Int. J. Advance Soft Compu. Appl, Vol. 7, No. 2, July 2015
ISSN 2074-2827

Empirical Mode Decomposition–Least Squares Support Vector Machine Based for Water Demand Forecasting

Ani Shabri[1] and Ruhaidah Samsudin[2]

[1]Faculty of Science, Universiti Teknologi Malaysia Johor, Malaysia
[2]Faculty of Computing, Universiti Teknologi Malaysia, Johor, Malaysia
e-mail: ani@utm.my

Abstract

Accurate forecast of water demand is one of the main problems in developing management strategy for the optimal control of water supply system. In this paper, a hybrid model which combines empirical mode decomposition (EMD) and least square support vector machine (LSSVM) model is proposed to forecast water demand. This hybrid is formulated specifically to address in modelling water demand that has high non-linear and non-stationary time series which can hardly be properly modelled and accurately forecasted by traditional statistical models. EMD is used to decompose the water demands into several intrinsic mode functions (IMFs) component and one residual component. LSSVM is built to forecast these IMFs and residual series individually, and all of these forecasting values are then aggregated to produce the final forecasted value for water demand series. To assess the effectiveness and predictability of proposed models, monthly water demand record data from Batu Pahat city in Johor of Peninsular Malaysia, has been used as a case study. Empirical results suggest that the proposed model outperforms the single LSSVM and artificial neural network (ANN) model without EMD preprocessing and EMD-ANN model. Thus, the EMD-LSSVM model is an effective method for water demand forecasting.

Keywords: *Water demand, forecasting, ANN, EMD, LSSVM.*

1 Introduction

Water is a basic need and known as the most important resource in any urban development program. Most of the decisions in urban planning and sustainable development are highly dependent on the forecasting of water demand. Accurate water demand forecasting is becoming an essential tool for designers and managers in the successful of an operation, management and modernization of water delivery systems.

Forecasting of water demand variables is a very active field of study and there is still a great deal of work to be done in this field. Previous works are done mainly by using traditional statistical models such as multiple regression [1-3] and autoregressive integrated moving average (ARIMA) models [4-6]. ARIMA model, the mostly used approach, relies on the direct identification of pattern existing in historical water demand data. The popularity of the ARIMA models is due to its statistical properties such as the well-known Box-Jenkins methodology, forecasting capabilities and richness of information regarding time-related changes. However, ARIMA models are basically linear assuming and have a limited ability to capture non-stationary and non-linearity that occurred in water demand data.

Recently, many artificial intelligence (AI) methods such as Artificial Neural Networks (ANN) [7], fuzzy methods [8], support vector machines (SVM) models [9,10] and dynamic artificial neural network models [4] have been successfully applied extensively to forecast the water demand data. Among these AI models, ANN has been frequently adopted as the modelling approach [2-4, 6, 7, 9-11]. Previous works on water demand forecasting [2, 5, 6, 10] show that the use of ANN provides very satisfactory results. More advanced, least square support vector machine (LSSVM), which is developed by Suykens and co-workers [12], is a very demanding research field nowadays. This model is a reformulation of the traditional support vector machines (SVM) that alters inequality constraints into equal conditions and employs a squared loss function, which is a differential setting relative to traditional SVM. The LSSVM can approach the non-linear system with high precision, making it a powerful tool for modelling and forecasting non-linear characteristic of time series. After several years of development, LSSVM model has been successfully used to solve forecasting problems in various fields such as wind speed [13], stream flow [14-16], water demand [17] and short term electric load [18]. Nevertheless, these AI models have their own disadvantages. For example, the performance of ANN in some specific situations is inconsistent and suffers from some weakness such as locally optimal solutions and over-fitting, which can make the forecasting precision unsatisfactory [19]. In addition, ANNs are unstable learning techniques where small changes in training data sets or parameter selection can produce a huge change in predicted outputs [20]. While other AI models such as SVM, LSSVM and ANN are sensitive to parameter selection [21].

In order to construct more efficient statistical and AI models, previous studies have proposed a number of hybrid models. For example, Pulido-Calvo and Gutierrez-Estrada [6] used hybrid methodology of ANN, fuzzy logic and genetic algorithm to forecast in Andalucía, Spain. Pulido-Calvo et al. [2] combined linear regressions and ANN to forecast water demand in irrigation districts with telemetry systems. Nasseri et al. [22] developed a hybrid model which combines Extended Kalman Filter and Genetic Programming for forecasting of water demand in Tehran. Jia and Hao [23] proposed a hybrid method based on extreme learning machine model with adaptive metrics of inputs for water demand forecasting. These papers showed that hybrid models outperform single methods. The basic idea of the above hybrid methods is to overcome the drawbacks of the single models and to generate a synergetic effect in forecasting.

