Soleh (2018), this simulation study is used to obtain empirical results about the performance of statistical methods in certain scenarios, as contrasting to more general analytic results, which may cover many situations. It is not always possible, or may be complex, to obtain analytic results. Simulation studies are used to verify the actual models when wrong assumptions are made since data simulation can access the flexibility of the method. This is not always possible with analytic results, where results may apply only when data arise from a specific model (Morris, White and Crowther, 2018).

Realistic and accurate data collection is needed for any decision-making process. Setting up the simulation

*Corresponding author Email address: rafidahali@unikl.edu.my

experiments enables the researchers to produce large data sets that represent many situations and people's profiles (Norrulashikin, Yusof and Kane, 2018). By building a process that is easy to understand and varying the simulation parameters, the researchers can test the efficiency and robustness of the developed algorithm (Duchêne, Garbay and Rialle, 2003).

Other studies that used the simulation technique to verify the model building process in hydrology application include the work of Haji, Sadik and Soleh (2018) who proposed a simulation method nonstationary time series for rainfall. ARIMA non-seasonal model (p,d, and q) was used to generate simulation data. The simulated process generated numerous synthetics with similar length to the original data set.

In this research, combination wavelet, empirical mode decomposition (EMD) and support vector machine (SVM) had been used and compared with hybrid wavelet SVM (WSVM) and SVM model. The support vector machine (SVM) is proposed by Vapnik and his co-workers in 1995 through statistical learning theory. The SVM is a powerful methodology and has become a hot topic of intensive study due to its successful employed to solve most non-linear regression and time series problem and becoming increasingly in the modelling and forecasting of number tourist arrivals. Several studies have been carried out using SVM in modelling forecasting such as tourism, rainfall runoff modelling and flood stage forecasting. The standard SVM is solved using quadratic programming methods. However, this method is often time consuming and has higher computational burden because of the required constrained optimization programming and the application of SVMs for tourist forecasting has not been widely explored and SVM method also challenging to estimate qualitatively how many time steps into the past would allow the greatest efficiency, the values of time for tourist arrival are not known in advance (Rajaee et al., 2011). Therefore, this research offers evidence on the predictive ability and the profitability of abnormal returns of a new combination forecasting model using EMD, wavelet and SVM (named EMD_WSVM) on future tourist forecasting. The main objective of this paper is to provide best practices for validating an application by using simulation results.

2. Literature Review

2.1 Time Series Forecasting

In recent decades, time series analysis has always been critically suggested on many applications including the control of physical systems, process of engineering, biochemistry, environmental economic system, number tourist arrivals, company management and economy Ghalehkhondabi, Ardjmand, Young, and Weckman (2019). More importantly, the forecast of tourist arrivals is important since it will affect macroeconomic policies, business management and individual decision-making establishment (Croce,2018). Many experts and scholars have recently put forward a lot of methods based on nonlinear theories and their combinations for time series forecasting.

Forecasting methods based on the artificial neural networks (ANN) have been applied in the many fields mentioned above and have been a pioneer in the field of tourism data analysis (Nguyen, Fernandes, and Teixeira, 2022) and have been widely used in time series prediction models with support vector machine (SVM), for chaotic time series prediction. Existing research based on computational intelligence, such as artificial neural networks (ANN) (Yu, Yang and Liu, 2013; Nor, Nurul and Rusiman, 2018) and support vector regression (SVR) (Lingyu, Jun and Chunyu, 2021), for time series prediction indicates that the latter emerges as the winner, especially in short-term forecasting (Adamowski and Karapataki, 2010). However, computational intelligence-based forecasting models have their own shortcomings and disadvantages such as local minima and over-fitting in ANN models and sensitivity to parameter selection in both SVR and ANN models.

In view of the limitations for computational intelligence-based forecasting models, recently, hybrid empirical mode decomposition (EMD)-based modelling framework has established itself as a promising alternative for nonlinear and non-stationary time series modelling and prediction introduced by (Huang et al., 1998). EMD-based modelling framework has been successfully applied into many different areas, including tourism demand (Wang and Liu, 2022; Shabri,2016), tourism management (C. F. Chen, Lai and Yeh, 2012), and hydrology (Napolitano, Serinaldi and See, 2011).

In the last decades, the application of wavelet transformation for analyzing variations, periodicities and trends in time series has received numerous attentions. Before the invention of wavelet and fractal methods in 1980‟ s, Fourier analysis and statistical methods were used for studying the behavior of time series. However, in the present days, wavelets and wavelet-based multi-fractal formalism are becoming more favorable among the developed and developing countries for scientific studies of real-world problems particularly climatic data. In this study is focused on the application of wavelet methods for the forecasting of tourism arrivals.

