

Time Series Forecasting Using Wavelet-Least Squares Support Vector Machines and Wavelet Regression Models for Monthly Stream Flow Data

Siraj Muhammed Pandhiani, Ani Bin Shabri

Faculty of Science, Department of Mathematics, Universiti Teknologi Malaysia, Skudai, Malaysia Email: pandhiani@hotmail.com, ani@utm.my

Received March 24, 2013; revised April 25, 2013; accepted May 2, 2013

Copyright © 2013 Siraj Muhammed Pandhiani, Ani Bin Shabri. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

This study explores the least square support vector and wavelet technique (WLSSVM) in the monthly stream flow forecasting. This is a new hybrid technique. The 30 days periodic predicting statistics used in this study are derived from the subjection of this model to the river flow data of the Jhelum and Chenab rivers. The root mean square error (RMSE), mean absolute error (RME) and correlation (R) statistics are used for evaluating the accuracy of the WLSSVM and WR models. The accuracy of the WLSSVM model is compared with LSSVM, WR and LR models. The two rivers surveyed are in the Republic of Pakistan and cover an area encompassing 39,200 km² for the Jhelum River and 67,515 km² for the Chenab River. Using discrete wavelets, the observed data has been decomposed into sub-series. These have then appropriately been used as inputs in the least square support vector machines for forecasting the hydrological variables. The resultant observation from this comparison indicates the WLSSVM is more accurate than the LSSVM, WR and LR models in river flow forecasting.

Keywords: River Flow; Time Series; Least Square Support Machines; Wavelet

1. Introduction

Pakistan is the home to the world's largest contiguous irrigation system. This is an economy that is largely dependent on the agricultural sector; therefore, it has a vast irrigation network. The dependence on this form of farming is developed by a persistent rainfall pattern that keeps the rivers flowing, which in turn leads to the availability of irrigation waters flowing downstream. The Jhelum and Chenab rivers provide the biggest percentage of irrigation waters in various provinces in Pakistan.

These days, river flow forecasting has a major role to play in water resources system planning. The sources of water with many activities such as planning and operating system component estimate for future demand. The composition of water is necessary for both short-term and long-term forecasts of the event flow to optimize system or an application for the growth or decline in the future. There are many mathematical models to predict future flow of rivers such as discussed by [1-4].

The Support Vector Machine (SVM) [5] forecasting method has been used in various studies of hydrology

water modeling and water resource processed such as flood stage forecasting [6], rainfall runoff modeling [7-9] and stream flow forecasting [10-12]. Researcher solved the standard SVM models using the complicated computational programming techniques, which are very expensive in time to handle the required optimization programming.

The Suykens and Vandewalls [13] proposed the least square support vector machines (LSSVM) model to simplify the SVM. In the various areas, such as regression problems and pattern recognition [14,15], the LSSVM model has been used successfully. LSSVM and SVM have almost similar advantages but the LSSVM has an additional advantage, e.g. it needs to solve only a linear system of equation, which is much easier to solve and predict results [1-4]. The LSSVM and SVM mathematical models have been used in predicting ad analyzing the future flow of rivers.

Recently, wavelet theory has been introduced in the field of hydrology, [16-18]. Wavelet analysis has recently been identified as a useful tool for describing both rainfall and runoff time series [16,19]. 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 [20-22]. Daubechies [22] employed the wavelets technique for signal transmission applications in the electronics engineering. Foufoula Georgiou and Kumar [23] used geophysical applications. Subsequently, [17] 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 [24] described the decomposition of a signal using a Haar wavelet technique, which is a very simple wavelet.

Wavelet spectrum, based on the continuous wavelet transform, has been proven to be a natural extension of the much more familiar conventional Fourier spectrum analysis which is usually associated with hydro metrological time series analysis [25]. Instead of the results being presented in a plot of energy *vis a vis* frequency for energy spectrum in Fourier Transform (FT) as well as FFT (Fast Fourier Transform), the wavelet spectrum is three dimensional and is plotted in the time frequency domain in which the energy is portrayed as contours. The wavelets are mathematical functions that break up a signal into different frequency components so that they are studied at different resolutions or scales. They are considered better than the Fourier analysis for their distinct signals that pose discontinuities and sharp spikes.

The main purpose of the this study is to investigate the performance of the WLSSVM model for streamflow forecasting and to compare it with the performance of the least square support vector machines (LSSVM), linear regression (LR) and wavelet regression models (WR).