Motivated by hybrid methodologies, this study attempts to apply the "decomposition-and-ensemble' principle based on empirical mode decomposition (EMD) to construct a hybrid water demand forecasting methodology. The EMD as a time-frequency resolution approach offers a new way which the stationary and nonlinear behaviour of time series can be decomposed into a series of valuable independent time resolutions. The main aim of decomposition is to simplify the difficult forecasting task by dividing it into some relatively simple but meaningful components for easy forecasting subtasks, while the goal of ensemble is to formulate a consensus forecasting result for the original data. Therefore, the decomposition can be helpful to transform non-linear and non-stationary time series to stationary time series and can be useful to improve the prediction capacity. Recent works have demonstrated the application of EMD methodology with other model outperformed individual forecasting model in many cases [24-27]. However, the applications of EMD in water demand are limited where only Ani & Samsudin [28] proposed a hybrid forecasting model based on EMD-ANN to forecast water demand. The literature review reveals that the EMD-LSSVM technique has not been used for forecasting water demand. Therefore, this study will provide important contributions to the literature of water demand forecasting.

In this paper, a hybrid water demand forecasting method based on EMD and LSSVM is proposed to further improve the forecasting accuracy. Hence, EMD is applied to decompose water demand series. Different LSSVM models are then constructed with each sub-series and the final forecasted value can be obtained through the conjunction of these models. In order to evaluate the performance of the proposed approach, the monthly water demand series from the city of Batu Pahat in Johor, Malaysia was chosen as the example to compare its prediction performance with some common individual methods such as ANN, LSSVM and EMD-ANN hybrid models.

2 Methodology

Owing to the inherent of non-linear and non-stationary of water demand series, the “decomposition-and-ensemble’ principle based on empirical mode decomposition (EMD) is introduced. In terms of this strategy, a hybrid method integrating EMD and LSSVM is proposed to enhance the quality of water demand forecast. In this methodology, the original water demand series is decomposed into a number of components depicting relatively simple but meaningful local time scales by using EMD technique. Then, the LSSVM algorithm which is a useful methodology and also a new kind of intelligent machine is applied to forecast future numerical values using the data of IMFs available in different frequencies. Finally, the forecasted values of the proposed model can be obtained by summing the forecasted value of all components respectively. This paper will demonstrate the effectiveness of the hybrid model for forecasting water demand. Before starting to use the hybrid method, it is necessary to describe the theory of the proposed approach. First, the decomposition techniques of EMD and the principle of LSSVM algorithm are presented. Then the EMD and LSSVM are combined into a developed approach.

2.1 Artificial Neural Network

Artificial neural networks (ANN) are computational models that have been extensively studied and widely ranges of applications included their great ability in modelling and forecasting in nonlinear time series since the early 1990s. Most popular neural network paradigm in water demand is multilayer perception (MLP). MLP usually consists of three layers: the first layer is the input layer where the data are introduced to the network, the second layer is the hidden layer where data are processed and the last layer is the output layer where the results of given input are produced. Fig. 1 illustrates the architecture of the proposed ANN for water demand forecasting. The general relationship between the input $(y_{t-1}, y_{t-2},..., y_{t-p})$ and output (y_t) in an ANN model can be expressed as:

$$y_t = g\left(w_0 + \sum_{j=1}^{q} w_j . f(\sum_{i=1}^{p} w_{i,j} y_{t-i} + w_{0,j}) \right) \tag{1}$$

where $w_{i,j}$ $(i = 0,1,2,..., p;\ j = 1,2,...,q)$ and $w_j\ (j = 0,1,2,...,q)$ are connection weights, p is the number of input nodes, q the number of hidden nodes and $f(\ )$ and $g(\ )$ is the transfer function of the hidden layer and the output layer.

