MJFAS MALAYSIAN JOURNAL OF FUNDAMENTAL AND APPLIED SCIENCES
PRINT ISSN: 2289-5981 | ONLINE ISSN: 2289-599X

RESEARCH ARTICLE

Comparative performance of support vector regressions for accurate streamflow predictions

Noraini Ibrahim*, Norhaiza Ahmad

Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia

* *Corresponding author: norainiibrahim89@gmail.com*

Article history
Received 24 October 2017
Accepted 8 November 2017

Abstract

Obtaining accurate streamflow predictions can be challenging due to the inherent variabilities and complex nonlinear nature in streamflow generation processes. Support vector regression model is an effective forecasting tool to forecast streamflow as it is able to capture the nonlinearity in the data and attain the global optimum parameters in the forecasted model. However, the efficiency of SVR might be hindered by noise that typically exists in any hydrological time series data through random influences and inaccuracies in recording. Thus, this condition could compromise the quality of input data into SVR. In this study, we investigate the effectiveness of forecasting monthly streamflow data using different settings of SVR in two ways. First, we use different variations of wavelet denoising technique using different selections of wavelet decomposition levels and mother wavelets in order to preserve information and reduce distortion of the original time series. For this purpose, we measured the impact of six different wavelets on SVR namely Daubechies of type *db3, db4, db5, db6* and *db7* with two different levels of decomposition which are level 3 and level 4. There is more information that may contribute to better performance of the model when the decomposition level is increase. Then, the data are applied using radial basis function (RBF) by performing K-fold cross-validation to obtain the optimal parameter for kernel function in forecasting streamflow. We illustrate the methods using the monthly streamflow data observed at Segamat River in the state of Johor. The results demonstrated that SVR based wavelet denoising for 1-month lead time streamflow forecasting of type *db5* with level 3 give better results using Gaussian (RBF) kernel function based on K-fold cross-validation compared to regular SVR. This implies that reduced variance in the denoising procedure and obtain optimal parameter in kernel function may improve forecasting accuracy.

***Keywords*:** Support vector regression; kernel functions; wavelet denoising; mother wavelets; wavelet decomposition levels.

INTRODUCTION

Streamflow forecasting is important towards an effective operation of a water resource system. It is a fundamental and critical component of global and regional hydrological cycles (Makkeasorn *et al*., 2008). It is associated with human water supply, the agricultural and industrial sectors and natural disasters including droughts and floods (Liu *et al*., 2014). The streamflow time series always tend to be nonlinear, time-varying and indeterminate since it is influenced by both known factors, including precipitation, evaporation, temperature and many unknown factors. Hence, it is very difficult to make exact prediction of the streamflow. Over the last few decades, streamflow forecasting become more important because of the fluctuations of global climate change that causes extreme drought and flood events (Adamowski *et al*., 2010). A variety of methods have been developed and used for streamflow forecasting including traditional statistical models such as multiple linear regression (MLR) and auto regressive integrated moving average (ARIMA) models (Mckerchar *et al*., 1974), and machine learning techniques such as artificial neural networks (ANNs) and support vector machine (SVM) (Kim *et al*., 2001; Sivapragasam *et al*., 2001; Kisi *et al*., 2011).

To date, SVM for regression or known as SVR, which is proposed by Cortes and Vapnik (1995), has attracted a great deal of interest as an effective forecasting tools and is considered as an alternative approach of ANNs. This forecasting tool uses machine learning theory to maximize predictive accuracy while automatically avoiding over-fitting to the data (Vapnik, 1995). SVR is based on the structural risk minimization (SRM) principle rather than the empirical risk minimization (ERM) principle and basically involves solving a quadratic programming problem, thus can obtain the global optimum result of the original problem (Yu *et al*., 2006; Wu *et al*., 2012; Guo *et al*., 2011). SRM is an inductive principle for model selection used for learning from finite training data sets.

SVR was proposed by Vapnik *et al.* (1997) after SVM for classification problem. It becomes successfully applied in the field of hydrology. Liong and Sivapragasm (2002) employed the SVR for flood stage forecasting; Xiong and Li (2005) used SVR to forecast the sediment-carrying capacity; Bray and Han (2004) analyzed the best model based on the optimized parameters of SVR to forecast streamflow accurately; Lin *et al.* (2006) demonstrated the application of SVR to forecast monthly river flow discharges in the Manwan Hydropower Scheme; Wang *et al.* (2009) presented that SVR performed better than the ANN and ARIMA for forecasting monthly discharge time series; Yu and Xia (2008) proposed a runoff prediction model based on SVR and chaos theory; Mohsen *et al.* (2009) concluded that SVR is better than ANN in some case in runoff modeling; Lin *et al.* (2006) used SVR to predict long-term discharge.

