Indian Journal of Science and Technology, *Vol 9(36), DOI: 10.17485/ijst/2016/v9i36/97773, September 2016*
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

A Hybrid of EEMD and LSSVM-PSO model for Tourist Demand Forecasting

Ani Shabri*

Department of Mathematics, Science Faculty, University Technology of Malaysia, Skudai, Johor, 81310, Malaysia; ani@utm.my

Abstract

In this research, hybrid model of Least Square Support Vector Machine (LSSVM) and Ensemble Empirical Mode Decomposition (EEMD) are presented to forecast tourism demand in Malaysia. Foremost, the original series of tourism arrivals data was separated using EEMD technique into residual and Intrinsic Mode Functions (IMFs) components. Next, both of IMFs and residual components were forecasted using Particle Swarm Optimization (LSSVM–PSO) method. In the end, the predicted result of IMFs and residual components from LSSVM–PSO method are sum together to produce the forecasted value for tourism arrivals in Malaysia. Empirical results showed that the presented model in this paper outperform individual forecasting model. The result indicated that LSSVM–PSO is a promising tool in time series forecasting by having the presence of non-stationary and non-linearity in the time series data.

Keywords: Article Swarm Optimization Forecasting, Ensemble Empirical Mode Decomposition, Forecasting, Least Square Support Vector Machines, Mutual Information, Tourism Demand

1. Introduction

Over the last few decades ago, one of the most swiftly rising aspect in global industries is tourism. A reliable forecasting for tourism demand is needed in service and tourism industry because it will accommodate information for the government and business sectors to devise a plan which can improve the tourism industry in the country. However, the tourism data is highly known for the presence of non-stationary and non-linearity in the data. Therefore, it may have a poor forecasting performance using traditional statistical and econometric models.

Artificial Neural Networks (ANN) model is extensively used method for tourism demand forecasting specifically to overcome the limitations of the linear models to forecast where nonlinear characteristics is presence in real tourism data[1–11]. Although ANN has the upper hand of reliable forecasting, their prediction in a few particular situation is conflicting and suffers from some weakness, such as local minima, over fitting and sensitiveness to parameter selection[12].

Another popular and useful method of AI for forecasting is Least Squares Support Vector Machines (LSSVM). LSSVM is a modified version of Support Vector Machines (SVM) proposed by[13], where it replaces the inequality constraints by equality constraints in solving a quadratic programming[14]. Extensive empirical studies indicated that LSSVM is as good as SVM in terms of generalization performance[15]. Nowadays, LSSVM is broadly applied in numerous numbers of fields namely nuclear energy consumption forecasting[16], container throughput forecasting[17], electrical load forecasting[18], tourism forecasting[19], stream flowforecasting[20] and industrial application[21].

Recently, the Ensemble Empirical Mode Decomposition (EEMD) introduced by[22] demonstrated it efficiency in extracting important characteristic information from non-stationary and non-linear that exist in time series da. Implementations of EEMD, it facilitates to decompose any complex data into sub-series of Intrinsic Mode Functions (IMFs). IMFs components are easier and more reliable to forecast because the IMFs components

**Author for correspondence*

have stronger correlations and simpler frequency. There are numerous number of successful applications EEMD in many different kind of time series data as described in the literature[16,23,24]. However, the applications of EEMD studies in tourism demand are mainly limited.

The main objective in this paper is to assess prediction accuracy of LSSVM model integrated with EEMD for tourist demand forecasting. Therefore, EEMD is used to decompose tourist arrival data. Then, LSSVM methods are applied to each sub-series obtain previously, and the end result of forecasted value can be acquired by the summation of predicted value of all LSSVM models. To guarantee broader applications of the presented model, the monthly tourist arrival data from Singapore to Malaysia are investigated.

2. Methodology Formulation

2.1 Mutual Information

The Mutual Information (MI) is a measures where it specify the amount of information obtained from X in the presence of Y. MI can be expressed as:

Formula (1):

$$I(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\log\frac{f(x,y)}{f(x)f(y)}dxdy$$

Where $f(.)$ a probability distribution is function and $f(x,y)$ is joint distribution function. Consequently, the Formula (1) can't be directly computed but it can be estimated. There are various methods for entropy estimation such as Miller Madow, shrink, Pearson and Spearman correlation has been introduced. In this paper, the Spearman correlation is used. Spearman Correlation is a distinct case of the Pearson correlation where the data are transformed into rankings before estimated the coefficient. The MI of two variables X_i and X_j based on Spearman correlation is given by:

Formula (2):

$$I(X_i,X_j) = -\frac{1}{2}\log(1-R^2)$$

Where R is Spearman correlation. After the MI is calculated, the second step, Context Likelihood of Relatedness (CLR) algorithm is implemented in MI to compute a score where it is used to determine the best input variables for LSSVM model. The CLR algorithm[25] computes the MI for each pair (X_i,X_j) and then a score, z_{ij} related to MI values is computed. The score z_{ij} is given by:

Formula (3):

$$z_{ij} = \sqrt{z_i^2 + z_i^2}$$

where $z_i = \max\left(0, \frac{I(X_i,X_j)-\mu_i}{\sigma_i}\right)$

and μ_i and σ_i are the mean and the standard deviation of the empirical distribution of the MI values respectively, , $j = 1, 2, 3, \ldots, n$.

2.2 The Least Square Vector Machines Model

LSSVM is an innovative version of SVM modified by[13]. LSSVM includes the resolution of a quadratic optimization issue with a least squares loss function and equality constraints as a replacement for of inequality constraints. LSSVM is briefly introduced in this part. Consider a sample data set (x_i, y_i) with input $x_i \in R^n$ and output $y_i \in R$ is expressed as below:

Formula (4):

$$\min \frac{1}{2}w^T w + \frac{\gamma}{2}\sum_{i=1}^{n} e_i^2$$

Subject to the equality constraints:

$$y(x) = w^T\phi(x_i) + b + e_i,\ i = 1, 2, \ldots, n$$

Lagrange function constructed to solve the optimization problem:

Formula (5):

$$L(w,b,e,\alpha) = \frac{1}{2}w^T w + \frac{\gamma}{2}\sum_{i=1}^{n} e_i^2$$

$$-\sum_{i=1}^{n}\alpha_i\{w^T\phi(x_i) + b + e_i - y_i\}$$

Where α_i is Lagrange multipliers. The solution of Formula (5) can be acquired by partially differentiating with respect to w, b, e and α_i. By eliminating w and e_i, we can obtain the following equation:

Formula (6):

$$\begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & K(x_1,x_1)+\frac{1}{\gamma} & \cdots & K(x_1,x_n) \\ \vdots & \vdots & \cdots & \vdots \\ 1 & K(x_n,x_1) & \cdots & K(x_n,x_n)+\frac{1}{\gamma} \end{bmatrix} \times \begin{bmatrix} b \\ \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} 0 \\ y_1 \\ \vdots \\ y_n \end{bmatrix}$$

Where $K(\,,\,)$ represents the kernel function and satisfies Mercer's theorem. Then, the parameters α_i $(i = 1, 2, \ldots, n)$ and b can be solved from the equation and the regression prediction function can be expressed by the following equation:

Formula (7):

$$y(x) = \sum_{i=1}^{n} \alpha_i K(x_i, x) + b$$

Where $K(x_i, x) = \exp(-\gamma \|x_i - x\|^2)$ is the Gaussian Radial Basis Function (RBF).

2.3 Particle Swarm Optimization Algorithm

PSO is a relatively recent heuristic pursuit which was introduced recently by[26], where it was motivated by social actions of bird collecting and fish swarms. The PSO begins through a random population and searchers for fitness optimum. To obtain the optimum solution, each particle regulates the route through the best knowledge which it has found (pbest) and the best knowledge been found by all other associates (gbest). Therefore, the particle goes around in a multidimensional space to the good area.

Each particle contains of three vectors: The position for ith individual particle are denoted as $X_i = (x_i^{(1)}, x_i^{(2)}, \ldots, x_i^{(D)})$, the best earlier position pbest that the ith particle has searched is $P_i = (p_i^{(1)}, p_i^{(2)}, \ldots, p_i^{(D)})$, the fly velocity of the ith is $V_i = (v_i^{(1)}, v_i^{(2)}, \ldots, v_i^{(D)})$. Each particle performance is obtained using a fitness function. When the procedure is repeated, the ith particle at iteration t are updated by:

Formula (8):

$$v_i^d(t+1) = \omega' \times v_i^d(t) + c_1 \times \phi_1 \times [p_i^d(t) - x_i^d(t)]$$

$$+ c_2 \times \phi_2 \times [p_g^d(t) - x_i^d(t)]$$

$$x_i^d(t+1) = x_i^d(t) + v_i^d(t+1)$$

Where φ_1 and φ_2 are stochastic value of [0, 1], ω is called inertia weight and c_1 and c_2 are acceleration constants. For PSO, the particle keeps changing their position until the best position is determined or the maximum iteration had been reached. Figure 1 illustrates the PSO algorithm.