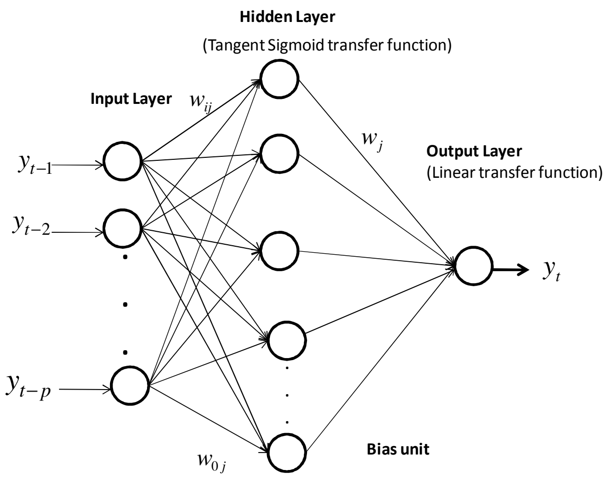

Fig. 1 Three-layered feed-forward ANN model

In order to find the optimal weight $w_{i,j}$ and w_j, learning or training processes must be employed to minimize the error. The objective function to minimize the error is the sum of the squares of the differences between the desirable output y_t and the predicted output $\hat{y}_t$, which given by

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

In this study, the training of the network is performed by the Levenberg-Marquardt (LM) algorithm. The LM algorithm is a modification of the classic Newton algorithm, it is often the fastest convergence in terms of iteration number and capable of finding better optima compared to other algorithms methods [29-31]. This optimisation technique is more powerful than the conventional gradient descent technique and has been recommended by several authors [2, 32-34].

2.2 Least Square Vector Machines Model

LSSVM is a new version of SVM which applied the linear system instead of solving a quadratic programming problem [12]. The basic principle of LSSVM is as follows: Suppose (X_t, Y_t) for t = 1 to n is a given set of data where $X_t = (y_{t-1}, y_{t-2}, \ldots, y_{t-p})$ is the input vector with p multiple variables and y_t is the corresponding output data at time t. By a nonlinear function φ, the data are mapped from the original feature space to a higher dimensional feature space. Thus, to approximate it in a linear way is as follows:

$$y_t = w^T \varphi(X_t) + b \quad (2)$$

LSSVM introduces a least square version to SVM regression by formulating the regression problem as

$$\min R(w,e) = \frac{1}{2} w^T w + \frac{\gamma}{2} \sum_{t=1}^{n} e_t^2 \quad (3)$$

subject to $\quad y_t = w^T \varphi(X_t) + b + e_t\,,\ \ t = 1, 2, ..., n$

where e_i is the error variable at time t and γ is a regulation constant. In order to solve the optimization problem, the Lagrange function is formulated as follow

$$L(w,e,\alpha) = \frac{1}{2} w^T w + \frac{\gamma}{2} \sum_{t=1}^{n} e_t^2 - \sum_{t=1}^{n} \alpha_t (w^T \varphi(X_t) + b + e_i - y_t) \quad (4)$$

where α_t is the Lagrange multipliers. By using partially differentiating $L(w,e,\alpha)$ in (4) with the variable w, b, e_i and α_i which is shown as follows

$$\begin{aligned}
&\frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{t=1}^{n} \alpha_t \varphi(X_t) \\
&\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{t=1}^{n} \alpha_t = 0 \\
&\frac{\partial L}{\partial e_t} = 0 \rightarrow \alpha_t = \gamma e_t\,,\ t = 1, 2, ..., n \\
&\frac{\partial L}{\partial \alpha_i} = 0 \rightarrow w^T \varphi(X_t) + b + e_t - y_t\,,\ t = 1, 2, ..., n
\end{aligned} \quad (5)$$

An alternative formulation of (5) can be transformed into the following linear solution:

$$\begin{bmatrix} 0 & \mathbf{1}^T \\ \mathbf{1} & K(X_t, X_l) + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}$$

where $K(X_t, X_l) = \langle \varphi(X_t), \varphi(X_l) \rangle$ is the kernel function. By solving the upper linear system, the final solution of LSSVM model for non-linear function can be written as:

$$y_t = \sum_{t=1}^{n} \alpha_t K(X, X_t) + b \tag{6}$$

where α_i, b are the solution to the linear system. There are several types of kernel function $K(X, X_t)$ such as sigmoid, polynomial and radial basis function (RBF). The most popular kernel function is RBF and is expressed as $K(X, X_t) = \exp(-\|X - X_t\| / 2\sigma^2)$ with a width σ and the polynomial kernel $K(X, X_t) = (a_1 X X_t + a_2)^d$ with an order d and constants a_1 and a_2.