2.2 Nonlinear classification

The original optimal hyperplane algorithm proposed by Vapnik in 1963 was a linear classifier. In 1992, Bernhard Boser, Isabelle Guyon and Vapnik suggested a way to create nonlinear classifiers by applying the kernel trick which was originally proposed by (Aizerman, Braverman and Rozonoer, 1964) to maximum-margin hyperplanes (Wu, 2010). The resulting algorithm is formally similar, except that every dot product is replaced by a nonlinear kernel function. This allows the algorithm to fit the maximum-margin hyperplane in a transformed feature space. The transformation may be nonlinear, and the transformed space may be of high dimensional, thus though the classifier is a hyperplane in the high-dimensional feature space, it may be nonlinear in the original input space (see Fig. 1).

2.3 Previous study on tourism

The importance of tourism forecasting in tourism planning and tourism policy formulation has been widely documented in studies such as Loeb, (1982) and Wong and Song (2012). Tourism forecasts may be generated by either quantitative approaches or qualitative approaches (WTO, 2016). The focus in this study is on quantitative forecasting methods. Tourist arrivals and tourist expenditures are still the common measure of tourism demand in the past decade. In his literature survey, Zhou et al., (1997) pointed out that amongst the 45 selected studies published since 1990, 37 of them used tourist arrivals as the dependent variable while only six employed tourist expenditures as the dependent variable.
During the past decades, several studies have developed methods of tourism demand time series predictions (Lim and Mcaleer, 2000; Louvieris, 2002; Brida and Garrido, 2011; Kulendran and Shan, 2012; Shu et al., 2014; Chinnakumy and Boonyasanaz, 2016; Liang, 2017). The most widely and comprehensive statistical methods used for time series forecasting are autoregressive moving average (ARMA) models (Nanthakumar and Ibrahim, 2010; Shu et al., 2014; Saayman and Botha, 2017). The popularity of the ARIMA model is due to its statistical properties as well as the well-known Box-Jenkins methodology; its forecasting capability and its wider information on time-related changes. The ARIMA model is only a class of linear model and thus it can only capture linear feature of time series data (Louvieris, 2002). However, most tourism time series data of practical relevance are nonlinear and chaotic in nature.
Kulendran and Shan, (2012) has pointed out that univariate forecasting techniques to forecast arrivals have no explanatory variables so the individual components are difficult to interpret. However, there are evidence that the forecasting records of many univariates' models have considerable forecasting accuracy and, in many cases, the variable pricing in structural models are difficult to predict (Gustavsson and Nordströrm, 2001).

3. Methodology/Materials

3.1 Simulation Data Generation

The simulation method in this research generate a SARIMA (p,d,q)(P,D,Q) of number tourist arrivals time series data sets. Let xt = (x1, x2,...,xt) be an observed univariate time series and t=1,2,...,n. Then the simulation method can be summarized as listed below:

Step 1: Define base process.
For seasonal ARIMA (SARIMA) processes with period s, a seasonal AR polynomial of order P, a seasonal MA polynomial of order Q and the seasonal difference operator of order D and the SARIMA(p,d,q)(P,D,Q)s is then defined by the equation:

$$\Phi(B^s)\varphi(B)(1-B)^d(1-B^s)^{\ddot{d}}X_t = \Theta(B^s)\theta(B)\varepsilon_t + \mu$$

Step 2: Identify the SARIMA model (p,d,q)(P,D,Q) which are ARIMA(p,d and q) represent the non-seasonal part of model and ARIMA(P,D and Q) represent seasonal part of model. In this simulation study, we have set the length of the data (n) is same with the actual data which is Brunei tourist arrival data.

Step 3: Simulate a long-time series data generating process from the base process and analyze it using
the proposed developed model.

```
oursimulation⟨sim.ssarima(order=list(ar=c(3,3),i=
c(1,1),ma=c(3,1),lags=c(1,12,constnt=TRUE,obs=2
52)
plot (our simulation)
```

3.2 Research method

Figure 2 shows the three models that used to forecast the simulation data. The three models that used in this study are SVM model, hybrid WSVM model and EMDWSVM model. The root means square errors (RMSE), mean absolute errors (MAE) and mean absolute percentage (MAPE) statistics used to evaluate EMDWSVM, conventional SVM and WSVM model. The correlation value shows the degree to which the two variables are linearly related. Different types of information about the predictive capabilities of the model are measured through RMSE, MAE and MAPE. The different program codes including wavelet toolbox are written in MATLAB language and R-code written the empirical mode decomposition for the development EMDWSVM model. The program codes are also written in MATLAB language and R-codes for other models; conventional SVM and WSVM in this study.