2. Methods and Materials

2.1. The Least Square Vector Machines Model

LSSVM is a new version of SVM modified by [13]. LSSVM involves the solution of a quadratic optimization problem with a least squares loss function and equality constraints instead of inequality constraints. In this section, we briefly introduce the basic theory LSSVM in time series forecasting. Consider a training sample set (x_i, y_i) with input $x_i \in \mathbb{R}^n$ and output $y_i \in \mathbb{R}$. In feature space SVM models take the form

$$y(x) = w^{\mathrm{T}}\phi(x) + b \tag{1}$$

where the nonlinear mapping $\phi(x)$ maps the input data into a higher dimensional feature space. LSSVM introduces a least square version to SVM regression by formulating the regression problem as

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

subject to the equality constraints

$$y(x) = w^{T}\phi(x_{i}) + b + e_{i}, i = 1, 2, \cdots, n$$
 (3)

To solve this optimization problem, Lagrange function is constructed as

$$L(w,b,e,\alpha) = \frac{1}{2}w^{\mathrm{T}}w + \frac{\gamma}{2}\sum_{i=1}^{n}e_{i}^{2} - \sum_{i=1}^{n}\alpha_{i}\left\{w^{\mathrm{T}}\phi(x_{i}) + b + e_{i} - y_{i}\right\}$$
(4)

where, α_i is Lagrange multipliers. The solution of (6) can be obtained by partially differentiating with respect to w,b,e_i and α_i

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

$$\frac{\partial L}{\partial b} = 0 \longrightarrow \sum_{i=1}^{n} \alpha_i = 0 \tag{6}$$

$$\frac{\partial L}{\partial e_i} = 0 \to \alpha_i = \gamma e_i \tag{7}$$

$$\frac{\partial L}{\partial \alpha_i} = 0 \to w^{\mathrm{T}} \phi(x_i) + b + e_i - y_i = 0, \qquad (8)$$

$$i = 1, 2, \cdots, n$$

then the weight w can be written as combination of the Lagrange multipliers with the corresponding data training x_i .

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

If we put the result of Equation (9) into Equation (3), then the following result is obtained:

$$y(\mathbf{x}) = \sum_{i=1}^{n} \alpha_{i} \phi(x_{i})^{\mathrm{T}} \phi(x_{i}) + b$$

$$= \sum_{i=1}^{n} \alpha_{i} K(\mathbf{x}_{i}, \mathbf{x}) + b$$
 (10)

where, a positive definite kernel is defined as follows:

$$K(\boldsymbol{x}_i, \boldsymbol{x}) = \phi(\boldsymbol{x}_i)^{\mathrm{T}} \phi(\boldsymbol{x}_i)$$
(11)

The α vector and b can be found by solving a set of linear equations:

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

Copyright © 2013 SciRes.

where, $y = [y_1; \dots; y_n], \mathbf{1} = [1; \dots; 1], \alpha = [\alpha_1; \dots; \alpha_n]$. This finally leads to the following LSSVM model for function estimation:

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

where α_i , *b* are the solution to the linear system. Kernel function, $K(\mathbf{x}_i, \mathbf{x})$ represents the high dimensional feature space that is nonlinearly mapped from input space *x*. The typical examples of the kernel function are as follows:

Linear :

$$K(x_{i}, x) = x_{i}^{T} x$$
Multi - Layer Perceptron Kernel :

$$K(x_{i}, x) = \tanh(\gamma x_{i}^{T} x + r)$$
Polynomial :

$$K(x_{i}, x) = (\gamma x_{i}^{T} x + r)^{d}, \gamma > 0$$
Radial basis function (*RBF*) :

$$K(x_{i}, x) = \exp(-\gamma ||x_{i} - x||^{2}), \gamma > 0$$
(14)

Here γ , *r* and *d* are the kernel parameters. The architecture of LSSVM is shown in **Figure 1**.

2.2. Wavelet Analysis

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:

$$\psi(t,s) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right) \tag{15}$$

where t stands for time, τ 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

$$W(\tau, s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} x(t) \overline{\psi} \left(\frac{t-\tau}{s}\right) dt$$
(16)

where $\overline{\psi}(t)$ stands for the complex conjugation of $\psi(t)$. $W(\tau, s)$ presents the sum of over all time period

Copyright © 2013 SciRes.