2.3 Empirical Mode Decomposition

EMD recently pioneered by Huang et al. [35], is a signal analysis method which can deal with non-linear and non-stationary data. The main idea of EMD method is to decompose the original data into a sum of intrinsic mode function (IMF) components with individual intrinsic time scale properties. IMFs have to satisfy the following two conditions (a) in the entire data set, the number of extreme values (maxima plus minima) and the number of zero-crossings must either be equal or differ by at most one; and (b) at any point, the mean value of the envelope, constructed by the local maxima and minima, is zero at any point. The detailed decomposition process of EMD is presented by Huang et al. [35]. The algorithm of EMD is described as follows:

(i) Identify all the local extremes including maxima and minima values in time series data $y(t)$.

(ii) Obtain the upper envelope $y_U(t)$ and the lower envelope $y_L(t)$.

(iii) Calculate the mean value $M_1(t) = (y_L(t) + y_U(t))/2$.

(iv) Evaluate the difference between the original time series $y(t)$ and the mean time series $M_1(t)$. The first IMF $h_1(t)$ is defined as $h_1(t) = y(t) - M_1(t)$.

(v) Check whether $h_1(t)$ satisfies the two conditions of an IMF property. If they are not satisfied, we repeat steps (i) – (iii) of the decomposition procedure to eventually find the first IMF.

(vi) After we obtained the first IMF, a repetition of the above steps are necessary to find the second IMF, until we reach the final time series $e(t)$ that satisfies one of the termination criteria suggesting to stop the decomposition procedure.

By using the above algorithm, the original time series $y(t)$ can be reconstructed by summing up all of the IMF components and one residual component as Eq. (7), following expresses.

$$y(t) = \sum_{i=1}^{m} h_i(t) + e_m(t) \tag{7}$$

where m is the number of IMs, $h_i(t)$ represents IMFs and $e_m(t)$ is the final residual, which is a constant or a trend. The EMD techniques provide a multi-scale analysis of the signal as a sum of orthogonal signals corresponding to different time scales and also be-taken as a filter of high pass, band pass or low pass.

2.4 The Architecture of Hybrid Intelligent Forecasting Model

Fig.2 describes the process of EMD-LSSVM forecasting method. As it can be seen from Fig.2, the EMD-LSSVM forecasting can be described by the following steps:

(i) Decompose the time series $y(t)$, t =1, 2, … , n is into m-IMFs $h_1(t)$, $h_2(t)$,..., $h_m(t)$ and one residual $e_m(t)$ by EMD.

(ii) Use the LSSVM model to model each of m-IMFs and the residue $e_m(t)$. The LSSVM models are then applied to forecast the future one-day values of these IMFs and the residual. The partial autocorrelation function (PACF) is used as an input data; pre-processing tool for LSSVM model.

(iii) Sum up of all the prediction results of IMF components and one residual component.

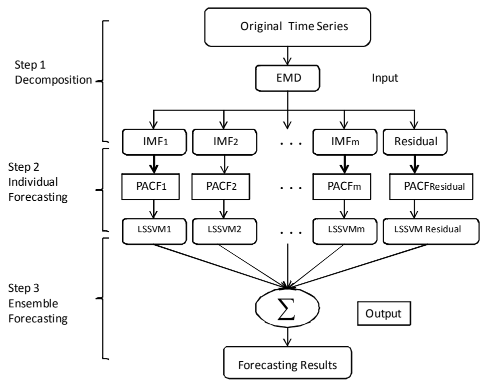

Fig. 2 The proposed EMD-LSSVM forecasting model for water demand data

The same EMD-based methodology steps are also fed into ANN in order to build the hybrid linear water demand forecasting model, namely, the EMD-ANN models.

3 Experimental Analysis

Time series data of monthly water demand data from Batu Pahat city in Johor, Malaysia obtained from Syarikat Air Johor (SAJ) was used in this study. The sample data covering the period dated from January 1995 to December 2011 with a total of 204 observations are used as shown in Fig. 3. The data dated from January 1995 to December 2010 (192 observations) are used as training dataset and the remaining data from January 2011 to December 2011 (12 observations) are chosen as testing dataset evaluate performance of prediction.

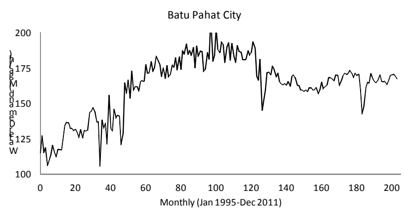

Fig. 3 The monthly water demand series from Batu Pahat city from Jan. 1995 to Dec. 2011

For comparison of the forecasting performance of the proposed model, three widely used performance indexes, root mean square error (RMSE), mean absolute error (MAE) and coefficient of correlation (R) are applied. RMSE, MAE and R are defined in following equations:

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

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|\ y_t - \hat{y}_t\ | \text{ and } R = \frac{\sum_{t=1}^{n}(y_t - \bar{y})(\hat{y}_t - \bar{y}')}{\sqrt{\sum_{t=1}^{n}(y_t - \bar{y})^2 \sum_{t=1}^{n}(\hat{y}_t - \bar{y}')^2}}$$

where y_t is the actual data, $\hat{y}_t$ is the forecasted value of period *t*, $\bar{y}$ and $\bar{y}'$ are the means of y_t and $\hat{y}_t$, respectively. Obviously the smaller the values of RMSE and MAE, and R closer to 1 indicate high efficiency of the model. *R* has been widely used for model evaluation, though they are oversensitive to high extreme values (outliers) and insensitive to additive and proportional differences between model predictions and measured data [36-38].

3 EMD-Based Methodology to the Data

In the EMD-based methodology, the first step is to apply EMD algorithm to decompose the original water demand data series into several IMF components and one residual. The EMD algorithm is implemented via R software package using EMD library. Fig. 4 shows the decomposition results for the water demand series is decomposed into seven IMFs and one residual, which exhibits a stable and regular variation. Clearly, all IMF components are listed in order from the highest frequency to the lowest frequency, and the last one is the residual.

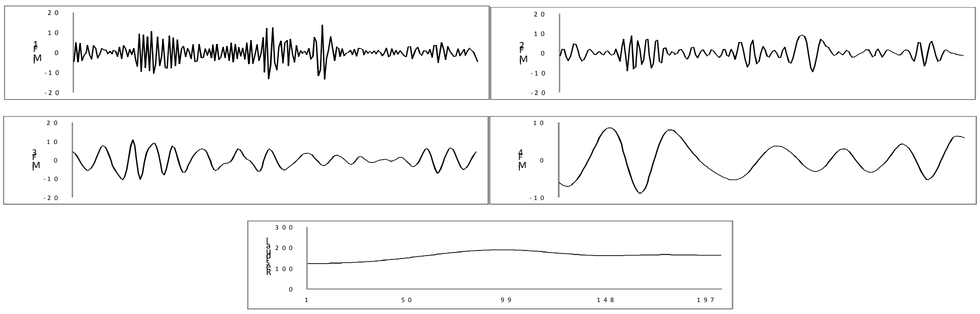

Fig. 4 The IMFs and one residual for water demand series via EMD

After the time series are decomposed, each component of IMFs and the residual are then used to build the LSSVM and ANN forecasting models. One of the most important steps in the model development process of LSSVM and ANN model is the determination of significant input variables. In the modelling of the single LSSVM and ANN models like the EMD-LSSVM and EMD-ANN models, the PACF graph which is simply the plot of PACF against the lag length is used to determine the input variables as studied by Hu et al. [39]. The PACF of original data, IMFs and one residual component are shown in Fig. 5. Through observing Fig. 5 with the output variable y_t, it is obvious that the input variables for LSSVM and ANN modelling are as shown in Table 1.

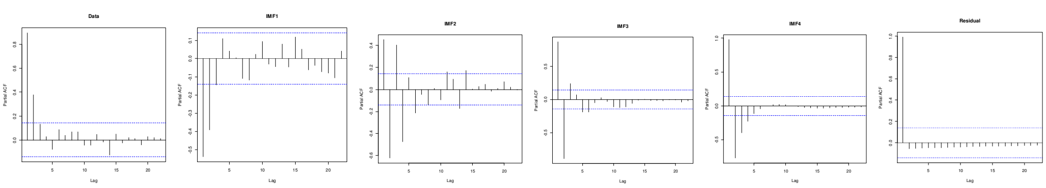

Fig. 5 PACF of original data, IMFs and one residual

Table 1: Input variables for LSSVM and ANN modelling based on PACF.

Data	Input Variables
Original	y_{t-1}, y_{t-2}
IMF1	$y_{t-1}, y_{t-2}, y_{t-3}$
IMF2	$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-6}, y_{t-8}, y_{t-11}, y_{t-13}, y_{t-14}$
IMF3	$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-5}, y_{t-6}$
IMF4	$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}$
Residual	y_{t-1}

In this study, in each LSSVM model, the most popular function kernel, Gaussian RBF is chosen and the kernel parameter σ^2 and γ were determined beforehand. Currently, many approaches have been applied in parameter optimization of LSSVM such as grid search, cross validation, genetic algorithm, particle swarm optimization, etc. In order to obtain the optimal model parameters γ and σ^2 of the LSSVM, the most used method, cross-validation grid search method were employed. To overcome parameter sensitiveness, 10-fold cross validation on the training set was performed to predict the prediction error. The best fit model structure for each model is determined according to the criteria of the performance evaluation. The LSSVMlab software package toolbox developed by Suykens et al. [14] for MATLAB platform is used to build the LSSVM models.