Although SVR is useful in forecasting hydrological time series, it has llimitation on the highly non-stationary data change over a range scales (Liu *et al*., 2006; Cannas *et al*., 2006; Adamowski *et al*., 2011). Wavelet transform is one of the path to deal with the non-stationary behavior in hydrological signals. The signals is presented as a time-frequency at different scales in the time domain and the time series data can be decomposed into various period while considering the physical structure of the data (Daubechies, 1990). The application of wavelet transforms for analyzing variabilities, periodicities and trends in time series has been widely used in recent years (Smith *et al*., 1998; Lu, 2002; Chou *et al*., 2002; Xingang *et al*., 2003; Coulibaly *et al*., 2004; Partal *et al*., 2006). Discrete wavelet transform (DWT) is applied by Smith *et al.* (1998) for assessing streamflow variability. Coulibaly and Burn (2004) employed DWT to characterize variability in annual Canadian streamflows. Partal and Kucuk (2006) performed DWT for identifying the possible trends in annual precipitation and concluded that the trend structure of data is well clearly explained based on analysis on DWT components of the precipitation. Milne *et al.* (2009) applied a wavelet packet transform to identify temporal variation of river water solutes.

The better forecasts can be possible to generate by combining the strengths of wavelet transform and SVR (or other data-driven models) (Kisi, 2008; Kisi, 2009; Nourani *et al*., 2009; Remesan *et al*., 2009; Pramanik *et al*., 2010; Shiri *et al*., 2010; Li, 2011; Tiwari *et al*., 2010; Kisi *et al*., 2012; Rasouli *et al*., 2012; Adamowski, 2013; Sang, 2013). The combination of wavelet and SVR has been shown that it can give better prediction compared to regular SVR model in hydrological forecasting (Kisi *et al*., 2011). The coupled model wavelet based data driven models including SVR has been demonstrated that it can yields more accurate forecasts than single data driven model (Kalteh, 2013).

However, there are some issues of the wavelet-SVR model including the choice of an appropriate wavelet and the decomposition level. The choice of choosing an appropriate mother wavelet is the most important part in wavelet analysis (Sang, 2013; Kalteh, 2013). In this paper, Daubechies (db) wavelets are considered because of its sensitivity in analyzing nonlinear time series (Guo *et al*., 2011); Brito *et al*., 1998). Besides, Daubechies wavelets are adopted because of consisting of certain characteristics that are vital for localizing events in time-dependent signals (Liu *et al*., 2014; Daubechies, 1990; Papivanov *et al*., 2002). They are also widely used as a mother wavelets in hydrological time series using discrete wavelet transform (Liu *et al*., 2014). Besides, it is important to select a suitable decomposition level for wavelet based SVR model since it can affects the accuracy of the model. When the decomposition level is increased, more detailed information of series at larger scales can be seen, but more input neutrals may reduce the computing efficiency which can decrease the stability of the model.

Therefore, this study aims to evaluate the performance discrepancies resulting from different mother wavelets and decomposition levels in wavelet based SVR model using *k-fold* cross-validation to obtain the optimal parameter in forecasting streamflow of Segamat River. Before starting training the SVR model, it is necessary to reduce the noise in streamflow time series since the hydrological data mostly consists of noise that can affects the forecasting accuracy of the streamflow. Hence, a denoise method based on wavelet is adopted which is called wavelet denoising. The SVR based wavelet denoising model is applied in one-step ahead forecasting for monthly streamflow and compare the results with those from the regular SVR model.

THEORITICAL BACKGROUND

It is necessary to reduce the noise in time series streamflow since it usually consists of noise which can influence the accuracy of the prediction. The wavelet is a powerful tool to process the time series streamflow signal due to the complexity of the streamflow process. The wavelet decomposition and reconstruction theory serve an effective denoise method which is known as wavelet denoising. The framework of the SVR based wavelet denoising is given in Fig. 1 while the details are shown as the follow of this section.