Figure 1. Architecture of LSSVM.

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 [26-28]. The DWT can be defined as

$$\psi_{m,n}\left(\frac{t-\tau}{s}\right) = \frac{1}{\sqrt{s_0^{m/2}}}\psi\left(\frac{t-n\tau_0 s_0^m}{s_0^m}\right)$$
(17)

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 τ_0 is the location parameter, which must be greater than zero. The most common choices for the parameters $s_0 = 2$ and $\tau_0 = 1$. For a discrete time series x(t) where x(t) occurs at discrete time *t*, the DWT becomes

$$W_{m,n} = 2^{-m/2} \sum_{t=0}^{N-1} \psi\left(2^{-m}t - n\right) x(t)$$
(18)

where $W_{m,n}$ is the wavelet coefficient for the discrete wavelet at scale $s = 2^m$ and $\tau = 2^m n$. In Equation (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 [20], 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 [20,26,27]

$$x(t) = T + \sum_{m=1}^{M} \sum_{t=0}^{2^{M-m-1}} W_{m,n} 2^{-m/2} \psi(2^{-m}t - n)$$
(19)

or in a simple format as

$$x(t) = A_M(t) + \sum_{m=1}^{M} D_m(t)$$
(20)

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. Study Area

The time series of monthly streamflow data of the Jhelum and Chenab river of Pakistan are used. The locations of the Jhelum and Chenab catchments are shown in **Figure 2**. The Jhelum River catchment covers an area of 21,359 km² and the Chenab catchment covers 28,000 km². The first set of data comprises of monthly streamflow data of Chanari station at Jhelum River from Jan 1970 to December 1996 and the second data of set of streamflow data of Marala station at Chenab River from April 1947 to March 2007. In the applications, the first 20 years and 51 year of flow Jhelum data and Chenab data (237 months and 430 months, 80% of the whole data set) were used for training the network to obtain the parameters model. Another dataset consisting of 75 monthly records (20% of the whole data) was used for testing.

The performances of each model for both training and testing data are evaluated by using the mean-square error (MSE), mean absolute error (MAE) and correlation coefficient (R) which is widely used for evaluating results of time series forecasting. MSE, MAE and R are defined as

$$MSE = \frac{1}{n} \sum_{t=1}^{n} (y_{t}^{o} - y_{t}^{f})^{2}$$

$$MAE = \frac{1}{n} \sum_{t=1}^{n} |y_{t}^{o} - y_{t}^{f}|$$

$$R = \frac{\frac{1}{n} \sum_{t=1}^{n} (y_{t}^{o} - \overline{y}_{t}^{o}) (y_{t}^{f} - \overline{y}_{t}^{f})}{\sqrt{\frac{1}{n} \sum_{t=1}^{n} (y_{t}^{o} - \overline{y}_{t}^{o})^{2}} \sqrt{\frac{1}{n} \sum_{t=1}^{n} (y_{t}^{f} - \overline{y}_{t}^{f})^{2}}$$
(21)

where y_t^o and y_t^f are the observed and forecasted values at time t, respectively and n is the number of data points. The criteria to judge for the best model are



Figure 2. Location map of the study area.

relatively small of MAE and MSE in the training and testing. Correlation coefficient measures how well the flows predicted, correlate with the flows observed. Clearly, the R value close to unity indicates a satisfactory result, while a low value or close to zero implies an inadequate result.

4. Model Structures

One of the most important steps in developing a satisfactory forecasting model such as LSSVM and LR models is the selection of the input variables. The appropriate input variables will allow the network to successfully map the desired output and avoid loss of important information. There are no fixed rules in selection of input variables for developing this model, even though a general framework can be followed based on previous successful application in water resources problems [12,29, 30]. In this study, the six model structures were developed to investigate variable enabling of input variables on model performance. Six model structures are accomplished by setting the input variables equal to the number of the lagged variables from monthly stream flows of previous periods data, $y_{t-1}, y_{t-2}, \dots, y_{t-p}$, where p is set 1, 2, 3, 4, 5 and 6 months. The model structure for original and DWT monthly streamflow data can be mathematically expressed as

$$y_t = f(y_{t-1}, y_{t-2}, \cdots, y_{t-p})$$

and

$$y_t = f\left(DW_{t-1}, DW_{t-2}, \cdots, DW_{t-p}\right)$$

186

where y_{t-1} denotes the streamflow value of 1 previous month, DW_{t-1} is obtained by adding the effective Ds (D2, D4 and D8), and approximately components of the y_{t-1} values. The model input structures for forecasting streamflow of Jehlum River and Chenab River is shown in **Table 1**.

