

# A COMPARITIVE ANALYSIS AND TIME SERIES FORECASTING OF MONTHLY STREAM FLOW DATA USING HYBRID MODEL

## Article history

Received  
6 April 2015  
Received in revised form  
8 April 2015  
Accepted  
2 August 2015

Siraj Muhammed Pandhiani\*, Ani Shabri

Department of Mathematical Sciences, Faculty of Science,  
Universiti Teknologi Malaysia, 81310, UTM Johor Bahru, Johor Darul  
Ta'azim, Malaysia

\*Corresponding author  
pandhiani@hotmail.com

## Graphical abstract



## Abstract

In this study, new hybrid model is developed by integrating two models, the discrete wavelet transform and least square support vector machine (WLSSVM) model. The hybrid model is then used to measure for monthly stream flow forecasting for two major rivers in Pakistan. The monthly stream flow forecasting results are obtained by applying this model individually to forecast the rivers flow data of the Indus River and Neelum Rivers. The root mean square error (RMSE), mean absolute error (MAE) and the correlation (R) statistics are used for evaluating the accuracy of the WLSSVM, the proposed model. The results are compared with the results obtained through LSSVM. The outcome of such comparison shows that WLSSVM model is more accurate and efficient than LSSVM.

**Keywords:** Artificial neural network, modeling, least square support system, discrete wavelet transform

## Abstrak

Dalam kajian ini, model hibrid baru dibangunkan dengan mengintegrasikan dua model, gelombang kecil diskret dan model Kuasa dua Terkecil Vektor Sokong (WLSSVM). Model hibrid kemudiannya digunakan untuk menguji ramalan aliran-aliran bulanan untuk dua sungai utama di Pakistan. Keputusan ramalan aliran bulanan diperolehi dengan menggunakan model ini secara berasingan untuk sungai data Sungai Indus dan Sungai Neelum. Punca min ralat kuasa dua (RMSE), min ralat mutlak (MAE) dan korelasi statistik (R) digunakan untuk menilai keupayaan model WLSSVM yang dicadangkan. Keputusan model yang dicadangkan dibandingkan dengan keputusan yang diperolehi menggunakan LSSVM. Hasil daripada perbandingan ini menunjukkan bahawa model WLSSVM adalah lebih tepat dan cekap daripada LSSVM.

**Kata kunci:** Rangkaian neural buatan, pemodelan, kuasa dua terkecil sistem sokongan, diskret gelombang kecil transformasi

© 2015 Penerbit UTM Press. All rights reserved

## 1.0 INTRODUCTION

The accuracy of stream flow forecasting is a key factor for reservoir operation and water resource management. However, streamflow is one of the most complex and difficult elements of the hydrological cycle due to the complexity of the atmospheric process. Pakistan is mainly an agricultural country with

the world's largest contiguous irrigation system. Therefore, the economy of Pakistan heavily depends on agriculture, hence rainfall is needed to keep its rivers flowing. The Indus River and Neelum River are the largest source of water in different provinces of Pakistan. It forms the backbone of agriculture and food production in Pakistan.

Recently, the support vector machine (SVM) method, which was suggested by [18], has been used in a range of application, including hydrological modeling and water resources process [2]. Several studies have been carried out using SVM in hydrological modeling such as streamflow forecasting [2], rainfall runoff modeling [6] and flood stage forecasting [20]. In the hydrology context, SVM has been successfully applied to forecast the flood stage [12] and to forecast discharge [11]. Previous studies have indicated that SVM is an effective method for streamflow forecasting [2]. As a simplification of SVM [16], [17], have proposed the use of the least squares support vector machines (LSSVM). LSSVM has been used successfully in various areas of pattern recognition and regression problems [8].

In order to make it computationally inexpensive without compromising over reliability and accuracy, LSSVM were introduced [17], this approach depends on computing least square error by considering the input vectors and the obtained vectors(results). The inclusion of this step ensures a higher level of accuracy in comparison to SVM. Besides, LSSVM is found suitable for solving linear equations which is a much needed characteristics. Unlike SVM, LSSVM works under the influence of equality constraints. Equality constraints are instrumental in reducing computational speed. Regarding convergence efficiency LSSVM provides an appreciative level of precision and carries good convergence [4], [5]. LSSVM is still considered in developing stages and is rarely used in addressing the problems of hydrological modeling [14]. However, the model was successfully applied for solving regression, pattern recognition problems [8] and for modeling ecological and environmental systems [21]. Technically both of these methods, i.e. SVM and LSSVM are equally effective. In terms of implementation LSSVM is comparatively easier than SVM. In terms of their generalized performance both of them are comparable [19] and are found reliable.