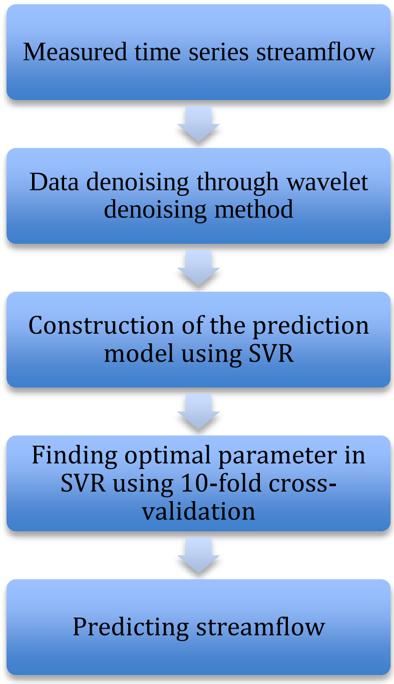

Fig. 1 The framework of the SVR based wavelet denoising

Support Vector Regression

In recent years, there has been a lot of interest in studying support vector machines (SVM) in the field of machine learning. SVR is one part of SVM to solve regression problems. The input vector for SVR is mapped to a high-dimensional feature space using a nonlinear mapping function (Wu *et al*., 2012). Performing linear regression in the feature space can solve the nonlinear problems. The nonlinear mapping of into a feature space by a nonlinear function $\phi(x)$ is given by

$$f(w,b) = w.f(\mathrm{x}) + b \tag{1}$$

The nonlinear regression problem can be expressed as

$$\begin{aligned}
&\min_{w,b,\xi,\xi^*} && \frac{1}{2}w^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \\
&\text{subject to} && y_i - (w.\phi(\mathrm{x}_i) + b \le \varepsilon + \xi_i \\
& && (w.\phi(\mathrm{x}_i) + b - y_i \le \varepsilon + \xi_i^* \\
& && \xi_i, \xi_i^* \ge 0, i = 1,2,\dots,n
\end{aligned} \tag{2}$$

The dual form of the nonlinear SVR is given by

$$\begin{aligned}
&\min_{a_i,a_i^*} \quad \frac{1}{2}\sum_{i,j=1}^{n}(a_i - a_i^*)(a_j - a_j^*)\langle\phi(\mathrm{x}_i).\phi(\mathrm{x}_j)\rangle + \varepsilon\sum_{i,j=1}^{n}(a_i \\
&\qquad\qquad + a_i^*)(a_i - a_i^*) \\
&\text{subject to} \sum_{i=1}^{n} = 0 \\
&\qquad 0 \le a_i \le C, i = 1,2,\dots,n \\
&\qquad 0 \le a_i^* \le C, i = 1,2,\dots,n
\end{aligned} \tag{3}$$

Functions that meet Mercer's condition (Guo *et al*., 2011; Brito *et al*., 1998) can be proven to correspond to dot products in a feature space. Hence, any functions that satisfy Mercer's theorem can be used as a kernel which is shown as follows.

$$K(\mathrm{x}_{i,}\mathrm{x}_j) = \langle f(\mathrm{x}_i).f(\mathrm{x}_j)\rangle \tag{4}$$

Therefore, the kernel function allows the decision function of nonlinear SVR to be expressed as follows.

$$f(\mathrm{x}_i) = \sum_{i,j=1}^{n} (a_k - a_k^*) K(\mathrm{x}_i, \mathrm{x}_k) + b \quad (k = 1,2,\dots,l) \qquad (5)$$

where x_k denoted as the support vector and l is the number of support vector.

Wavelet denoise method

Wavelet analysis is a multi-resolution analysis that represented in time and frequency of time series data. The wavelet transform decomposes time series data into different components at different resolution levels using the mother wavelets (Tiwari et al., 2010). Wavelet denoising is useful to extract the high frequency from the signal specifically in streamflow time series that is always has high frequency signals. This high frequency of signals is called as a noise in streamflow time series. The wavelet function $\psi(t)$ which is called the mother wavelet can be defined as $\int_{-\infty}^{\infty} \psi(t)dt = 0$ and the $\psi_{a,b}(t)$ can be obtained by expanding $\psi(t)$ as follows [41]:

$$\psi_{a,b}(t) = |a|^{\frac{1}{2}} \psi\left(\frac{t-b}{a}\right) \; (b \in R, a \in R, a \neq 0 \qquad (6)$$

where $\psi_{a,b}(t)$ denotes the successive wavelet, a represent the scale or frequency factor, b is a time factor and R is the domain of real numbers (Kisi *et al*., 2012). The successive wavelet transform of f(t) for any time series $f(t) \in L^2(R)$ can be defined as (Guo *et al*., 2011):

$$W_f(a,b) = a^{\frac{1}{2}} \int_{-\infty}^{\infty} \psi\left(\frac{t-b}{a}\right) dt \qquad (7)$$

where $\overline{\psi(t)}$ is the complex conjugate functions of $\psi(t)$ and the wavelet transform is the decomposition of $f(t)$ using different resolution level. Using the successive wavelet transform $W_f(a,b)$ the original time series $f(t)$ can be expressed as in Eq. (8) which is through wavelet reconstruction.