SVM Model

The procedure of developing a support vector machine for time series forecasting is illustrated in Fig. 2. From Fig. 2, the flow chart can be divided into three stages. This research was carried out based on this procedure. Stage 1: Data sampling – data training, validating, and testing data need to be collected.
Stage 2: Sample pre-processing - it included two steps: data division. In any model development process, familiarity with the available data is of the utmost importance. After that, the data collected were split into two sub-sets: training sample data and testing sample data which were used for model development and model evaluation respectively.

Stage 3: Implementation SVM training and learning - this phase included three main tasks: determination of SVM architecture, sample training and sample validation. It is the core process of an SVM model. In this stage, time-delay τ, embedding dimension d, ε, regularization constant λ and the choice of the kernel function can be determined.
Stage 4: Sample forecasting - when the three steps are complete, the SVM can be used as a forecaster or predictor for out-of-sample forecasting of time series.

WSVM Model

For the WSVM model inputs (see Fig. 2), two resolution levels from high frequency are employed in this study. The two resolutions were D1 and D2. To the writer's knowledge, there is no existing theory to tell how many resolution levels are needed for any given time series. According to Wang and Ding (2003), however, in general there is log (n) resolution level, where n is length of the time series. The WSVM is constructed in which the Ds of original input time series are the input of SVM, and the original output time series are the output of SVM. In this study, the type of wavelet is Daubechies wavelet (db2) had been carried out on the discrete wavelet transform. Daubechies wavelet is continuous in extracting the signals which in contrast to other type of wavelet such as Haar wavelet. Haar wavelet was tending to exhibit jump discontinuity the extracted signals (Addison, 2001).
In the wavelet analysis, the original signal time series is broken into the low frequency which is presented by a scale of approximate mode and the five resolution levels of high frequency. The signal is then represented by its features, which are becoming the wavelet coefficients. The sets of wavelet coefficients that generated from both high and low frequencies were sum up and tally for each input variables which were for data simulation s1-s10 tourist arrivals of Brunei. The total wavelet coefficients from every single input variable were divided into training and testing data. The input variables that carry through the wavelet coefficients then iterated into SVM model. Figure 2 shows the flowchart of implementing the algorithm.

EMDWSVM Model

For this model (see Fig.2), in step 1, decomposition data into EMD method. Data simulations were decomposed into n-IMF components and one residue, using EMD technique. The extracted IMF components represent a range of high to low frequencies where each IMF component represents the local characteristic time scale by itself. Then, the IMF and residual are identified as stochastics or deterministic components based on is deterministic components otherwise stochastics (Aamir and Shabri, 2018). The average mutual information (AMI) plot is also used for visual stochastic and deterministic components. The stochastic components be dealt individually, and all deterministic IMFs be summed to make another component.
Then step 2, the data decompose into wavelet method. In this step, the IMF stochastics and deterministic component decomposed into Discrete Wavelet Transformation (DWT). In the wavelet analysis two the original signal time series is broken into the low frequency which presented by a scale of approximate mode and the two resolution levels of high frequency.
In the final step, which is forecasting step, the SVM forecasting technique was employed on the chosen signals. The new dataset is further input into the SVM forecasting technique model accordingly and compared it with single model SVM using several performance measurements.

3.3 Error measurement

To evaluate the performance of the constructed model, error measurements were used to compare the performance and accuracy of each model used. Different statistical tests are performed because each method measures the error for different purpose. There were three measurement errors used namely Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE). These measurements are employed as performance indicators and calculated based on the tested data. For instance, EMD-WSVM proposed hybrid model is considered the best alternative model for forecasting univariate time series data if it gives lowest values of MAE, RMSE and MAPE.

Mean Absolute Error (MAE)

The MAE measures the average magnitude of the errors in a forecasting model, without considering their direction. MAE is the average over the verification sample of the absolute values of the differences between forecast and the corresponding observation. It is a linear score which means that all the individual differences are weighted equally in the average. (S.S. et al., 2009) It is given as;

$$\text{MAE} = \frac{1}{N}\sum_{t=1}^{n}|y_t - \hat{y}_t|$$

where;

N = The number of observations
y_t = The observed values at the time t.
$\hat{y}_t$= The forecast values at the time *t*.