For ANN model, the training and testing data were normalized within the range of [-1, 1]. The hidden nodes use the hyperbolic tangent sigmoid transfer function and the output layer uses the linear function because the prediction performance is the best when these transfer functions are used. There is no greed standard on how to determine the optimal number of hidden nodes in the hidden layer. Berry and Linoff [40] claimed that the number of hidden nodes should never be more than 2I, where I is the number of inputs. Hect-Nielsen [41] claimed that the number of hidden neuron is equal 2I +1. In the present study, the number of hidden nodes was progressively increased from 1 to 2I +1.

A program code including the wavelet toolbox was written in MATLAB language for the development of the ANN model. The optimal complexity of the ANN model, that is the number of input and hidden nodes, was determined by a trial-and-error approach.

In order to verify the forecasting capability of the proposed EMD-based model with two popular single forecasting approaches, including single LSSVM and ANN models are employed for comparison to forecast water demand. Table 2 summarizes the prediction performance of difference prediction models on the water demand data mentioned above. It is obvious that the proposed EMD-LSSVM model yields better results than EMD-ANN model with the lowest

RMSE and MAE, and the highest R. When comparing single forecasting models, the ANN model mostly ranks the last, while LSSVM model produce better results in all cases. It is worth noting that RMSE, MAE and R results of the three hybrid ensemble learning models (i.e. EMD-LSSVM and EMD-ANN) are significantly better than those of the two single models (i.e. ANN and LSSVM). It shows that the decomposition strategy does effectively improve prediction performance.

The actual water demand data and forecasted values for the ANN, LSSVM, EMD-ANN and EMD-LSSVM models are illustrated in Fig. 6. From Fig. 6, it can be seen that the proposed hybrid ensemble learning models provide good forecasting results. The forecasted values of the proposed model are very close to the actual values than those obtained from other models.

Table 2: Performance of the four forecasting methods

Statistics	ANN	LSSVM	EMD-ANN	EMD-LSSVM
MAE	2.1625	1.8116	1.8688	**1.6485**
RMSE	2.5610	2.1012	2.7130	**2.0021**
R	0.3533	0.4391	0.5439	**0.6819**

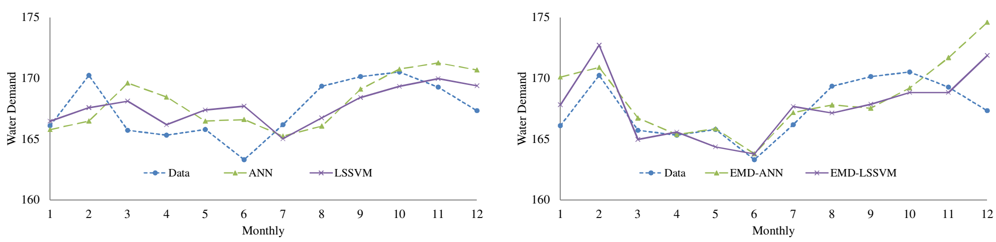

Fig. 6 Forecasts of water demand from Batu Pahat city in Johor of Malaysia.

4 Conclusion

The development of accurate forecasting methods is an important topic of continuing interest and research. However, it is widely known that water demand series often is highly non-stationary and non-linearity. Therefore, it may have poor prediction performance from applying traditional statistical models. This study attempts to apply a hybrid forecasting method which is an integration of empirical model decomposition EMD and LSSVM model to predict water demand series from Batu Pahat city in Johor Malaysia. This methodology first decomposes the original water demand series into several intrinsic model function (IMFs) components and one residual component by EMD method. This can reduce the non-stationary of the water demand series and enhance the prediction accuracy. Then, the IMFs components and the residual components are forecasted respectively using LSSVM model whose input variables are selected by using partial autocorrelation function (PACF). The final forecasted result for water demand series is produced by aggregating all the forecasted results. Empirical results show that the proposed EMD-LSSVM model outperforms the EMD-ANN as well as the LSSVM and ANN models without time series decomposition. The performance of the EMD-LSSVM and EMD-ANN are generally better than the single methods. This indicates that the "decomposition and ensemble" strategy can effectively improve the prediction performance of water demand series. Thus, it can be concluded that the proposed EMD-LSSVM model may be an effective tool as a very promising methodology for complex problems such as water demand series forecasting with highly non-stationary and non-linearity.