$$f(t)\left(\int_{-\infty}^{\infty} \frac{|\hat{\psi}(\varphi)|^2}{\varphi} d\varphi\right) \int_{-\infty}^{\infty} \frac{1}{a^2} W_f(a,b)\psi_{a,b}(t)da db \qquad (8)$$

where $\hat{\psi}(\varphi)$ denoted the Fourier transform of $\psi(t)$. The wavelet denoise procedure consists of three parts including wavelet decomposition, threshold processing and wavelet reconstruction which presented as follows:

- Wavelet decomposition. The proper wavelet function and suitable decomposition level, N are selected. The one low-frequency wavelet coefficients and N high-frequency wavelet coefficients series can be obtained after calculate the wavelet transform of the original time series by applying Eq. (7).
- Threshold processing. In order to determine insignificant high-frequency wavelet coefficients in wavelet transform coding of signals, a threshold T is used. The threshold is set as a standard threshold in this approach.
- Wavelet reconstruction. The denoised time series can be obtained through wavelet reconstruction by employing Eq. (8) after acquire the low-frequency wavelet coefficients series and N high-frequency wavelet coefficients after threshold processing.

PERFORMANCE CRITERIA

One of the statistical method that can be used to evaluate the performance of machine learning based prediction algorithms is cross-validation (Cheng *et al*., 2012). In this paper, cross-validation is used to select a predictor and model selection for predicting streamflow of Johor River. Cross-validation is the process on dividing the datasets into two segments, which are for training and validation. The K-fold cross-validation is applied in this paper since it can maintain the accuracy of the estimation and reduce the computation time. The performance of the prediction algorithms can be estimated by the root mean squared error of cross-validation (RMSECV).

The appropriate model can be obtained by comparing RMSECV using four types of kernel function which are linear, polynomial, RBF and sigmoid kernel based on 10 types of denoised predictors. We acquire 10-fold cross-validation which means the data were divided into 10 equal sizes and the inner sum of RMSECV is taken over the observations in 10th segment. By applying K-fold cross-validation, all the datasets are eventually used for both training and testing and the best model will be selected based on lowest RMSECV. The RMSECV for parameter selection and root mean squared error (RMSE) for validation of the best model are given as follows:

$$RMSECV = \sqrt{\frac{1}{K}\frac{1}{Q}\sum_{k=1}^{10}\sum_{i=1}^{Q}(\mathbf{y}_{ki} - \hat{\mathbf{y}}_{ki})^2} \qquad (9)$$

$$RMSE_k = \sqrt{\frac{1}{Q}\sum_{i=1}^{Q}(\mathbf{y}_{ki} - \hat{\mathbf{y}}_{ki})^2} \qquad (10)$$

where $k = 1,2,....,10, Q = 12$ and $K = 10$. $\mathbf{y}_{ki}$ denoted as the original denoised streamflow data and $\hat{\mathbf{y}}_{ki}$ denoted as predicted denoised streamflow data using SVR.

CASE STUDY

Introduction of the area study

The Segamat River Basin is located in the state of Johor, in the southern Peninsular Malaysia. The river is about 777 km long and one of its tributaries is the Muar River which is about 23 km in lenghth and 14m above the sea level that flows through the Segamat town. The location of the Segamat River Basin of Segamat Station is shown in Fig. 2. The streamflow data was observed from specifically in Segamat station for 11 years which is from January 2000 to December 2010 in cubic meter per second (m^3/s). The data that we obtained was taken from Department of Irrigation & Drainage Malaysia. An automatic Water Level Recorder was used to measure and record the river flow data. In the applications, the first 32-year of flow data (75% of the whole datasets) were used for training and the remaining 10-year (25% of the whole datasets) were used for testing.

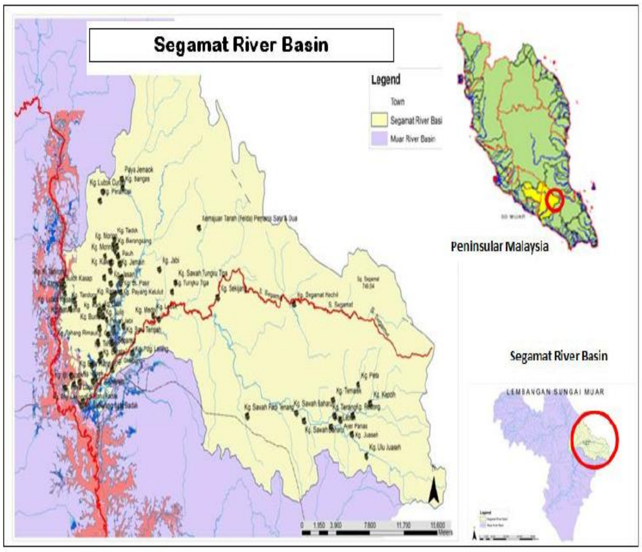

Fig. 2 The Segamat River Basin

Monthly streamflow forecasting

The monthly streamflow data measured from January 2000 until December 2010 of the Segamat station (2528414) is chosen in this research. According to previous research, there are some issues in wavelet decomposition including wavelet choice and decomposition level (Kisi *et al*., 2012). Mostly previous studies have been ignored on these issues and they decide the wavelet choice and decomposition level based on personal preferences when using wavelet-based methods for streamflow predictions (Liu *et al*., 2014). In this paper, these two issues are considered to improve the performance of the SVR based wavelet denoising model.