Root Mean Squared Error (RMSE)

RMSE is a quadratic scoring rule which measures the average magnitude of the error. The RMSE serves to aggregate the magnitudes of the errors in predictions for various times into a single measure of predictive power. RMSE is a good measure of accuracy, but only to compare forecasting errors of different models for a variable and not between variables, as it is scale dependent. RMSE were written as;

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{n}(y_t - \hat{y}_t)}$$

where;

N = The number of observations
y_t= The observed values at the time *t*.
$\hat{y}_t$= The forecast values at the time *t*.

Mean Absolute Percentage Error (MAPE)

In statistical test MAPE is favoured and has gained popularity in the literature because it is not prone to change in the magnitude of time series data (Park, Yoon and Kim, 2000).

$$\text{MAPE} = \frac{1}{N}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \times 100$$

where;

N = The number of observations
y_t= The observed values at the time *t*.
$\hat{y}_t$= The forecast values at the time *t*.

According to (Baldigara, 2013)"a rough scale for the accuracy of a model can be based on MAPE" following the suggestions gave in the Table 1 below.

Table1
Scale for the accuracy (Source: Baldigara, 2013)

MAPE	Forecasting accuracy
Less than 10%	Highly accurate
10-20%	Good
20-50%	Reasonable
Greater than 50%	Inaccurate

4. Results and Findings

In this section, simulation studies on prediction performance of SVM, WSVM, and EMDWSVM model were conducted. A simulation of a system is the operation of a model of the system where the model can be reconfigured and experimented (Carson, 2005; Altiok and Melamed, 2007). Simulation is a tool to evaluate the performance existing and proposed under configure conditions of the simulation data.

The simulation dataset used in this study were generated from the best SARIMA model of Brunei country. The total numbers of the datasets used were the same for each model. The inputs for SVM, WSVM and EMDWSVM were based on the PACF graphs.

Table 2 shows the results of RMSE, MAE and MAPE for SVM, WSVM, and EMDWSVM model. The winning model was the model that has the lowest RMSE, MAE and MAPE for prediction. The summary of the simulation study is shown in Table 2 where it shows the win frequency of model for Brunei country

Table 3 shows the win frequency for EMDWSVM, SVM, and WSVM models based on the RMSE, MAPE and MAE. From Table 2, it can be summarized that EMDWSVM model outperformed other models for dataset from Brunei. Meanwhile, SVM model for s6 outperformed all the other models. On average, the winning frequency percentages for EMDWSVM, SVM, and WSVM model were 90%, 10%, and 0% respectively. In addition, EMDWSVM model outperformed SVM model when the residual p-value Ljung Box-statistics were independent which was indicated as the values above the dotted blue line (see Figure 3). Besides that, SVM model outperformed EMDWSVM when the residuals were dependent (see Figure 4).

Other than that, the EMDWSVM model has a higher winning percentage compared to WSVM model where 90% for EMDWSVM model and 10% and 0% for SVM and WSVM model, respectively were recorded. It shows that the implementation of EMD into WSVM model has improved the prediction performance of WSVM and SVM model. Therefore, based on win frequency percentage, EMDWSVM model was the most efficient and robust in terms of prediction performance compared to other models when the dataset was decomposed with wavelet and SVM combined EMD method.