Though LSSVM is found good in terms of stability and accuracy and generally it seems to be a good choice for training data. The only concern is that for getting generalized results, like any other AI based technique LSSVM also requires huge amount of data for training purposes. With this idea i.e. training it on large database would make it able to deal with most of the variations which are likely to be there in the collected dataset. So there is strong need to utilize the capability of LSSVM by optimizing the input data. The present research addresses this issue and successfully optimizes the training and testing data.

Recently, wavelet theory has been introduced in the field of hydrology, [3]. Wavelet analysis has recently been identified as a useful tool for describing both rainfall and runoff time series [3]. In this regard there has been a sustained explosion of interest in wavelet in many diverse fields of study such as science and engineering. During the last couple of decades, wavelet transform (WT) analysis has become an ideal tool studying of a measured non-stationary times series, through the hydrological process.

An initial interest in the study of wavelets was developed by [9]. Daubechies employed the wavelets technique for signal transmission applications in the electronics engineering. Foufoula Georgiou and Kumar [7] used geophysical applications. Subsequently, [15] attempted to apply wavelet transformation to daily river discharge records to quantify stream flow variability. The wavelet analysis, which is analogous to Fourier analysis is used to decomposes a signal by linear filtering into components of various frequencies and then to reconstruct it into various frequency resolutions. Rao and Bopardikar [13] described the decomposition of a signal using a Haar wavelet technique, which is a very simple wavelet.

The main contribution of this paper is to propose a novel hybrid integrating model of least square support vector machine with wavelet transform model for streamflow river data. In order to achieve this target, the daily streamflow data of Indus River and Neelum River were decomposed into subseries at different scale by Mallat algorithm. Then, effective subseries were summed together and the used as inputs into the LSSVM model for streamflow forecasting. Finally to evaluate the model ability, the proposed model was compared with individual model LSSVM.

## 2.0 METHODS AND MATERIALS

In the subsections, Least square support vector machine models, Discrete wavelet transform method along with the data used for research purpose is explained in details.

### 2.1 Least Square Support Vector Machines (LSSVM) Model

LSSVM optimizes SVM by replacing complex quadratic programming. It achieves this by using least squares loss function and equality constraints. For the purpose of understanding about the construction of the model consider a training sample set represented by  $(x_i, y_i)$  where  $x_i$  represents the input training vector. Suppose that this training vector belongs to 'n' dimensional space i.e.  $R^n$ , so we can write  $x_i \in R^n$ . Similarly, suppose that  $y_i$  represents the output and this output can be described as,  $y_i \in R$ . SVM can be described with the help of Equation (1)

$$y(x) = w^T \phi(x) + b \quad (1)$$

Where  $\phi(x)$  is a function that ensures the mapping of nonlinear values into higher dimensional space. LSSVM formulates the regression problem according to Equation (2)

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

The regression model shown in Equation (1) works under the influence of equality constraints

$$y(x) = w^T \phi(x_i) + b + e_i, \quad i = 1, 2, \dots, n \quad (3)$$

It introduces Lagrange multiplier for:

$$L(w, b, e, \alpha) = \frac{1}{2} w^T w + \frac{\gamma}{2} \sum_{i=1}^n e_i^2 - \sum_{i=1}^n \alpha_i \{w^T \phi(x_i) + b + e_i - y_i\} \quad (4)$$

Where  $\alpha_i$  represents Lagrange multipliers. Since Equation (4) involves more than one variable so, for studying the rate of change partial differentiation of Equation (5-8) is required. Therefore differentiating Equation (4) with respect to  $w, b, e_i$  and  $\alpha_i$  and equating them equal to zero yields the following set of Equations.

$$\frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^n \alpha_i \phi(x_i) \quad (5)$$

$$\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^n \alpha_i = 0 \quad (6)$$

$$\frac{\partial L}{\partial e_i} = 0 \rightarrow \alpha_i = \gamma e_i \quad (7)$$

$$\frac{\partial L}{\partial \alpha_i} = 0 \rightarrow w^T \phi(x_i) + b + e_i - y_i = 0, \quad i = 1, 2, \dots, n \quad (8)$$

Substituting Equation (5-7) in Equation (4) we get the value of 'w'. This 'w' is described according to Equation (8)

$$w = \sum_{i=1}^n \alpha_i \phi(x_i) = \sum_{i=1}^n \gamma e_i \phi(x_i) \quad (9)$$

Putting Equation (9) in Equation (3)

$$y(x) = \sum_{i=1}^n \alpha_i \phi(x_i)^T \phi(x_j) + b = \sum_{i=1}^n \alpha_i K(x_i, x) + b \quad (10)$$