5. Results and Discussion

5.1. Fitting LSSVM to the Data

The selection of appropriate input data sets is an important consideration in the LSSVM modeling. In the training and testing of LSSVM model, the same input structures of the data set (M1 - M6) were used. In order to obtain the optimal model parameters of the LSSVM, a grid search algorithm and cross-validation method was employed. Many work on the use of the LSSVM in time series modeling and forecasting have demonstrated favorable performances of the RBF [6,31,32]. Therefore, RBF is used as the kernel function for streamflow forecasting in this study. The LSSVM model used herein has two parameters (γ, σ^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 (γ, σ^2) with γ in the range 10 to 1000 and σ^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 (γ, σ^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.

Tables 2(a) and **2(b)** show the performance results obtained in the training and testing period of the regular LSSVM approach (*i.e.* those using original data). For the

Table 1. The model structures for forecasting streamflow of Jhelum and Chenab Rivers.

| Model | Original streamflow data | Model | DWT of streamflow data |
|------------|--|-------|--|
| <i>M</i> 1 | <i>Y</i> _{r-1} | MW1 | DW_{r-1} |
| <i>M</i> 2 | $\mathcal{Y}_{t-1}, \mathcal{Y}_{t-2}$ | MW2 | DW_{r-1}, DW_{r-2} |
| М3 | $y_{t-1}, y_{t-2}, y_{t-3}$ | MW3 | $DW_{_{t-1}}, DW_{_{t-2}}, DW_{_{t-3}}$ |
| <i>M</i> 4 | $y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}$ | MW4 | $DW_{_{t-1}}, DW_{_{t-2}}, DW_{_{t-3}}, DW_{_{t-4}}$ |
| M 5 | $y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}$ | MW5 | $DW_{_{t-1}}, DW_{_{t-2}}, DW_{_{t-3}}, DW_{_{t-4}}, DW_{_{t-5}}$ |
| <i>M</i> 6 | $y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}, y_{t-6}$ | MW6 | $DW_{_{t-1}}, DW_{_{t-2}}, DW_{_{t-3}}, DW_{_{t-4}}, DW_{_{t-5}}, DW_{_{t-6}}$ |
| | | | |

Table 2. Forecasting performance indicates of LSSVM for (a) Jhelum River of Pakistan; (b) Chenab River of Pakistan.

| | | | (a) | | | |
|-------|--------|----------|--------|--------|---------|--------|
| Madal | | Training | | | Testing | |
| Woder | MSE | MAE | R | MSE | MAE | R |
| M1 | 0.0160 | 0.0974 | 0.7757 | 0.0133 | 0.0892 | 0.8064 |
| M2 | 0.0104 | 0.0730 | 0.8615 | 0.0102 | 0.0727 | 0.8556 |
| M3 | 0.0077 | 0.0626 | 0.8993 | 0.0075 | 0.0662 | 0.8943 |
| M4 | 0.0096 | 0.0715 | 0.8730 | 0.0090 | 0.0744 | 0.8777 |
| M5 | 0.0087 | 0.0673 | 0.8851 | 0.0081 | 0.0703 | 0.8719 |
| M6 | 0.0084 | 0.0655 | 0.8885 | 0.0090 | 0.0752 | 0.8775 |
| | | | (b) | | | |
| Madal | | Training | | | Testing | |
| Woder | MSE | MAE | R | MSE | MAE | R |
| M1 | 0.0123 | 0.0810 | 0.7622 | 0.0083 | 0.0694 | 0.7644 |
| M2 | 0.0047 | 0.0427 | 0.9167 | 0.0041 | 0.0421 | 0.9098 |
| M3 | 0.0039 | 0.0391 | 0.9307 | 0.0036 | 0.0416 | 0.9229 |
| M4 | 0.0041 | 0.0406 | 0.9277 | 0.0039 | 0.0438 | 0.9143 |
| M5 | 0.0037 | 0.0391 | 0.9349 | 0.0035 | 0.0430 | 0.9203 |
| M6 | 0.0039 | 0.0401 | 0.9318 | 0.0035 | 0.0429 | 0.9162 |

training and testing phase in Jehlum stations, the best values of the MSE, MAE and R were obtained using M3. In the model M3 had the smallest MAE and MSE whereas it had the highest value of the R. For Chenab River training and testing phase, the best value of MSE and MAE were obtained using M5, whereas the best value of R for testing was obtained using M3.