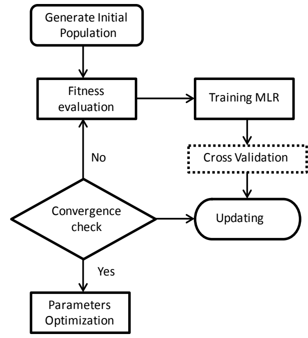

Figure 1. Flowchart of PSO algorithm.

2.4 Ensemble Empirical Mode Decomposition

EEMD is an improved method from Empirical Model Decomposition (EMD)[22]. Objective of EEMD is to decompose the simultaneous oscillatory function that is IMFs and a residual from the original data. If IMF meets two conditions: 1. The number of extreme values and zero-crossings either are equivalent or contrast at the utmost is one; and 2. At every point, if the mean value of the envelope constructed by the local maxima and minima must be zero. Thus, the original data $y(t)$ can be put as the sum of k IMFs and a residual:

Formula (9):

$$y(t) = \sum_{i=1}^{k} h_i(t) + r_k(t)$$

Where $h_i(t)$ represents IMFs, k is the number of IMFs and $r_k(t)$ denotes the final residual. EEMD includes an extra step for addition of white noise which differs from EMD where the added white noise effect can be contained by using well known statistical rule[22].

Formula (10):

$$\varepsilon_n = \frac{\varepsilon}{\sqrt{N}}$$

Where N is the ensemble number, ε is the amplitude of the added white noise and ε_n is the standard deviation of error between original series and the corresponding IMFs.

2.4 The Hybrid EEMD-LSSVM Model

The following Figure 2 describes the process of EEMD-LSSVM forecasting method. As can be seen from

Figure 2, the EEMD-LSSVM forecasting can be described as following steps:

- First, apply the EEMD technique to decompose the original data into k-IMFs and a residual component $r_k(t)$.
- Then, LSSVM model used to form a forecasting model for each extracted IMF and the residual components. Each LSSVM model forecasted one day ahead of the components.
- In the end, the forecasted values of all extracted IMF and residual components gained by LSSVM model are summed together to produce the final forecasting for the originals time series.

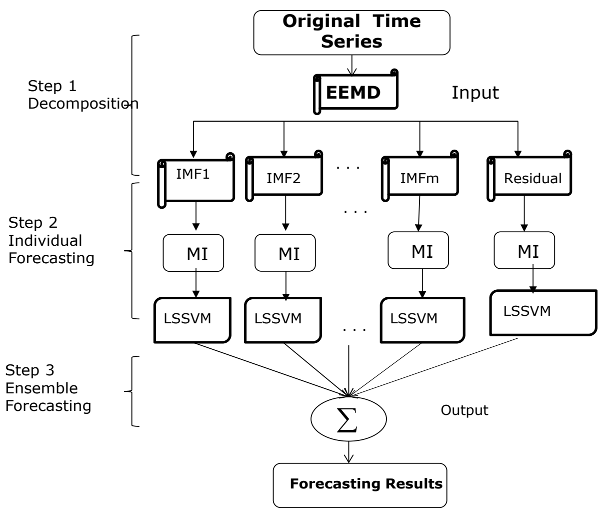

Figure 2. The hybrid EEMD-LSSVM model.

3. Experimental Results

3.1 Data Sets

The data use in this study was obtained from Malaysia tourist authorities. The data contains number of tourist arrived in Malaysia from Singapore in monthly period. The period of data were taken from January 2000 to December 2014. The data is shown on Figure 3. The data is separated into training and testing. The training data set consists of data within period of January 2000 to December 2012 and testing data consists of data from January 2013 to December 2014. The testing data used to obtain the unknown parameters of proposed models and testing data to evaluate the forecasting performance of the models.

3.2 Performance criteria

In order to evaluate the forecasting performance of different models, and Mean Absolute Percentage Error (MAE) and Root Mean Squared Error (RMSE) are applied as the criteria for assessing the prediction accuracy, which are calculated as:

Formula (11):

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}$$

and $$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|y_t - \hat{y}_t\right|$$

Where n is the sample size, y_t is the observation and $\hat{y}_t$ is the predicted value in the tth month. RMSE and MAE is assessor of deviation between predicted values and original observation. Model that has lower value of both RMSE and MAE indicated the reliability of the model.