Daubechies wavelets db3-db7 were used for monthly time series to identify the influence of different mother wavelets on streamflow forecasting. These mother wavelets are used according to experiments and previous researches. Nowadays, there is no existing approach to choose the suitable mother wavelet for the applications in hydrolody. Hence, in this study, we consider a set of daubechies (db) wavelets to select the most suitable mother wavelet for streamflow prediction. The decision of optimal decomposition level is important since it can protect the information and reduces the distortion in original time series datasets (Liu *et al*., 2014). It is depends on the size of time series datasets and the mother wavelet. Based on previous study by de Artigas *et al.* (2006) and Nalley *et al.* (2012), the highest decomposition level, *D*, for the monthly time series of Segamat can be calculated as:

$$D = \frac{Log\left(\dfrac{N}{2m-1}\right)}{Log(2)} \tag{11}$$

where N is the size of the monthly series and m is the number of vanishing moments of a db wavelet. Daubechies wavelets (db3-db7) were used for each monthly series. The maximum decomposition level for different Daubechies wavelets (db3-db7) were between 3.21 and 4.58 for the monthly series. Therefore, the numbers of decomposition level to be used in wavelet denoising are 3 and 4 levels.

The original and de-noised signals of Segamat River is shown in Fig. 3. The red line shows the original streamflow signal while the black line shows the denoised signal after applying wavelet denoising. There exist noise in the original streamflow since we can see the different between original and de-noised signals. The noise in stremaflow time series always with high frequency signal and the wavelet denoising is applied in this research to extract the high frequency signals.

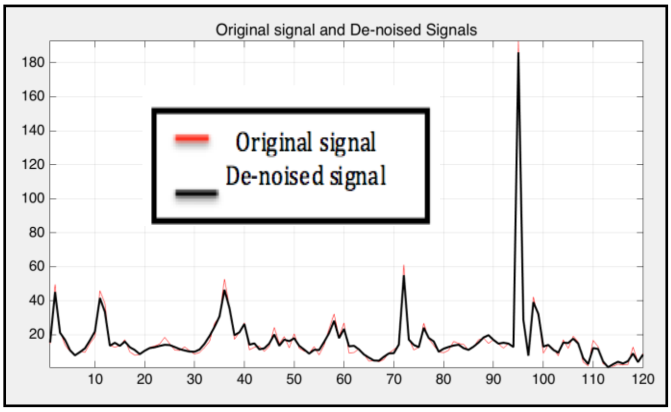

Fig. 3 The Original and De-noised signals of Segamat River

With the results of de-noised signal using wavelet denoising, it will be implemented in SVR to evaluate the SVR based wavelet denoising models for 1-month lead time. In order to obtain the optimal parameters of the model, 10-fold cross-validation is used in SVR. According to Table 1, there are three parameters to be tuned including C, σ and d while ε is set as 0.001.Parameter C represents for the soft margin cost function, σ is the parameter of RBF kernel function, d denotes the degree polynomial kernel function and ε represents the epsilon-insentive value in SVR.

Table 1 SVR parameter settings

Parameter	Value
C	1, 5, 10, 20, 30, ¥
ε	0.001
σ	0.01-22
d	1-10

RESULTS AND DISCUSSION

After performing the forecasting models, the RMSECV statistics of the SVR based wavelet denoising models were compared over the training period to obtain the optimal parameters using *10*-fold cross-validation based on different wavelets, decomposition levels and four types of kernel functions. Table 2, 3 and 4 show the performance statistics of the models during the training period at Segamat station for a 1-month lead time. It can be noted from Table 3 that the best performance with lowest RMSECV was obtained using 3 decomposition levels of db5 wavelet using RBF kernel function. According to this result, it is found that the suitable kernel function of SVR based wavelet denoising is RBF and the optimal parameters are $C = 5, and\ \sigma = 2.09$. Besides, previous researches also have shown that RBF is proved as the best kernel function in SVR for nonlinear forecasting (Wu *et al*., 2009). The results highlight the importance of considering different decomposition levels, mother wavelets and kernel functions in SVR based wavelet denoising analysis. The best SVR based wavelet denoising model is indicated in bold font.