ACKNOWLEDGEMENTS.
The authors wish to thank the Ministry of Higher Education (MOHE) for the financial support of the work under the Fundamental Research Grant Scheme (FRGS No. 4F275). Also the authors gratefully acknowledge the critical comments and corrections of the anonymous reviews; their comments significantly improved the original manuscript.

References

[1] Haque, M.M., Rahman,A., Hagare, D. and G. Kibria. 2013. Principal Component Regression Analysis in Water Demand Forecasting: An Application to the Blue Mountains, NSW, Australia. *Journal of Hydrology and Environment Research*, Vol 1, No.1, 49-59.

[2] Pulido-Calvo, I., Montesinos, P., Roldan, J., and F. Ruiz-Navarro. 2007. Linear Regression and Neural Approaches to Water Demand Forecasting in Irrigation Districts with Telemetry Systems. *Biosystem Engineering*, 97, 283-293.

[3] Bougadis, J., Adamowski, K., and R. Diduch. 2005. Short-term Municipal Water Demand Forecasting. *Hydrological Processes*, Vol. 19, No.1, 137-148.

[4] Ghiassi, M., Zimbra, D.K., and H. Saidane. 2008. Urban Water Demand Forecasting with a Dynamic Artificial Neural Network Model. *Journal of Water Resources Planning and Management*, Vol.134, No.2, 138-146.

[5] Lipae, J.L., and E.P. Deligero. 2012. On forecasting water consumption in Davao city using Autoregressive Integrated Moving Average (ARIMA) Models and the Multilayer Perceptron Neural Network (MLPNN) Processes. *International Journal of Humanities and Applied Sciences*, Vol.1, No.4, 2277-4386.

[6] Pulido-Calvo, I., and J.C. Gutierrez-Estrada. 2009. Improved irrigation water demand forecasting using a soft-computing hybrid model. *Biosystem Engineering*, 102, 202-218.
[7] Liu, J., Savenije, H.H.G., and J. Xu. 2003. Forecast of Water Demand in Weinan City in China Using WDF-ANN model. *Physics and Chemistry of the Earth*,. 28, 219-224.
[8] Tabesh,M., and M. Dini. 2009. Fuzzy and Neuro-Fuzzy Models for Short-Term Water Demand Forecasting in Tehran. *Iranian Journal of Science & Technology*, Transaction B, Engineering, 33, 61-77.
[9] Msiza, I.S., Nelwamondo, F.V., T. Marwala. 2008. Water Demand Prediction Using Artificial Neural Networks and Support Vector Regression. *Journal of Computer*,. 3(11), 1-8.
[10] Herrea, M., Torgo, L., Izquierdo, J., and R. Perez-Garcia. 2010. Predictive Models For Forecasting Hourly Urban Water Demand. *Journal of Hydrology*, 387, 141-150.
[11] Ajbar A., and E.M. Ali. 2013. .Prediction of municipal water production in touristic Mecca City in Saudi Arabia using neural network. *Journal of King Saud University-Engineering Sciences*, 1-9.
[12] Suykens, J.A.K, Van Gestel, De Brabanter, T., De Moor, J., and J. Vandewalle. 2002. .*Least squares support vector machines*. World Scientific Singapore,
[13] Zhao, X.H., Wang, G., Zhao, K.K., and D.J. Tan, 2009.*On-line least squares support vector machine algorithm in gas prediction.* Min. Sci. Technol. 19, 194–198.
[14] Ismail, S., Samsudin, R., and A. Shabri, 2010. *River flow forecasting: a hybrid model of self organizing maps and least square support vector machine*. Hydrol. Earth Syst. Sci. Discuss, 7, 8179–8212.
[15] Samsudin, R., Saad, P. and A. Shabri, 2011. *River flow time series using least squares support vector machines*. Hydrol. Earth Syst. Sci. 15, 1835-1852.
[16] Ani, S., Suhartono, 2012. Streamflow forecasting using least-squares support vector machines. *Hydrological Sciences Journal*, 57(7), 1275-1293.
[17] Xuemei L., Lixing D. and L. Jinhu, 2010. Agriculture Irrigation Water Demand Forecasting Based on Rough Set Theory and Weighted LS-SVM. Second International Conference on Communication Systems, *Networks and Applications, IEEE*, 371-374.
[18] Wu, Q., 2011. Hybrid model based on wavelet support vector machine and modified genetic algorithm penalizing Gaussian noises for power load forecasts. *Expert Syst. Appl..* 38, 379–385.
[19] Ming, Y., 2009. Application of Particle Swarm Optimization and RBF Neural Network in Fault Diagnosis of Analogue Circuits. *IEEE-Third International Symposium on Intelligent Information Technology Application*, 176-178.