Where,  $K(x_i, x)$ , represents a kernel such that:

$$K(x_i, x) = \phi(x_i)^T \phi(x_j) \quad (11)$$

The  $\alpha$  vector which is a Lagrange Multiplier and the Biased can be computed by solving a set of linear equations shown in Equation (12)

$$\begin{bmatrix} 0 & 1^T \\ 1 & \phi(x_i)^T \phi(x_j) + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (12)$$

Where,  $y = [y_1; \dots; y_n]$ ,  $1 = [1; \dots; 1]$ ,  $\alpha = [\alpha_1; \dots; \alpha_n]$  This eventually constitutes LSSVM model which is described according to Equation (13).

$$y(x) = \sum_{i=1}^n \alpha_i K(x_i, x) + b \quad (13)$$

The model shown in Equation (13) deals with the linear system and solution this linear system is provided by  $\alpha_i, b$ . The high dimensional feature space is defined by a function. This function is generally known as 'kernel function' and is represented by  $K(x_i, x)$

There are various choices available for picking up this function.

## 2.2 Discrete Wavelet Transform (DWT)

A wavelet based forecasting method for time series is introduced which construct by a multiple resolution decomposition of the signal using the redundant trous wavelet transform and has the advantage of being shift – invariant. Wavelet decomposition is used to propose as a learning tool to predict consecutive application data.

Wavelets are becoming an increasingly important tool in time series forecasting. The basic objective of wavelet transformation is to analyze the time series data both in the time and frequency domain by decomposing the original time series in different frequency bands using wavelet functions. Unlike the Fourier transform, in which time series are analyzed using sine and cosine functions, wavelet transformations provide useful decomposition of original time series by capturing useful information on various decomposition levels.

Assuming a continuous time series  $x(t)$ ,  $t \in [-\infty, \infty]$ , a wavelet function can be written as

$$\psi(t, s) = \frac{1}{\sqrt{s}} \psi\left(\frac{t - \tau}{s}\right) \quad (14)$$

Where  $t$  stands for time,  $\tau$  for the time step in which the window function is iterated, and  $s \in [0, \infty]$  for the wavelet scale.  $\psi(t)$  called the mother wavelet can be defined as  $\int_{-\infty}^{\infty} \psi(t) dt = 0$ . The continuous wavelet transform (CWT) is given by

$$W(\tau, s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} x(t) \overline{\psi}\left(\frac{t - \tau}{s}\right) dt \quad (15)$$

Where  $\overline{\psi}(t)$  stands for the complex conjugation of  $\psi(t)$ .  $W(\tau, s)$  presents the sum over all time of the time series multiplied by scale and shifted version of wavelet function  $\psi(t)$ . The use of continuous wavelet transform for forecasting is not practically possible because calculating wavelet coefficient at every possible scale is time consuming and it generates a lot of data.

Therefore, Discrete Wavelet Transformation (DWT) is preferred in most of the forecasting problems because of its simplicity and ability to compute with less time. The DWT involves choosing scales and position on powers of 2, so-called dyadic scales and translations, then the analysis will be much more efficient as well as more accurate. The main advantage of using the DWT is its robustness as it does not include any potentially erroneous assumption or parametric testing procedure [23][27][30]. The DWT can be defined as

$$\psi_{m,n}\left(\frac{t - \tau}{s}\right) = \frac{1}{\sqrt{s_0^{m/2}}} \psi\left(\frac{t - n\tau_0 s_0^m}{s_0}\right) \quad (16)$$

Where,  $m$  and  $n$  are integers that control the scale and time, respectively;  $s_0$  is a specified, fixed dilation step greater than 1; and  $\tau$  is the location parameter, which must be greater than zero. The most common choices for the parameters  $s_0 = 2$  and  $\tau = 1$ . For a discrete time series  $x(t)$  where  $x(t)$  occurs at discrete time  $t$ , the DWT becomes

$$W_{m,n} = 2^{-m/2} \sum_{t=0}^{N-1} \psi(2^{-m}t - n) x(t) \quad (17)$$

Where,  $W_{m,n}$  is the wavelet coefficient for the discrete wavelet at scale  $s = 2^m$  and  $\tau = 2^m n$ . In Eq. (18),  $x(t)$  is time series ( $t = 1, 2, \dots, N-1$ ), and  $N$  is an integer to the power of 2 ( $N = 2^M$ );  $n$  is the time translation parameter, which changes in the ranges  $0 < n < 2M - m$ , where  $1 < m < M$ . According to Mallat's theory [1], the original discrete time series  $x(t)$  can be decomposed into a series of linearity independent approximation and detail signals by using the inverse DWT. The inverse DWT is given by [1][23][30].

$$x(t) = T + \sum_{m=1}^M \sum_{t=0}^{2M-m-1} W_{m,n} 2^{-m/2} \psi(2^{-m}t - n) \quad (18)$$

or in a simple format as

$$x(t) = A_M(t) + \sum_{m=1}^M D_m(t) \quad (19)$$

which  $A_M(t)$  is called approximation sub-series or residual term at levels  $M$  and  $D_m(t)$  ( $m = 1, 2, \dots, M$ ) are detail sub-series which can capture small features of interpretational value in the data.