With respect to the monthly streamflow data, the 1-month lead time forecasts over the testing period of 10 years for Segamat station demonstrated the varying accuracy of the selected SVR based wavelet denosing model and regular SVR model. The performance statistics of this comparison models in the testing period are given in Table 5. As can be seen from Table 5, it is noted that the SVR based wavelet denoising model for monthly streamflow forecasting at Segamat station was superior to the regular SVR. In general, the SVR based wavelet denoising seems to be more accurate than the regular SVR model for forecasting streamflow. The original signal in monthly streamflow consists of certain high frequency which is referred as noise that can give influence in streamflow forecasting accuracy. Wavelet denoising method is adopted to eliminate the influence of noise existed in streamflow time series by extracting the high frequency part from the signals. This is why the SVR based wavelet denoising performs better than the regular SVR.

Table 2 Performance of SVR based wavelet denoising models with different mother wavelets and decomposition

Wavelet	Decomposition level	
	Level 3	Level 4
	(RMSECV (m^3/s)	(RMSECV (m^3/s)
db3	13.7375	13.4958
db4	14.0252	13.9878
db5	**13.2398**	13.3347
db6	13.5670	13.7090
db7	13.4833	13.5261

Table 3 Performance statistics of optimal parameter using four different types of kernel functions for SVR based wavelet Denoising

Optimal Value	Model Input	Kernel Function (RMSECV) (m^3/s)			
		RBF	Polynomial	Sigmoid	Linear
RMSEC	db5	**13.239**	14.8048	14.8052	15.171
V	with 3	**8**			5
C	decom	5	¥	1	¥
σ	positio	2.09	-	2.09	-
d	n level	-	6	-	-

Table 4 Performance statistics of optimal parameter using four different types of kernel functions for regular SVR

Optimal Value	Kernel Function (RMSECV) (m^3/s)			
	RBF	**Polynomial**	**Sigmoid**	**Linear**
RMSECV	14.0497	14.8052	14.8009	15.0610
C	5	1	¥	¥
σ	2.09	-	6.91	-
d	-	5	-	-

Table 5 Comparisons of Different Models in the Testing Period for 1-Month ahead Forecasting

Model	(RMSE) (m^3/s)
SVR based wavelet denoising	4.0304
Regular SVR	4.3444

CONCLUSION

The accuracy of the SVR based wavelet denoising model has been investigated for forecasting monthly streamflow of Segamat River in Johor. The SVR based wavelet denoising were obtained using wavelet denoising and support vector regression. We consider two important factors of the wavelet decomposition phase, which are decomposition levels and mother wavelets and different types of kernel function in SVR that will affect the performance of SVR, based wavelet denoising model. The SVR based wavelet denoising is adopted to find the optimal parameter in SVR accoding to lowest RMSECV using *10*-fold cross-validation in the training phase. There are four types of kernel functions have been used including RBF, polynomial, linear and Sigmoid kernel. The forecasting skill of the models was tested using monthly streamflow from Segamat station located in Johor. The test results were compared with the regular SVR for 1-month ahead streamflow forecasts and it demonstrated noticebale differences in the SVR based wavelet denoising models with different combinations of decomposition levels, mother wavelets, and kernel functions. Based on this analysis, the results show that RBF kernel is the most suitable kernel for db5 with 3 level decompositions.

ACKNOWLEDGEMENT

This work was financially supported by the Universiti Teknologi Malaysia under STEM Grant with vote no. A. J091002.5600.07397 and Ministry of Higher Education Malaysia (MOHE).

REFERENCES

Adamowski, J. and Chan, H. F. 2011. A wavelet neural network conjunction model for groundwater level forecasting. *J. Hydrol*. 407: 28–40.

Adamowski, J. and Sun, K. 2010. Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds. *J. Hydrol*. 390 (1–2), 85–91.

Adamowski, J. 2013. Using support vector regression to predict direct runoff, base flow and total flow in a mountaineous watershed with limited data. *Land Reclam.* 45 (1): 71–83.

Bray, M. and Han, D. 2004. Identification of support vector machines for runoff modeling. *J. Hydroinform*. 6 (4): 265–280.

Brito, N. S. P., Souza, B. A. and Pires, F. A. C. 1998. Daubechies wavelets in quality of electrical power. *The international conference on harmonics and quality of power*. 511–515.

Cannas, B., Fanni, A., See, L and Sias, G. 2006. Data preprocessing for river flow forecasting using neural networks: wavelet transforms and data partitioning. *Phys. Chem. Earth*. 31 (18): 1164–1171.

Cheng S. and Pecht, M. 2012. Using cross-validation for model parameter selection of sequencial probability ratio test, *Expert Systems with Applications*. 39: 8467-8473.