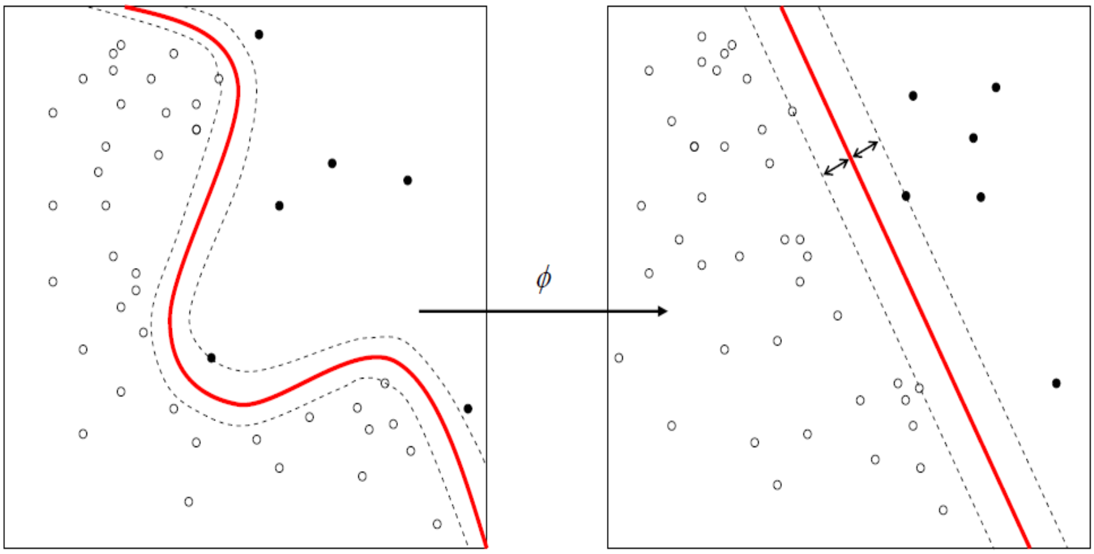

Fig. 1. Kernel machines are used to compute a non-linearly separable function into higher dimension linearly separable function.

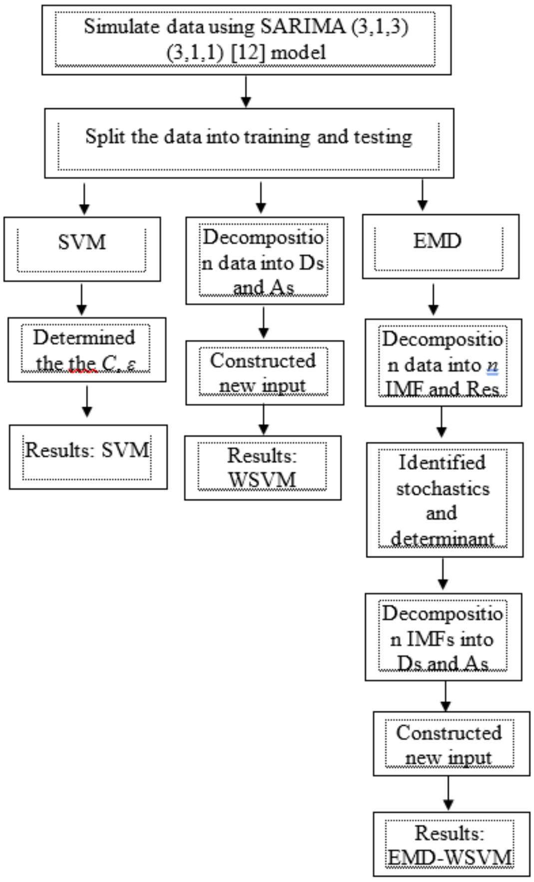

Fig. 2. Model for forecasting

Table 2
RMSE, MAPE and MAE of SVM, WSVM and EMDWSVM

DATA SIMULATION	MODEL	RMSE	MAPE (%)	MAE	DATA SIMULATION	MODEL	RMSE	MAPE (%)	MAE
s1	SVM	11889.82	43.51	11714.29	s6	SVM	1282.36	13.10	1242.64
	WSVM	11435.41	41.78	11255.08		WSVM	576.18	5.62	534.82
	EMDWSVM	**553.60**	**1.78**	**474.97**		**EMD_WSVM**	**246.38**	**1.72**	**172.27**
s2	SVM	3248.76	35.18	2878.51	s7	SVM	32136.96	51.60	30781.36
	WSVM	3248.76	35.18	2878.51		WSVM	26713.46	44.51	26250.73
	EMDWSVM	**1173.13**	**13.73**	**1026.41**		**EMDWSVM**	**3872.35**	**6.32**	**3739.88**
s3	SVM	17471.43	48.32	16411.83	s8	SVM	21679.80	45.12	19804.94
	WSVM	17255.14	47.58	16191.05		WSVM	22378.83	48.35	20989.96
	EMDWSVM	**2310.43**	**6.37**	**2143.92**		**EMDWSVM**	**3211.40**	**6.36**	**2740.10**
s4	SVM	28.86	2.41	22.22	s9	**SVM**	**2291.06**	**11.45**	**2100.21**
	WSVM	21.21	1.85	17.05		WSVM	4206.14	18.72	3278.19
	EMDWSVM	**20.67**	**1.59**	**14.65**		EMDWSVM	2694.26	13.64	2166.01
	SVM	405.19	6.81	378.02		SVM	745.66	16.04	697.41
s5	WSVM	341.73	5.89	326.61	s10	WSVM	824.05	17.94	781.51
	EMDWSVM	**207.17**	**3.49**	**193.79**		**EMDWSVM**	**197.49**	**3.67**	**159.51**