### 3.0 APPLICATION

In this study the time series of monthly streamflow data of the Neelum and Indus river of Pakistan are used. The Neelum River catchment covers an area of 21359 km<sup>2</sup> and the Indus River catchment covers 1165000 km<sup>2</sup>. The first set of data comprises of monthly streamflow data of Neelum River from January 1983 to February 2012 and the second data of set of streamflow data of Indus River January 1983 to March 2013. In the application, the first 75% of the whole data set were used for training the network to obtain the parameters model. Another 20% of the whole dataset was used for testing.

The comprehensive assessments of model performance atleast mean absolute error (MAE) measures, root mean square error (RMSE) and correlation coefficient (R). Evaluated the results of time series forecasting data to check the performance of all models for forecasting data and training data.

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad (20)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2} \quad (21)$$

$$R = \frac{\frac{1}{n} \sum_{t=1}^n (y_t - \bar{y})(\hat{y}_t - \bar{\hat{y}})}{\sqrt{\left[ \frac{1}{n} \sum_{t=1}^n (y_t - \bar{y})^2 \right] \left[ \frac{1}{n} \sum_{t=1}^n (\hat{y}_t - \bar{\hat{y}})^2 \right]}} \quad (22)$$

Where  $n$  is the number of observation,  $\hat{y}_t$  stands for the forecasting rainfall,  $y_t$  is the observed rainfall at all-time  $t$ .

## 4.0 RESULTS AND DISCUSSION

The analysis of data and design and comparison of model results and discussions are presented in following section.

### 4.1 Fitting Ann Model And Lssvm Model To The Data

In this part, the LSSVM model is tested. In this study the same inputs structures of the datasets which is M1 to M6 were used. In order to obtain the optimal model parameters of the LSSVM, a grid search algorithm and cross-validation method were employed. Many works on the use of the LSSVM in time series modeling and forecasting have demonstrated favorable performances of the RBF [20]. Therefore, RBF is used as the kernel function for streamflow forecasting in this study. The LSSVM model used herein has two parameters ( $\gamma, \sigma^2$ ) to be determined. The grid search method is a common method which was applied to calibrate these parameters more effectively and systematically to overcome the potential shortcomings of the trails and error method. It is a straightforward and exhaustive method to search parameters. In this study, a grid search of ( $\gamma, \sigma^2$ ) with  $\gamma$  in the range 10 to 1000 and  $\sigma^2$  in the range 0.01 to 1.0 was conducted to find the optimal parameters. In order to avoid the danger of over fitting, the cross-validation scheme is used to calibrate the parameters. For each hyper parameter pair ( $\gamma, \sigma^2$ ) in the search space, 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.

For training and forecasting period, obtain the best results for MAE, RMSE and R. Six models (M1 – M6) having various input structures are trained and test by ANN models. The network was trained for 5000 epochs using the back-propagation algorithm with a learning rate of 0.001 and a momentum coefficient of 0.9. (Table 1) lists model performance evaluation results of the M1 – M6 models.

$$y_t = \frac{y}{1.24 y_{\max}} + 0.1 \quad (23)$$

Where  $y_t$  = standardized flow; and  $y_{\max}$  = maximum of the flow values.

Let  $y_t$  represent the river flow at time  $t$ . In the present study, the following combinations of input data of flow were evaluated. As shown in (Table 2) the performance results obtained in the training and testing period of the regular LSSVM approach (i.e. those using original data). For the training and testing phase in Neelum River, the best values of the MSE (0.0021), MAE (0.0281) and R (0.9712) were obtained using input model 5. In the input model 6 has the

smallest MSE (0.0279) and MAE (0.111) whereas it has the highest value of the R (0.8470). For Indus River training and testing phase, the best value of MSE (0.0004) and MAE (0.0095) and R (0.9071) were obtained using input model 2, whereas the for the testing phase the best value of MSE (0.0085), MAE (0.0267) and R (0.8968) were obtained using input model 3.