3.3 Forecasting Results

In this paper, the EEMD algorithm is employed through R software package that is EEMD library. By employing the EEMD technique, original data of monthly tourist arrivals are decomposed into both IMFs and residual components. The plot of EEMD outcomes are showed in Figure 3. Figure 3 shows six IMFs and one residual component.

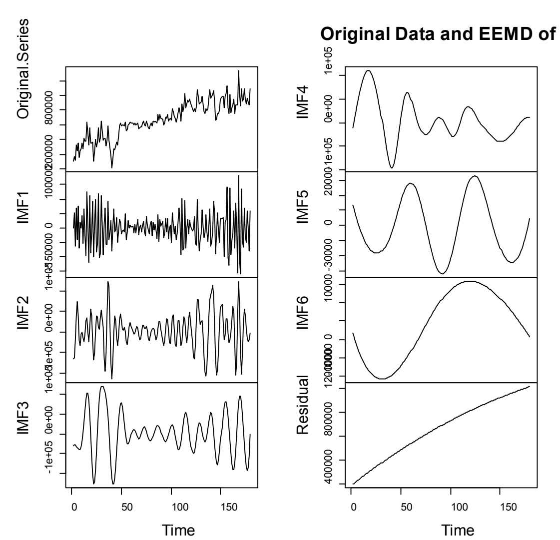

Figure 3. Original series and EEMD of tourist arrivals series.

In this paper, the parameters of LSSVM models were estimated using the PSO in the training phase. When LSSVM is used to forecast, the main problem arise on how to choose the input variables. There are no rules in the selection of input variables for developing LSSVM model. For LSSVM modelling, the Mutual Information (MI) using CLR algorithm is applied to decide the input variables. Entire data series were normalized into the range of -1 and 1. In this study, the most well-known kernel function implemented in LSSVM model is Gaussian RBF to obtain the optimal model parameters. To overcome parameter sensitiveness, PSO method is employed to determine values of optimal parameters γ and σ^2 which yield the lowest error in the training dataset.

In Table 1, for time series of monthly tourist arrivals from Singapore, the EMD–LSSVM model obtained the best MAE and RMSE of 44352.32 and 6313.5 respectively. From the result that had been obtained, the presented model EMD-LSSVM model was capable to increase the prediction accuracy compare to single LSSVM model which is 22.9% and 14.2% for MAE and RMSE.

Table 2. Comparison of LSSVM and EEMD-LSSVM

Model	MAE	RMSE
LSSVM	57562.99	73645.53
EEMD-LSSVM	44352.32	63173.05

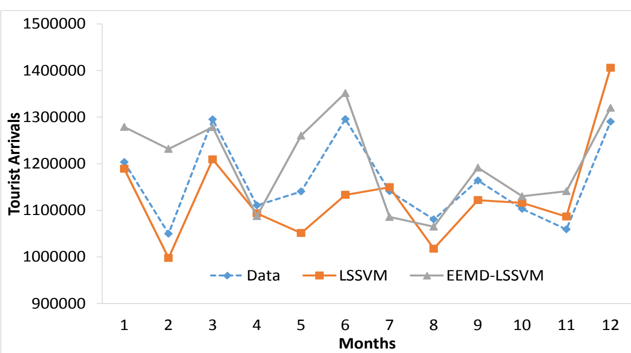

Figure 4. Comparison of LSSVM and EEMD-LSSVM.

From Table 1 and Figure 4, the conclusion can be made is that the presented EEMD-LSSVM model outperforms the single LSSVM model. The outcomes gained in this research show that due to the non-linear and non-stationary presence in monthly tourist arrivals data, EEMD-based models are more suitable for forecasting than single forecasting models. The proposed nonlinear EEMD-LSSVM, which is of effective decomposition and nonlinear prediction, can be used as a promising tool for time series forecasting with nonlinear and non-stationary.

4. Conclusion

In this study, hybrid EEMD and LSSVM model is proposed to enhance the forecasting accuracy compared to single model LSSVM to forecast tourist arrival from Singapore to Malaysia. Then, PSO algorithm is implemented to obtain optimal LSSVM model parameters. The results indicated that EEMD can efficiently improve prediction accuracy and the presented EMD-LSSVM model can significantly enhance the single LSSVM model data decomposition for tourist arrivals forecasting. Accordingly, the implementation of EEMD in hybrid forecasting can produce more stable and reliable forecasting result. Other than that, the proposed model can which are more reliable can contribute to more effective tourism planning and policy making

5. Acknowledgments

The authors gratefully recognized the financial support from MOE, UTM and GUP Grant (VOT 4F681).