5.2. Fitting Hybrid Models Wavelet LSSVM and Wavelet LR to the Data

Two hybrid wavelet-LSSVM (WLSSVM) model and wavelet-LR (WR) model are obtained by combining two methods, DWT with LSSVM and DWT with LR. Before LSSVM and LR applications, the original time series data were decomposed into periodic components (DWs) by Mallat DWT algorithm [20,26,27]. 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 [26].

$$M = \log(n)$$

where, *n* is length of the time series and *M* is decomposition level. In this study, n = 315 and 552 monthly data are used for Jhelum and Chenab, respectively, which approximately gives M = 3 decomposition levels. Three

decomposition levels are employed in this study, the same as studies employed by [28]. 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. Tables 3(a) and 3(b) show 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 and LR model. Figure 3 shows the structure of the WLSSVM model. Figures 4 and 5 show the original streamflow data time and their Ds, 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 and LR models. The forecasting performances of the wavelet-LSSVM (WLSSVM) models, Linear Regression (LR) and wavlet-regression (WR) are presented in **Ta**-

Table 3. The correlation coefficients between each of sub-time series and original monthly streamflow data.

| (a) | | | | | | | |
|-----------------------|--------------------------|---------------|---------------|---------------|---------------|---------------|-------------------------|
| Discrete | Correlations Jehlum data | | | | | | |
| Wavelet Components | D_{r-1}/Q_r | D_{r-2}/Q_r | D_{r-3}/Q_r | D_{r-4}/Q_r | D_{r-5}/Q_r | D_{r-6}/Q_r | Absolute Correlation |
| D1 | 0.053 | 0.140 | 0.003 | 0.011 | 0.014 | 0.017 | 0.040 |
| D2 | 0.209 | 0.189 | 0.203 | 0.154 | 0.097 | 0.081 | 0.156 |
| D3 | 0.815 | 0.824 | 0.809 | 0.809 | 0.800 | 0.768 | 0.804 |
| A3 | 0.466 | 0.416 | 0.512 | 0.546 | 0.583 | 0.627 | 0.525 |
| | | | (b |)) | | | |
| Discrete | Correlations Chenab data | | | | | | |
| Wavelet Components | D_{r-1}/Q_r | D_{t-2}/Q_t | D_{r-3}/Q_r | D_{r-4}/Q_r | D_{r-5}/Q_r | D_{r-6}/Q_r | Absolute Correlation |
| D1 | 0.110 | 0.055 | 0.016 | 0.015 | 0.047 | 0.040 | 0.047 |
| D2 | 0.338 | 0.353 | 0.285 | 0.172 | 0.074 | 0.173 | 0.233 |
| D3 | 0.835 | 0.846 | 0.875 | 0.902 | 0.903 | 0.867 | 0.871 |
| A3 | 0.251 | 0.261 | 0.281 | 0.304 | 0.338 | 0.394 | 0.305 |







Figure 4. Decomposed wavelet sub-series components (Ds) of streamflow data of Jhelum Station.

bles 4(a) and 4(b), Tables 5(a) and 5(b) and Tables 6(a) and 6(b) respectively, in terms of MSE, MAE and *R* in training and testing periods.

Tables 4(a) and 4(b) show that WLSSVM model has a

significant positive effect on streamflow forecast. As seen from **Table 4(a)**, for the Jhelum station, the MW3 model has the smallest MSE (0.0031) and MAE (0.0413) and the highest R (0.9614) in the training phase. How-



Figure 5. Decomposed wavelet sub-series components (Ds) of streamflow data of Chenab Station.