**Table 1** The model structures for forecasting streamflow

| Input | Original streamflow data                               | DWT of streamflow data                                       |
|-------|--|--|
| 1     | $y_{t-1}$  | $DW_{t-1}$   |
| 2     | $y_{t-1}, y_{t-2}$                                     | $DW_{t-1}, DW_{t-2}$   |
| 3     | $y_{t-1}, y_{t-2}, y_{t-3}$                            | $DW_{t-1}, DW_{t-2}, DW_{t-3}$                               |
| 4     | $y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}$                   | $DW_{t-1}, DW_{t-2}, DW_{t-3}, DW_{t-4}$                     |
| 5     | $y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}$          | $DW_{t-1}, DW_{t-2}, DW_{t-3}, DW_{t-4}, DW_{t-5}$           |
| 6     | $y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}, y_{t-6}$ | $DW_{t-1}, DW_{t-2}, DW_{t-3}, DW_{t-4}, DW_{t-5}, DW_{t-6}$ |

**Table 2** Training and testing performance indicates of LSSVM Model

| Data   | Input | Training      |               |               | Testing       |               |               |
|--------|-------|---------------|---------------|---------------|---------------|---------------|---------------|
|        |       | MSE           | MAE           | R             | MSE           | MAE           | R             |
| Neelum | 1     | 0.0129        | 0.0863        | 0.8007        | 0.0271        | 0.1180        | 0.7869        |
|        | 2     | 0.0029        | 0.0354        | 0.9597        | 0.0297        | 0.1152        | 0.8174        |
|        | 3     | 0.0025        | 0.0313        | 0.9651        | 0.0352        | 0.1213        | 0.7869        |
|        | 4     | 0.0023        | 0.0300        | 0.9685        | 0.0396        | 0.1243        | 0.7684        |
|        | 5     | <b>0.0021</b> | <b>0.0281</b> | <b>0.9712</b> | 0.0458        | 0.1352        | 0.7336        |
|        | 6     | 0.0031        | 0.0344        | 0.9565        | <b>0.0279</b> | <b>0.1111</b> | <b>0.8470</b> |
| Indus  | 1     | 0.0005        | 0.0110        | 0.8745        | 0.0080        | 0.0250        | 0.7991        |
|        | 2     | <b>0.0004</b> | <b>0.0095</b> | <b>0.9071</b> | 0.0087        | 0.0266        | 0.8561        |
|        | 3     | 0.0004        | 0.0097        | 0.9035        | <b>0.0085</b> | <b>0.0267</b> | <b>0.8968</b> |
|        | 4     | 0.0004        | 0.0096        | 0.9055        | 0.0083        | 0.0265        | 0.8958        |
|        | 5     | 0.0004        | 0.0098        | 0.9043        | 0.0089        | 0.0275        | 0.8825        |
|        | 6     | 0.0004        | 0.0096        | 0.9057        | 0.0089        | 0.0274        | 0.8774        |

#### 4.2 Fitting Hybrid Models Wavelets-ANN Model And Wavelet-LSSVM Model To The Data

A hybrid model Wavelet-LSSVM (WLSSVM) model is obtained by combining two methods, discrete transform (DWT) and LSSVM. In WLSSVM, the original time series was decomposed into a certain number of sub-time series components which were entered LSSVM in order to improve the model accuracy. In this study, the Deubechies wavelet, one of the most widely used wavelet families, is chosen as the wavelet function to decompose the original series [1][10]. The observed series was decomposed into a number of wavelet components, depending on the selected decomposition levels. Deciding the optimal decomposition level of the time series data in wavelet analysis plays an important role in preserving the information and reducing the distortion of the datasets. However, there is no existing theory to tell how many decomposition levels are needed for any time series. To select the number of decomposition levels, the following formula is used to determine the decomposition level [10].

$$M = \log(n)$$

Where,  $n$  is length of the time series and  $M$  is decomposition level. In this study,  $n = 350$  and  $n = 483$ , monthly data are used for Neelum and Indus, respectively, which approximately gives  $M = 3$  decomposition levels. Three decomposition levels are employed in this study, the same as studies employed by [12]. The observed time series of discharge flow data was decomposed at 3 decomposition levels (2 – 4 – 8 months).