Table 3
Prediction Model Winning Frequency Based on RMSE, MAPE and MAE

Data	Win Frequency		
	EMDWSVM	SVM	WSVM
Simulation	9	1	0

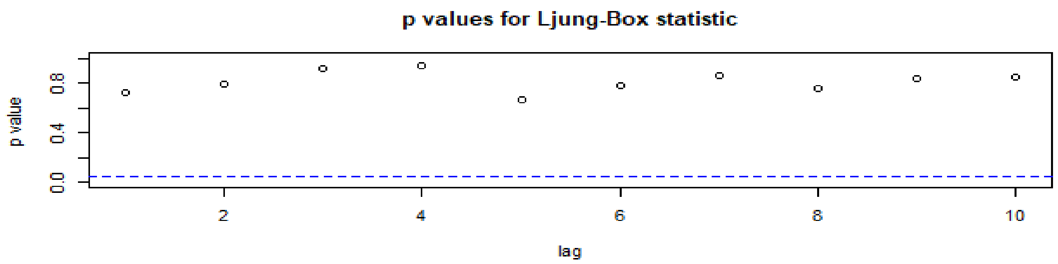

Fig. 3. p-Value for Ljung Box Statistics for residual are Independent

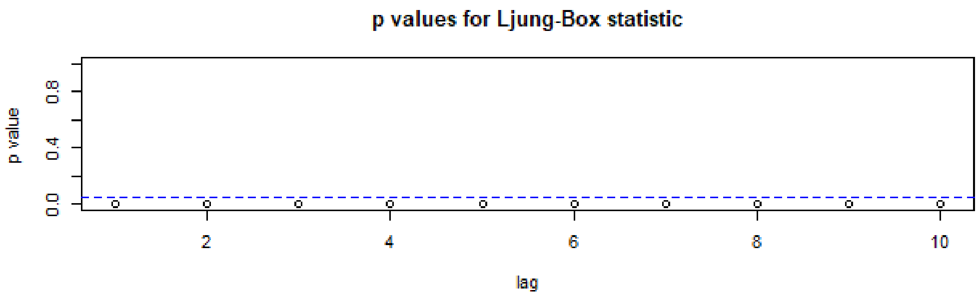

Fig. 4. p-Value for Ljung Box Statistics for residual are Dependent

To support the findings, Figure 4 shows the boxplot for each model for simulation data tourist arrivals. From Figure 5, the EMDWSVM model was the best as the spacing between different parts of the box were smaller and closer to zero. The results obtained in this study indicate that EMDWSVM model is a powerful forecasting tool and can produce better prediction performance compared to SVM and WSVM modelling.

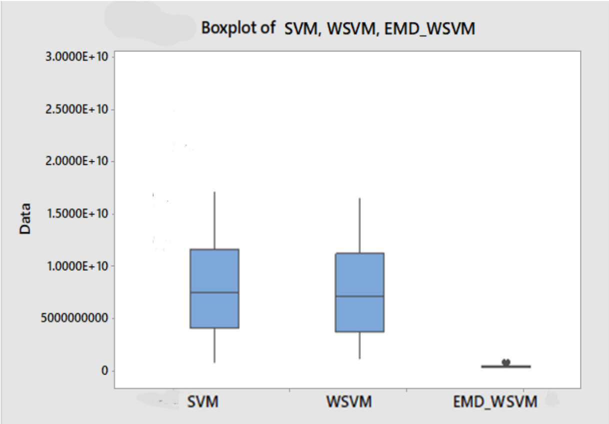

Fig. 5. The boxplot for predicted tourist arrivals using SVM, WSVM and EMD_WSVM

5. Conclusion

It can be summarized that the proposed EMDWSVM model was the best forecasting model for tourist arrivals for most of the simulation data of Brunei. Besides, this study has proven that the combination of EMD with wavelet and SVM is possible and can provide good prediction accuracy as previous researchers have only combined wavelet with SVM model or EMD with SVM model.

Simulation studies have been conducted to measure how much EMD reduces the complexity of WSVM model. The simulation results show that the proposed model EMDWSVM reduced the complexity of WSVM model. However, the SVM was more superior when the residual Ljung Box was dependent. The results also indicated that EMDWSVM model has proven to be the most efficient and robust model due to its superior performance compared to the other models for all simulation data.