Chou, C.-M. and Wang, R.-Y. 2002. On-line estimation of unit hydrographs using the wavelet-based LMS algorithm. *Hydrol. Sci. J*. 47 (5): 721–738.

Cortes, C. and Vapnik, V. 1995, Support-vector networks. *Machine Learning*. 20(3): 273–297.

Coulibaly, P. and Burn, H.D. 2004. Wavelet analysis of variability in annual Canadian streamflows. *Water Resour. Res*. 40: W03105.

Daubechies, I. 1990. The wavelet transform, time–frequency localization and signal analysis. *IEEE Trans. Inform. Theory*. 36 (5): 6–7.

de Artigas, M.Z., Elias, A.G. and *de* Campra, P. F. 2006. Discrete wavelet analysis to assess long-term trends in geomagnetic activity. *Phys. Chem. Earth.* 31 (1–3):77–80.

Guo, J., Zhou, J., Qin, H., Zou, Q. and Li, Q. 2011. Monthly streamflow forecasting based on improved support vector machine model. Expert Systems with Applications. 38: 13073–13081. http://doi.org/10.1016/j.eswa.2011.04.114

Guo, J., Zhou, J., Qin, H., Zou, Q. and Li, Q. 2011. Monthly streamflow forecasting based on improved support vector machine model. Expert Systems with Applications. 38: 13073–13081. http://doi.org/10.1016/j.eswa.2011.04.114

Kalteh, A. M. 2013. Monthly River flow forecasting using artificial neural network and support vector regression models coupled with wavelet transform. Comput. *Geosci*. 54: 1–8.

Kim, G. and Barros, A.P. 2001. Quantitative flood forecasting using multisensor data and neural networks. *J. Hydrol*. 246: 45–62.

Kisi, O. and Cimen, M. 2011. A wavelet-support vector machine conjunction model for monthly streamflow forecasting. *J. Hydrol*. 399: 132–140.

Kisi, O. and Cimen, M. 2012. Precipitation forecasting by using wavelet-support vector machine conjunction model. *Eng. Appl. Artif. Intell*. 25: 783–792.

Kisi, O. 2009. Neural networks and wavelet conjunction model for intermittent stream flow forecasting. *J Hydrol Eng*. 14(8):773–782.

Kisi, O. 2008. Stream flow forecasting using neuro-wavelet technique. *Hydrol Process*. 22:4142–4152.

Li, L. 2011. Water resource requirement prediction based on the wavelet–bootstrap–Svm hybrid model. *J. Inf. Comput. Sci*. 8 (13):2563–2568.

Lin, J. Y., Cheng, C. T. and Chau, K. W. 2006. Using support vector machines for long-term discharge prediction. *Hydrol. Sci. J*. 51 (4): 599–612.

Liong, S. and Sivapragasam, C. 2002. Flood stage forecasting with support vector machines. *J. Am. Water Resour. Assoc*. 38 (1): 173–186.

Liu, F., Zhou, J. Z., Qiu, F. P., Yang, J. J. and Liu, L. 2006. Nonlinear hydrological time series forecasting based on the relevance vector regression. *Neural information processing, LNCS II*. 4233: 880-889.

Liu, Z., Zhou, P., Chen, G. and Guo, L. 2014. Evaluating a coupled discrete wavelet transform and support vector regression for daily and monthly streamflow forecasting. *Journal of Hydrology*. 519: 2822–2831. http://doi.org/10.1016/j.jhydrol.2014.06.050.

Lu, R.Y. 2002. Decomposition of interdecadal and interannual components for North China rainfall in rainy season. *Chinese J. Atmos*. 26:611–624.

Makkeasorn, A., Chang, N. B. and Zhou, X. 2008. Short-term streamflow forecasting with global climate change implications—a comparative study between genetic programming and neural network models. J. *Hydrol*. 352: 336–354.

McKerchar, A. I. and Delleur, J. W. 1974. Application of seasonal parametric linear stochastic models to monthly flow data. *Water Resour. Res*. 10: 246–255.

Milne, A. E., Macleod, C. J. A., Haygarth, P. M., Hawkins, J. M. B. and Lark, R. M. 2009. The wavelet packet transform: A technique for investigating temporal variation of river water solutes. *J. Hydrol*. 379 (1–2): 1–19.

Mohsen, B., Keyvan, A., Morteza, E. and Palhang, M. 2009. Generalization performance of support vector machines and neural networks in runoff modeling. *Expert Systems with Applications*. 36(4):7624–7629.

Nalley, D., Adamowski, J. and Khalil, B. 2012. Using discrete wavelet transforms to analyze trends in streamflow and precipitation in Quebec and Ontario (1954– 2008). *J. Hydrol*. 475: 204–228.