The effectiveness of wavelet components is determined using the correlation between the observed streamflow data and the wavelet coefficients of different decomposition levels. (Table 3) shows the correlations between each wavelet component time series and original monthly stream flow data. It is observed that the D1 component shows low correlations. The correlation between the wavelet component D2 and D3 of the monthly stream flow and the observed monthly stream flow data show significantly higher correlations compared to the D1 components. Afterward, the significant wavelet components D2, D3 and approximation (A3) component were added to each other to constitute the new series. For the WLSSVM model, the new series is used as inputs to the LSSVM model. (Figure 1) and (Figure 2) shows the original streamflow data time and

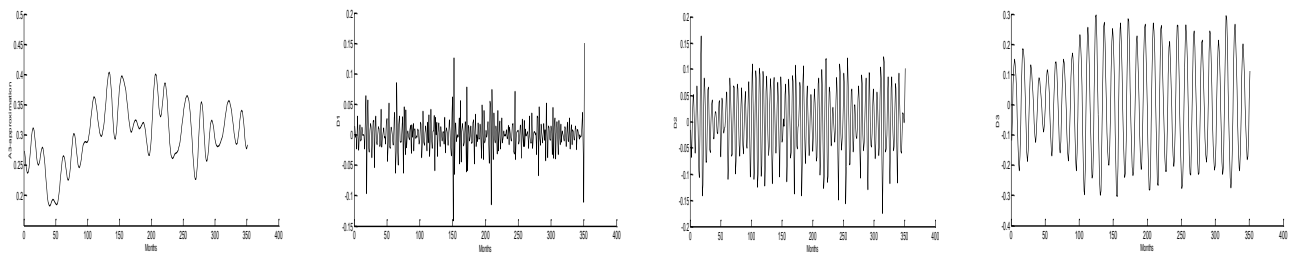
their  $D_s$ , that is the time series of 2-month mode (D1), 4-month mode (D2), 8-month mode (D3), approximate mode (A3), and the combinations of effective details and approximation components mode (A2 + D2 + D3). Six different combinations of the new series input data (Table 1) is used for forecasting as in the previous application.

A program code including wavelet toolbox was written in MATLAB language for the development of LSSVM. The forecasting performances of the Wavelet-LSSVM (WLSSVM) model is presented in (Table 4) respectively, in terms of MSE, MAE and R in training and testing periods.

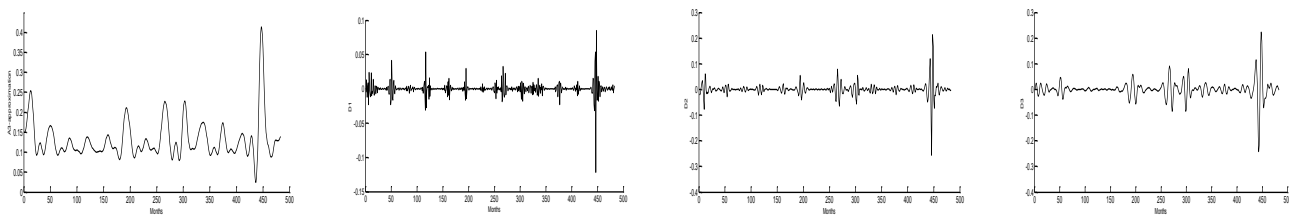
As seen in the (Table 4), the WLSSVM models are evaluated based on their performances in the training data and testing data. For the training phase of Neelum river, the best value of the MSE (0.0006), MAE (0.0168) and R (0.9919) statistics for input data model 6. However, for the testing phase, the best MSE (0.0053), MAE (0.0520) and R (0.9705) were obtained for the input combination model 5. In the other hand, for the Indus river, the input model 6 obtained lowest value of the MSE (0.0000) and MAE (0.0032) and the highest R (0.9926) in the training phase. However, for the testing phase, the best MSE (0.0020) and MAE (0.0178) and R (0.9398) was obtained for the input combination model 2.

**Table 3** Correlation coefficients between each of sub-time series for streamflow data

| Data   | Discrete Wavelet Components | $D_{t-1}/Q_t$ | $D_{t-2}/Q_t$ | $D_{t-3}/Q_t$ | $D_{t-4}/Q_t$ | $D_{t-5}/Q_t$ | $D_{t-6}/Q_t$ | Mean Absolute Correlation |
|--------|-----------------------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------------------|
| Neelum | D1                          | -0.0900       | -0.0320       | 0.0670        | 0.0030        | -0.0350       | -0.0070       | 0.0157                    |
|        | D2                          | 0.0790        | -0.2880       | -0.3700       | -0.1340       | 0.1650        | 0.2530        | 0.0492                    |
|        | D3                          | 0.7930        | 0.4690        | 0.0080        | -0.4560       | -0.7700       | -0.8660       | 0.1370                    |
|        | A3                          | 0.2830        | 0.2700        | 0.2450        | 0.2200        | 0.1770        | 0.1500        | 0.2242                    |
| Indus  | D1                          | 0.1030        | -0.2880       | -0.3540       | -0.1180       | 0.1200        | 0.2020        | 0.0558                    |
|        | D2                          | 0.4450        | 0.2230        | -0.0830       | -0.3570       | -0.5030       | -0.4850       | 0.1267                    |
|        | D3                          | 0.4450        | 0.2230        | -0.0830       | -0.3570       | -0.5030       | -0.4850       | 0.1267                    |
|        | A3                          | 0.7690        | 0.7440        | 0.6830        | 0.5910        | 0.4720        | 0.3370        | 0.5993                    |