Table 4. Forecasting performance indicates of WLSSVM for (a) Jhelum River of Pakistan; (b) Chenab River of Pakistan.

| | | | (a) | | | |
|--------|--------|----------|--------|---------|---------|--------|
| Madal | | Training | | Testing | | |
| Widdei | MSE | MAE | R | MSE | MAE | R |
| MW1 | 0.0103 | 0.0813 | 0.8638 | 0.0100 | 0.0786 | 0.8585 |
| MW2 | 0.0054 | 0.0515 | 0.9312 | 0.0046 | 0.0489 | 0.9385 |
| MW3 | 0.0031 | 0.0413 | 0.9614 | 0.0049 | 0.0487 | 0.9326 |
| MW4 | 0.0041 | 0.0271 | 0.9832 | 0.0039 | 0.0505 | 0.9479 |
| MW5 | 0.0071 | 0.0316 | 0.9784 | 0.0038 | 0.0476 | 0.9482 |
| MW6 | 0.0081 | 0.0304 | 0.9777 | 0.0043 | 0.0502 | 0.9409 |
| | | | (b) | | | |
| Madal | | Training | | | Testing | |
| Widdei | MSE | MAE | R | MSE | MAE | R |
| MW1 | 0.0122 | 0.0793 | 0.7640 | 0.0078 | 0.0647 | 0.7803 |
| MW2 | 0.0017 | 0.0297 | 0.9705 | 0.0016 | 0.0294 | 0.9610 |
| MW3 | 0.0018 | 0.0290 | 0.9684 | 0.0013 | 0.0291 | 0.9671 |
| MW4 | 0.0018 | 0.0302 | 0.9698 | 0.0014 | 0.0298 | 0.9672 |
| MW5 | 0.0021 | 0.0224 | 0.9801 | 0.0011 | 0.0248 | 0.9729 |
| MW6 | 0.0012 | 0.0231 | 0.9796 | 0.0012 | 0.0264 | 0.9695 |

| | | | (a) | | | |
|--------|--------|----------|--------|--------|---------|--------|
| Madal | | Training | | | Testing | |
| Widder | MSE | MAE | R | MSE | MAE | R |
| LR1 | 0.1271 | 0.0981 | 0.7736 | 0.1162 | 0.0912 | 0.8826 |
| LR2 | 0.1157 | 0.0850 | 0.8171 | 0.1045 | 0.0806 | 0.8447 |
| LR3 | 0.1122 | 0.0835 | 0.8298 | 0.1058 | 0.0838 | 0.8395 |
| LR4 | 0.1084 | 0.0808 | 0.8418 | 0.1030 | 0.0835 | 0.8496 |
| LR5 | 0.1055 | 0.0779 | 0.8561 | 0.0995 | 0.0798 | 0.8611 |
| LR6 | 0.1048 | 0.0769 | 0.8497 | 0.1018 | 0.0819 | 0.8534 |
| | | | (b) | | | |
| Madal | | Training | | | Testing | |
| Widder | MSE | MAE | R | MSE | MAE | R |
| LR1 | 0.1165 | 0.0854 | 0.7317 | 0.0918 | 0.0732 | 0.7549 |
| LR2 | 0.0980 | 0.0699 | 0.8192 | 0.0727 | 0.0581 | 0.8580 |
| LR3 | 0.0971 | 0.0703 | 0.8231 | 0.0729 | 0.0590 | 0.8571 |
| LR4 | 0.0963 | 0.0687 | 0.8267 | 0.0722 | 0.0578 | 0.8613 |
| LR5 | 0.0924 | 0.0654 | 0.8420 | 0.0707 | 0.0555 | 0.8713 |
| LR6 | 0.0893 | 0.0630 | 0.8536 | 0.0766 | 0.0569 | 0.8767 |

Table 5. Forecasting performances indicates of LR for (a) Jhelum River of Pakistan; (b) Chenab River of Pakistan.

Table 6. Forecasting performances indicates of WR for (a) Jhelum River of Pakistan; (b) Chenab River of Pakistan.