Nourani V, Komasi, M. and Mano, A. 2009. A multivariate ANN-wavelet approach for rainfall-runoff modeling. *Water Resour Manag*. 23(14):2877–2894.

Nourani, V., Kisi, O. and Komasi, M. 2011. Two hybrid artificial intelligence approaches for modeling rainfall–runoff process. *J. Hydrol.* 402 (1–2): 41–59.

Partal, T. and Kucuk, M. 2006. Long-term trend analysis using discrete wavelet components of annual precipitations measurements in Marmara region (Turkey). *Phys. Chem. Earth*. 31: 1189–1200.

Popivanov, I. and Miller, R.J. 2002. Similarity search over time-series data using wavelets. *Proceedings 18th International Conference on Data Engineering.* 212–221.

Pramanik, N., Panda, R. K., Singh, A. 2010. Daily river flow forecasting using wavelet ANN hybrid models. *J. Hydroinform*. 13 (1): 49–63.

Rasouli, K., Hsieh, W. W. and Cannon, A. J. 2012. Daily streamflow forecasting by machine learning methods with weather and climate inputs. *J. Hydrol*. 414– 415: 284–293.

Remesan, R., Shamim, M. A., Han, D. and Mathew, J. 2009. Runoff prediction using an integrated hybrid modelling scheme. *J. Hydrol*. 372: 48–60.

Sang, Y. F. 2013. A review on the applications of wavelet transform in hydrology time series analysis. *Atmos. Res*. 122: 8–15.

Shiri, J. and Kisi, O. 2010. Short-term and long-term stream flow forecasting using a wavelet and neuro-fuzzy conjunction model. *J. Hydrol*. 394 (3–4): 486–493.

Sivapragasam, C., 2001. Liong, S. Y. and Pasha, M. F. K. Rainfall and runoff forecasting with SSA–SVM approach. *J. Hydroinform*. 3 (3): 141–152.

Smith, L. C., Turcotte, D. L. and Isacks, B. 1998. Stream flow characterization and feature detection using a discrete wavelet transform. *Hydrol. Process.* 12:233–249.

Tiwari, M. K. and Chatterjee, C. 2010. Development of an accurate and reliable hourly flood forecasting model using wavelet–bootstrap–ANN (WBANN) hybrid approach. *J. Hydrol.* 1 (394): 458–470.

Torrence, C. and Compo, G. P. 1998. A practical guide to wavelet analysis. *Bull. Am. Meteorol. Soc*. 79: 61–78.

Vapnik, V. 1995. The Nature of Statistical Learning Theory. Springer Verlag, New York, USA.

Vapnik, V., Golwich, S. and Smola, A. J. 1997. Support vector method for function approximation, regression estimation, and signal processing. In: Mozer, M., Jordan, M., Petsche, T. (Eds.), *Advances in Neural Information Processing Systems 9*. MIT Press, Cambridge, Massachusetts, USA. 281–287.

Vapnik, V.N. 1999. An overview of statistical learning theory. *IEEE Transactions on Neural Networks.* 10 (5): 988–999.

Wang, W. C., Chau, K. W., Cheng, C. T. and Qiu, L. 2009. A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *J. Hydrol*. 374 (3–4): 294–306.

Wu, C. H., Tzeng, H., and Lin, R. H. 2009. A novel hybrid genetic algorithm for kernel function and parameter optimization in support vector regression. *Expert Systems with Applications*. (36):4725-4735.

Wu, M. C., Lin, G. F. and Lin, H. Y. 2012. Improving the forecasts of extreme streamflow by support vector regression with the data extracted by self-organizing map. *Hydrol. Process*. http://dx.doi.org/10.1002/ hyp.9584.

Xingang, D., Ping, W. and Jifan, C. 2003. Multiscale characteristics of the rainy season rainfall and interdecadal decaying of summer monsoon in North China. *Chinese Sci. Bull*. 48: 2730–2734.

Xiong, J. Q. and Li, Z. Y. 2005. Sediment-carrying capacity forecasting based on support vector machine. *Journal of Hydraulic Engineering*. 36(10): 1171–1175.

Yu, G. R. and Xia, Z. Q. 2008. Prediction model of chaotic time series based on support vector machine and its application to runoff. *Advances in Water Science*. 19(1): 116–122.

Yu, P., Chen, S. and Chang, I. 2006. Support vector regression for real-time flood stage forecasting. 704–716. http://doi.org/10.1016/j.jhydrol.2006.01. 021.

Yu, P.S., Chen, S. T. and Chang, I. F. 2006. Support vector regression for real-time flood stage forecasting. *J. Hydrol*. 328 (3–4): 704–716.