Based on the findings of this study, it is suggested that EMD can combined with SVM in developing tourist arrival forecasting model in the future and different types of mother wavelet can be considered in decomposing the tourist arrivals. The other mother wavelets are Haar wavelet, Symlets wavelet and Coiflets wavelet.

Acknowledgement: Short Term Research Grant UNIKL for funding this project 2020 – 2021. Project ID: str10958

References

Adamowski, J., & Karapataki, C. (2010). Comparison of multivariate regression and artificial neural networks for peak urban water-demand forecasting: evaluation of different ANN learning algorithms. Journal of Hydrologic Engineering, 15(10), 729-743.

Aamir, M., & Shabri, A. (2018). Improving crude oil price forecasting accuracy via decomposition and ensemble model by reconstructing the stochastic and deterministic influences. Advanced Science Letters, 24(6), 4337-4342.

Aizerman, M. A. (1964). Theoretical foundations of the potential function method in pattern recognition learning. Automation and remote control, 25, 821-837.

Altiok, T., & Melamed, B. (2010). Simulation modeling and analysis with Arena. Elsevier.

Baldigara, T. (2013). Forecasting Tourism Demand inCroatia: A Comparison of Different Extrapolative Methods. Journal of Business Administration Research, 2(1), 84.

Brida, J. G., & Garrido, N. (2011). Tourism forecasting using SARIMA models in Chilean regions. International Journal of Leisure and Tourism Marketing, 2(2), 176-190.

Chinnakum, W., & Boonyasana, P. (2016). Forecasting International Tourism Demand in Thailand. Thai Journal of Mathematics, 231-244.

Croce, V. (2018), “With growth comes accountability:could a leisure activity turn into a driver for sustainable growth?”, Journal of Tourism Futures, Vol. 4 No. 3, pp. 218-32.

Carson, J. S. (2005, December). Introduction to modeling and simulation. In Proceedings of the Winter Simulation Conference, 2005. (pp. 8-pp). IEEE.

Chen, C. F., Lai, M. C., & Yeh, C. C. (2012). Forecasting tourism demand based on empirical mode decomposition and neural network. Knowledge-Based Systems, 26, 281-287.

Duchêne, F., Garbay, C., & Rialle, V. (2003, October). An hybrid knowledge-based methodology for multivariate simulation in home health telecare. In Proc. of the Joint Workshop Intelligent Data Analysis in Medicine and Pharmacology (IDAMAP) of the 9th Artificial Intelligence in Medicine Europe conference (AIME) (pp. 87-94).

Gustavsson, P., & Nordström, J. (2001). The impact of seasonal unit roots and vector ARMA modelling on forecasting monthly tourism flows. Tourism Economics, 7(2), 117-133.

Ghalehkhondabi, I., Ardjmand, E., Young, W. A., & Weckman, G. R. (2019). A review of demand forecasting models and methodological developments within tourism and passenger transportation industry. Journal of Tourism Futures, 5(1), 75-93.

Haji, H. A., Sadik, K., & Soleh, A. M. (2018). A Comparative Simulation Study of ARIMA and Fuzzy Time Series Model for Forecasting Time Series Data.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., ... & Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences, 454(1971), 903-995.

Jensen, A., & la Cour-Harbo, A. (2001). Ripples in mathematics: the discrete wavelet transform. Springer Science & Business Media.

Kulendran, N., & Witt, S. F. (2001). Cointegration versus least squares regression. Annals of Tourism Research, 28(2), 291-311.

Liang, Y. H. (2014). Forecasting models for Taiwanese tourism demand after allowance for Mainland China tourists visiting Taiwan. Computers & Industrial Engineering, 74, 111-119.

Loeb, P. D. (1982). International travel to the United States: an econometric evaluation. Annals of Tourism Research, 9(1), 7-20.

Lim, C., & McAleer, M. (2000). A seasonal analysis of Asian tourist arrivals to Australia. Applied Economics, 32(4), 499-509.

Louvieris, P. (2002). Forecasting international tourism demand for Greece: A contingency approach. Journal of Travel & Tourism Marketing, 13(1-2), 21-40.

Liang, Y. H. (2016, July). Using the combined model for forecasting the tourism demand. In 2016 International Conference on Machine Learning and Cybernetics (ICMLC) (Vol. 2, pp. 612-615). IEEE.

Loganathan, N., & Ibrahim, Y. (2010). Forecasting international tourism demand in Malaysia using Box Jenkins Sarima application. South Asian Journal of Tourism and Heritage, 3(2), 50-60.