| | | | (a) | | | |
|--------|--------|----------|--------|--------|---------|--------|
| Madal | | Training | | | Testing | |
| Model | MSE | MAE | R | MSE | MAE | R |
| WR1 | 0.1125 | 0.0914 | 0.8281 | 0.1023 | 0.0781 | 0.8505 |
| WR2 | 0.0839 | 0.0644 | 0.9089 | 0.0809 | 0.631 | 0.9089 |
| WR3 | 0.0736 | 0.0525 | 0.9305 | 0.0801 | 0.0626 | 0.9105 |
| WR4 | 0.0690 | 0.0465 | 0.9391 | 0.0882 | 0.0586 | 0.8935 |
| WR5 | 0.0706 | 0.0486 | 0.9358 | 0.0768 | 0.0404 | 0.9181 |
| WR6 | 0.0677 | 0.0440 | 0.9402 | 0.0917 | 0.0561 | 0.8848 |
| | | | (b) | | | |
| Madal | | Training | | | Testing | |
| widder | MSE | MAE | R | MSE | MAE | R |
| WR1 | 0.1124 | 0.0836 | 0.7532 | 0.0881 | 0.0685 | 0.7768 |
| WR2 | 0.0666 | 0.0533 | 0.9209 | 0.0536 | 0.0435 | 0.9248 |
| WR3 | 0.0577 | 0.0449 | 0.9413 | 0.0449 | 0.0356 | 0.9468 |
| WR4 | 0.0525 | 0.0394 | 0.9519 | 0.0372 | 0.0307 | 0.9677 |
| WR5 | 0.0461 | 0.0345 | 0.9632 | 0.0356 | 0.0280 | 0.9682 |
| WR6 | 0.0423 | 0.0308 | 0.9690 | 0.0409 | 0.0281 | 0.9607 |

ever, for the testing phase, the best MSE (0.0038), MAE (0.0476) and R (0.9482) was obtained for the model input combination MW5. From **Tables 5(a)** and **5(b)** for Chenab station, the MW6 model has the smallest MSE

(0.0012) and MAE (0.0231) and the highest R (0.9796) in the training phase. However, for the testing phase, the best MSE (0.0011) and MAE (0.0248) and R (0.9729) was obtained for the model input combination MW5.

5.3. Comparisons of Forecasting Models

For further analysis, the best performances of the LSSVM, LR, WLSSVM and WR models in terms of the MSE, MAE and R at testing phase are compared.

In **Tables 7(a)** and **7(b)**, it shows that WLSSVM has good performance during the testing phase, and they outperform LSSVM, LR and WR in terms of all the standard statistical measures. The correlation coefficient (R) for Jehulm River and Chenab River data obtained by LSSVM models is 0.8943 and 0.9203 and by WR models is 0.9181 and 0.9682 respectively, with WLSSVM model, the R value is increased to 0.9482 and 0.9729. The MSE obtained by LSSVM models is 0.0075 and 0.0035 for both data sets respectively with WLSSVM model this value is decreased to 0.0038 and 0.0011. Similarly, while the MAE obtained by LSSVM is 0.0662 and 0.0430, the MAE value of WLSSVM model is decreased to 0.0476 and 0.0248. The WLSSVM model obtained the best value of MSE and MAE decrease 49% and 69%, respectively, and the best of R increases by 6% compared with single LSSVM model for Jhulem data. For Chenab data the best of R increases by 6% and the best value obtained for MSE and MAE decreases 28% and 42%.

Figures 6 and **7** show the hydrograph and scatter plot for the LSSVM, WLSSVM, WR and LR models for the testing period. It can be seen that the WLSSVM forecasts quite close to the observed data for both station.

The performance of WLSSVM in predicting the streamflow is superior to the classical LSSVM model. As seen from the fit line equations (assume that the equation is y = a + bx) in the scatterplots that a and b coefficients for the LSSVM, WLSSVM, WR and LR models, respectively, the WLSSVM has less scattered estimates and the R value of WLSSVM model close to 1 (R = 0.9482 and R = 0.9729) compared to the LSSVM, WR and LR models for both data sets respectively. Overall, it can be

| Table 7. The | performance result | s LSSVM. V | WLSSVM. | LR and WR | approach duri | ng testing period |
|---------------|--------------------|------------|---------|-------------|---------------|-------------------|
| Lable /. Life | perior mance resul | | | Lix and wix | approach uur | ng woung periou |

| (a) | | | | | | | | | |
|--------------|-----------------------|-----------------------|-----------------------|--|--|--|--|--|--|
| Jehlum River | | | | | | | | | |
| Model | MSE | MAE | R | | | | | | |
| LSSVM | 0.0075 (49%) | 0.0662 (28%) | 0.8943 (6%) | | | | | | |
| WLSSVM | 0.0038 | 0.0476 | 0.9482 | | | | | | |
| LR | 0.0995 (96%) | 0.0798 (52%) | 0.8611 (10%) | | | | | | |
| WR | 0.0768 (95%) | 0.0404 (38%) | 0.9181 (3%) | | | | | | |
| | (b) | | | | | | | | |
| | Chenab | River | | | | | | | |
| Model | MSE | MAE | R | | | | | | |
| LSSVM | 0.0035 (69%) | 0.0430 (42%) | 0.9203 (6%) | | | | | | |
| WLSSVM | 0.0011 | 0.0248 | 0.9729 | | | | | | |
| LR | 0.0707 (98%) | 0.0555 (55%) | 0.8773 (11%) | | | | | | |
| WR | 0.0356 (97%) | 0.0280 (11%) | 0.9682 (1%) | | | | | | |



Figure 6. Predicted and observed streamflow during testing period by WLSSVM, LSSVM, LR and WR for Jhelum Station.