**Figure 1** Decomposed wavelets sub-series components ( $D_s$ ) of streamflow data of Neelum River



**Figure 2** Decomposed wavelets sub-series components ( $D_s$ ) of streamflow data of Indus River

**Table 4** Training and testing performance indicates of WLSSVM model

| Data   | Input | Training      |               |               | Testing       |               |               |
|--------|-------|---------------|---------------|---------------|---------------|---------------|---------------|
|        |       | MSE           | MAE           | R             | MSE           | MAE           | R             |
| Neelum | 1     | 0.0116        | 0.0823        | 0.8238        | 0.0232        | 0.1146        | 0.8098        |
|        | 2     | 0.0017        | 0.0303        | 0.9766        | 0.0084        | 0.0653        | 0.9463        |
|        | 3     | 0.0014        | 0.0281        | 0.9799        | 0.0068        | 0.0595        | 0.9607        |
|        | 4     | 0.0008        | 0.0196        | 0.9893        | 0.0105        | 0.0740        | 0.9316        |
|        | 5     | 0.0010        | 0.0219        | 0.9862        | <b>0.0053</b> | <b>0.0520</b> | <b>0.9705</b> |
|        | 6     | <b>0.0006</b> | <b>0.0168</b> | <b>0.9919</b> | 0.0141        | 0.0802        | 0.9019        |
| Indus  | 1     | 0.0004        | 0.0098        | 0.8994        | 0.0026        | 0.0188        | 0.9137        |
|        | 2     | 0.0001        | 0.0065        | 0.9624        | <b>0.0020</b> | <b>0.0178</b> | <b>0.9398</b> |
|        | 3     | 0.0001        | 0.0056        | 0.9763        | 0.0053        | 0.0230        | 0.9050        |
|        | 4     | 0.0001        | 0.0041        | 0.9860        | 0.0068        | 0.0251        | 0.7872        |
|        | 5     | 0.0001        | 0.0047        | 0.9836        | 0.0076        | 0.0246        | 0.8320        |
|        | 6     | <b>0.0000</b> | <b>0.0032</b> | <b>0.9926</b> | 0.0067        | 0.0227        | 0.8052        |

## 5.0 COMPARISON OF FORECASTING MODELS

Finally, in order to evaluate the efficiency of the proposed hybrid model, the obtained results was also compared with the results of LSSVM model using the same data. The comparison has been summarized in the (Table 5). (Table 5) shows that the hybrid model WLSSVM has good performance during the testing

phase, and outperform single model LSSVM in term of all the standard statistical measures. It is observed that the proposed model yields better result than the other models for both streamflow data. This result shows that the new input series from discrete wavelet transforms have significant extremely positive effect on LSSVM model results.

**Table 5** The performance results ANN, LSSVM, WANN and WLSSVM Approach during testing period

| Data   | Model  | MSE    | MAE    | R      |
|--------|--------|--------|--------|--------|
| Neelum | ANN    | 0.1467 | 0.1072 | 0.8866 |
|        | LSSVM  | 0.0279 | 0.1111 | 0.8470 |
|        | WANN   | 0.0011 | 0.0247 | 0.9735 |
|        | WLSSVM | 0.0053 | 0.0520 | 0.9705 |
| Indus  | ANN    | 0.0922 | 0.0288 | 0.9195 |
|        | LSSVM  | 0.0085 | 0.0267 | 0.8968 |
|        | WANN   | 0.0015 | 0.0290 | 0.9597 |
|        | WLSSVM | 0.0020 | 0.0178 | 0.9398 |

## 6.0 CONCLUSION

In this study a new method based on the WLSSVM is developed by combining the discrete wavelet transforms (DWT) and LSSVM model for forecasting streamflows. The monthly streamflow time series is decomposed at different decomposition levels by DWT. Each of the decompositions carried most of the information and plays distinct role in original time series. The correlation coefficients between each of sub-series and original streamflow series are used for the selection of the LSSVM model inputs and for the determination of the effective wavelet components on streamflow. The monthly streamflow time series data is decomposed at 3 decomposition levels (2–4–8 months). The sum of effective details and the approximation component are used as inputs to the LSSVM model. The WLSSVM model are trained and tested by applying different input combinations of monthly streamflow data of Neelum River and Indus River of Pakistan. Then, LSSVM models are constructed with new series as inputs and original streamflow time series as output. The performance of the proposed

WLSSVM model is compared to the regular LSSVM model for monthly streamflow forecasting. Comparison results indicated that the WLSSVM model is substantially more accurate than LSSVM model. The study concludes that the forecasting ability of the LSSVM model is found to be improved when the wavelet transformation technique is adopted for the data pre-processing. The decomposed periodic components obtained from the DWT technique are found to be most effective in yielding accurate forecast when used as inputs in the LSSVM model.