Loganathan, N., Han, A. S., & Kogid, M. (2013). Demand for Indonesia, Singapore and Thailand tourist to Malaysia: seasonal unit root and multivariate analysis. International Journal of Economics and Empirical Research (IJEER), 1(2), 15-23.

Lingyu, T., Jun, W., & Chunyu, Z. (2021). Mode decomposition method integrating mode reconstruction, feature extraction, and ELM for tourist arrival forecasting. Chaos, Solitons & Fractals, 143, 110423.

Morris, T. P., White, I. R., & Crowther, M. J. (2019). Using simulation studies to evaluate statistical methods. Statistics in medicine, 38(11), 2074-2102.

Norrulashikin, S. M., Yusof, F., & Kane, I. L. (2018). Performance Evaluation of a New Hybrid Multivariate Meteorological Model Analysis: A Simulation Study. MATEMATIKA: Malaysian Journal of Industrial and Applied Mathematics, 73-85.

Nor, M. E., Nurul, A. I., & Rusiman, M. S. (2018, April). A hybrid approach on tourism demand forecasting. In Journal of Physics: Conference Series (Vol. 995, No. 1, p. 012034). IOP Publishing.

Nguyen, L. Q., Fernandes, P. O., & Teixeira, J. P. (2022). Analyzing and forecasting tourism demand in Vietnam with artificial neural networks. Forecasting, 4(1), 36-50.

Napolitano, G., Serinaldi, F., & See, L. (2011). Impact of EMD decomposition and random initialisation of weights in ANN hindcasting of daily stream flow series: an empirical examination. Journal of Hydrology, 406(3-4), 199-214.

Rajaee, T., Nourani, V., Zounemat-Kermani, M., & Kisi, O. (2011). River suspended sediment load prediction: application of ANN and wavelet conjunction model. Journal of Hydrologic Engineering, 16(8), 613-627.

Robert Davies, Tim Coole, D. O. (2014) 'The Application of Time Series Modelling and Monte Carlo Simulation: Forecasting Volatile Inventory Requirements', Applied Mathematics, i(May), pp. 1152–1168.

Robinson, S. (2005). Distributed simulation and simulation practice. Simulation, 81(1), 5-13.

Shabri, A., & Suhartono. (2012). Streamflow forecasting using least-squares support vector machines. Hydrological Sciences Journal, 57(7), 1275-1293.

Shu, M. H., Hung, W. J., Nguyen, T. L., Hsu, B. M., & Lu, C. H. U. N. W. E. I. (2014). Forecasting with Fourier residual modified ARIMA model-An empirical case of inbound tourism demand in New Zealand. WSEAS Transactions on Mathematics, 13(1), 12-21.

Saayman, A., & Botha, I. (2017). Non-linear models for tourism demand forecasting. Tourism Economics, 23(3), 594-613.

Song, H., & Li, G. (2008). Tourism demand modelling and forecasting—A review of recent research. Tourism management, 29(2), 203-220.

Song, H., Li, G., Witt, S. F., & Athanasopoulos, G. (2011). Forecasting tourist arrivals using time-varying parameter structural time series models. International Journal of Forecasting, 27(3), 855-869.

Vapnik, V., Golowich, S. E., & Smola, A. (1997). Support vector method for function approximation, regression estimation, and signal processing. Advances in neural information processing systems, 281-287.

Wang, W., & Ding, J. (2003). Wavelet network model and its application to the prediction of hydrology. Nature and Science, 1(1), 67-71.

Werthner, H., & Ricci, F. (2004). E-commerce and tourism. Communications of the ACM, 47(12), 101-105.

Wu, Q. (2010). Product demand forecasts using wavelet kernel support vector machine and particle swarm optimization in manufacture system. Journal of computational and applied mathematics, 233(10), 2481-2491.

Wong, K. K. F., & Song, H. (2012). Tourism forecasting and marketing. Psychology Press.

Wang, H., & Liu, W. (2022). Forecasting Tourism Demand by a Novel Multi-Factors Fusion Approach. IEEE Access, 10, 125972-125991.

Yu, S. P., Yang, J. S., & Liu, G. M. (2013). A novel discussion on two long-term forecast mechanisms for hydro-meteorological signals using hybrid wavelet-NN model. Journal of hydrology, 497, 189-197.

Zhou, D., Yanagida, J. F., Chakravorty, U., & Leung, P. (1997). Estimating economic impacts from tourism. Annals of Tourism Research, 24(1), 76-89.