6. References

1. Uysal M, El Roubi MS. Artificial Neural Networks versus multiple regression in tourism demand analysis. J Trav Res. 1999; 38:111–8.
2. Law R. Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting. Tour Manage. 2000; 21:331–40.
3. Burger C, Dohnal M, Kathrada M, Law R. A practitioner's guide to time-series methods for tourism demand forecasting: A case study of Durban, South Africa. Tour Manage. 2001; 22:403–9.
4. Law R. The impact of the Asian financial crisis on Japanese demand for travel to Hong Kong: A study of various forecasting techniques. J Trav Tour Mark. 2001; 10(2-3):47–66.
5. Tsaur SH, Chiu YC, Huang CH. Determinants of guest loyalty to international tourist hotels: A neural network approach. Tour Manag. 2002; 23:397–405.
6. Cho V. A comparison of three different approaches to tourist arrival forecasting. Tour Manage. 2003; 24:323–30.
7. Kon SC, Turner WL. Neural network forecasting of tourism demand. Tour Econ. 2005; 11:301–28.

8. Palmer A, Jose Montano JJ, Sese A. Designing an Artificial Neural Network for forecasting tourism time series. Tour Manage. 2006; 27:781–90.
9. Chen KY. Combining linear and nonlinear model in forecasting tourism demand. Expert Systems with Applications. 2011; 38:10368–76.
10. Chen CF, Lai MC, Yeh C. Forecasting tourism demand based on empirical mode decomposition and neural network. Knowledge-Based Systems. 2012; 26:281–7.
11. Oscar C, Salvador T. Forecasting tourism demand to Catalonia: Neural networks vs. time series models. Economic Modelling. 2014; 36:220–8.
12. Wang S, Yu L, Tang L, Wang S. A novel seasonal decomposition based least squares support vector regression ensemble learning approach for hydropower consumption forecasting in China. Energy. 2011; 36:6542–54.
13. Suykens JAK, Van Gestel T, De Brabanter J, De Moor J, Vandewalle J. Least squares Support Vector Machines. World Scientific Singapore; 2002.
14. Yang LS, Yang S, Zhang R, Jin H. Sparse Least Square Support Vector Machine via coupled compressive pruning. Neurocomputing. 2014; 131:77–86.
15. Wang H, Hu D. Comparison of SVM and LS-SVM for regression. IEEE International Conference Neural Networks and Brain; 2005. p. 279–83.
16. Tang L, Yu L, Wang S, Li J, Wang S. A novel hybrid ensemble learning paradigm for nuclear energy consumption forecasting. Applied Energy. 2012; 93:432–43.
17. Xie G, Wang S, Zhao Y, Lai KK. Hybrid approaches based on LSSVR model for container throughput forecasting: A comparative study. Applied Soft Computing. 2013; 13(5):2232–41.
18. Niu D, Kou B, Zhang Y, Gu Z. A short-term load forecasting model based on LS-SVM optimized by dynamic inertia weight Particle Swarm Optimization Algorithm. Advances in Neural Networks. 2009; 5552:242–50.
19. Fupeng T, Chun-Li W. Tourist number forecast of Gansu Province based on LS-SVM. Proceedings of the 2013 Fifth International Conference on Multimedia Information Networking and Security; 2013. p. 789–91.
20. Ani S, Suhartono S. Stream flow forecasting using Least Squares Support Vector Machines. Hydrological Sciences Journal. 2012; 57(7):1275–93.
21. Zhang HG, Zhang S, Yin YX. A novel improved LSSVM algorithm for a real Industrial application. Mathematical Problems in Engineering. 2014; 2014:1–7.
22. Wu Z, Huang N. Ensemble Empirical Mode Decomposition: A noise assisted data analysis method. Adv Adapt Data Anal. 2009; l(1):1–41.
23. Wang S, Yu L, Tang L, Wang S. A novel seasonal decomposition based least squares support vector regression ensemble learning approach for hydropower consumption fore casting in China. Energy. 2011; 36(11):6542–54.
24. Liu Z, Sun W, Zeng J. A new short-term load forecasting method of power system based on EEMD and SS-PSO. Neural Comput and Applic. 2014; 24:973–83.
25. Meyer PE, Lafitte F, Bontempi G. Minet: A R/Bioconductor package for inferring large transcriptional networks using mutual information. BMC Bioinformatics. 2008; 9:461.
26. Kennedy J, Eberhart RC. Particle Swarm Optimization. IEEE International Conference on Neural Networks; 1995. p. 1942–8.