## Acknowledgement

We are grateful for the UTM giving us opportunity to express our research ability.

## References

- [1] Grossman, A., and Morlet, J. 1984. Decomposition of Harley Functions into Square Integral Wavelets of Constant Shape. *SIAM J. Math. Anal.* 15: 723-736.
- [2] Asefa, T., Kemblowski, M., McKee, M., and Khalil, A. 2006. Multi-time Scale Stream Flow Predictions: The Support Vector Machines Approach. *J. Hydrol.* 318(1-4): 7-16.
- [3] Krishna, B., Satyajai Rao, Y. R., Nayak, P. C. 2011. Time Series Modeling of River Flow Using Wavelet Neural Networks. *Journal of Water Resources and Protection.* 3: 50-59.
- [4] Bose, G. E. P., and Jenkins, G. M. 1970. Time Series Analysis Forecasting and Control. Holden Day, San Francisco.
- [5] Delleur, J. W., Tao, P. C., and Kavvas, M. 1976. An Evaluation of the Practicality and Complexity of Some Rainfall and Runoff Time Series Model. *Water Resources.* 12(5): 953-970.
- [6] Dibike, Y. B., Velickov, S., Solomatine, D. P., and Abbott, M. B. 2001. Model Induction with Support Vector Machines: Introduction and Applications. *ASCE J. Comput. Civil Eng.* 15(3): 208-216.
- [7] Foufoula-Georgiou & Kumar, P. E. 1994. *Wavelets in Geophysics.* Ed. E. Foufoula-Georgiou & P. Kumar. San Diego and London: Academic.
- [8] Hanbay, D. 2009. An Expert System Based on Least Square Support Vector Machines for Diagnosis of Valvular Heart Disease. *Expert Syst. Appl.* 36(4): 8368-8374.
- [9] Daubechies, I. 1988. Orthogonal Bases of Compactly Supported Wavelets. *Commun. Pure Applied Math.* 41: 909-996.
- [10] Kisi, O. 2010. Wavelet Regression Model for Short-term Streamflow Forecasting. *J. Hydrol.* 389: 344-353.
- [11] Lin, J. Y., Cheng, C. T., and Chau, K. W. 2006. Using Support Vector Machines for Long-term Discharge Prediction. *Hydrology Sci. J.* 51 (4): 599-612.
- [12] Ma, P. Y. 2006. *A Fresh Engineering Approach for the Forecast of Financial Index Volatility and Hedging Strategies.* PhD Thesis, Quebec University, Montreal, Canada.
- [13] Rao, R. M. and Bopardikar, A. S. 1998. *Wavelet Transforms: Introduction to Theory and Applications.* Addison Wesley Longman, Inc. 310.
- [14] Ismail, S. Samsuddin, R. and Shabri, A. 2010. A Hybrid Model of Self Organizing Maps and Least Square Support Vector Machine. *Hydrol. Earth Syst. Sci. Discuss.* 7: 8179-8212.
- [15] Mallat, S. G. 1998. A Theory for Multi Resolution Signal Decomposition: The Wavelet Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 11(7): 674-693. Doi:10.1109/34.192463.
- [16] Pandhiani, S. and Shabri, A. 2013. Time Series Forecasting Using Wavelet-Least Squares Support Vector Machines and Wavelet Regression Models for Monthly Stream Flow Data. *Open Journal of Statistics.* 3(3): 183-194. doi: 10.4236/ojs.2013.33021.
- [17] Suykens, J. A. K., and Vandewalle, J. 1999. Least Squares Support Vector Machine Classifiers. *Neural Processing Letters.* 3: 293-300.
- [18] Vapnik, V. 1995. *The Nature of Statistical Learning Theory.* Springer Verlag, Berlin.
- [19] Wang, H. and Hu, D. 2005. Comparison of SVM and LS-SVM for Regression. *IEEE.* 279-283.
- [20] 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.
- [21] Yunrong, X. and Liangzhong, J. 2009. Water Quality Prediction Using LS-SVM With Particle Swarm Optimization. *Second International Workshop on Knowledge Discovery and Data Mining.* 900-904.