[20] Naftaly, U. , Intrator, N., and D. Horn, 1997. Optimal ensemble averaging of neural networks*, Network. Comput. Neural Syst.*, 8, 283–296.

[21] Wang, S., Yu, L., Tang, L., Wang, S., 2011. A novel seasonal decomposition based least squares support vector regression ensemble learning approach for hydropower consumption forecasting in China. *Energy*, 36, 6542-6554.

[22] Nasseri, M., Moieini, A., and M. Tabesh, 2011. Forecasting Monthly Urban Water demand Using Extended Kalman Filter and Genetic Programming. *Expert Systems with Applications*, 38, 7387-7395.

[23] Jia, J., and S. Hao, Water Demand Forecasting Based On Adaptive Extreme Learning Machine. 2013. *International Conference on Artificial Intelligence and Software Engineering*, 42-45.

[24] Tang, L., Yu, L., Wang, S., Li, J., and S. Wang, 2012. A novel hybrid ensemble learning paradigm for nuclear energy consumption forecasting. *Applied Energy*, 93, 432-443.

[25] Lin, C.S., Chiu, S.H., and T.Y. Lin, 2012. Empirical mode decomposition–based least squares support vector regression for foreign exchange rate forecasting. *Economic Modelling*, Vol. 29 No. 6, 2583-2590.

[26] Wei, Y., and M.C. Chen, 2012. Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks. *Transportation Research Part C*, 21, 148-162.

[27] Chen, C.F., Lai, M.C., and C.C. Yeh, 2012. Forecasting tourism demand based on empirical mode decomposition and neural network. *Knowledge-Based Systems*, 26, 281-287.

[28] Ani, S., Samsudin, R., 2014. A new approach for water demand forecasting based on empirical mode decomposition. *IEEE Software Engineering Conference (MySEC), 8th Malaysian*, 284-288.

[29] Hagan M.T., and M.B. Menhaj, 1994. Training Feedforward Networks with the Marquardt Algorithm. *IEEE Transactions On Neural Networks*, 5(6): 989-993.

[30] Zayani R., Bouallegue R., and D. Roviras, 2008. Levenberg-Marquardt Learning Network dor Adaptive Predistortion for Time-Varying HPA with Memory in OFDM Systems. 16th *European Signal Processing Conference* (EUSIPCO 2008), Lausanne, Switzerland, August, 25-29.

[31] Beale, M.H., Hagan, M.T., H.B. 2011. Demuth, Neural Network Toolbox Getting Started Guide R2011b,

[32] Anctl, F., A. Rat, 2005. Evaluation of Neural Network Stream Flow Forecasting on 47 Watersheds. *Journal of Hydrologic Engineering*, 10, 85-88.

[33] Adamowski, J., and C. Karapataki, 2010. Comparison of multivariate regression and artificial neural networks for peak urban water-demand forecasting: evaluation of different ann learning algorithms. *J. Hydrol. Eng.*, 15(10), 729–743.

[34] Piotrowski A.P., and J.J. Napiorkowski, 2011. Optimizing neural networks for river flow forecasting-Evolutionary Computation methods versus the Levenberg-Marquardt approach. *Journal of Hydrology*, 407, 12-27.

[35] Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N., Tung, C.C., and H.H.Liu, H.H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc*. London, 454 903-995.

[36] Legates, D.R., McCabe, G.J. 1999. Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Res. Res*. 35, 233–241.

[37] Wang, W.C., Chau, K.W., Cheng, C.T., Qiu, L. 2009. A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *Journal of Hydrology*. 374, 294–306.

[38] Wang, W., Chau, K., Qiu, L., Chen, Y. 2015. Improving forecasting accuracy of medium and long-term runoff using artificial neural network based on EEMD decomposition. *Environmental Research*. 139, 46-54.

[39] Hu, J,, Wang, J., Zeng, G., 2013. A hybrid forecasting approach applied to wind speed time series. *Renewable Energy*, 60, 185-194.

[40] Berry, M.J., and G. Linoff, 2007. *Data mining techniques*. John Wiley, Hoboken, N.J.

[41] Hecht-Nielsen, R., 1990. *Neurocomputing*, Reading, MA: Addison-Wesley.