Figure 7. Predicted and observed streamflow during testing period by LSSVM and WLSSVM for Chenab Station.

concluded the WLLSVM model at both studies provided more accurate forecasting results than the LSSVM, WR and LR models for streamflow forecasting.

6. Conclusions

The new method based on the WLSSVM is developed by combining the discrete wavelet transforms (DWT) and least square support vector machines (LSSVM) model for forecasting streamflows. The monthly streamflow time series is decomposed at different decomposition levels by DWT. Each of the decompositions carries most of the information and plays a distinct role in original time series. The correlation coefficients between each of the 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 are decomposed at 3 decomposition levels (2 - 4 - 8 months). The sum of effective details and the approximation component were used as inputs to the LSSVM model. The WLSSVM models are trained and tested by applying different input combinations of monthly streamflow data of Chanari station in Jhelum River and Marala station in Chenab in Punjab of Pakistan. Then, LSSVM model is constructed with new series as inputs and original streamflow time series as output. The performance of the proposed WLSSVM model is then compared to the regular LSSVM model for monthly streamflow forecasting.

Comparison results carried out in the study indicated that the WLSSVM model was substantially more accurate than LSSVM, LR and WR models. The study concludes that the forecasting abilities of the LSSVM model are 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 models.

REFERENCES

- H. E. Hurst, "Long Term Storage Capacity of Reservoirs," *Transactions of ASCE*, Vol. 116, 1961, pp. 770-799.
- [2] N. C. Matalas, "Mathematical Assessment of Symmetric Hydrology," *Water Resources Research*, Vol. 3, No. 4, 1967, pp. 937-945. <u>doi:10.1029/WR003i004p00937</u>
- [3] G. E. P. Box and G. M. Jenkins, "Time Series Analysis Forecasting and Control," Holden Day, San Francisco, 1970.
- [4] J. W. Delleur, P. C. Tao and M. L. Kavvas, "An Evaluation of the Practicality and Complexity of Some Rainfall and Runoff Time Series Model," *Water Resources Research*, Vol. 12, No. 5, 1976, pp. 953-970. doi:10.1029/WR012i005p00953
- [5] V. Vapnik, "The Nature of Statistical Learning Theory," Springer Verlag, Berlin, 1995. doi:10.1007/978-1-4757-2440-0
- [6] P. S. Yu, S. T. Chen and I. F. Chang, "Support Vector Regression for Real-Time Flood Stage Forecasting," *Journal of Hydrology*, Vol. 328, No. 3-4, 2006, pp. 704-716. doi:10.1016/j.jhydrol.2006.01.021
- [7] Y. B. Dibike, S. Velickov, D. P. Solomatine and M. B. Abbott, "Model Induction with Support Vector Machines: Introduction and Applications," *Journal of Computing in Civil Engineering*, Vol. 15, No. 3, 2001, pp. 208-216. doi:10.1061/(ASCE)0887-3801(2001)15:3(208)
- [8] A. Elshorbagy, G. Corzo, S. Srinivasulu and D. P. Solomatine, "Experimental Investigation of the Predictive Capabilities of Data Driven Modeling Techniques in Hydrology, Part 1: Concepts and Methodology," *Hydrology and Earth System Sciences Discussions*, Vol. 6, 2009, pp. 7055-7093.
- [9] A. Elshorbagy, G. Corzo, S. Srinivasulu and D. P. Solomatine, "Experimental Investigation of the Predictive Capabilities of Data Driven Modeling Techniques in Hydro-

logy, Part2: Application," *Hydrology and Earth System Sciences Discussions*, Vol. 6, 2009, pp. 7095-7142.

- [10] T. Asefa, M. Kemblowski, M. McKee and A. Khalil, "Multi-Time Scale Stream Flow Predictions: The Support Vector Machines Approach," *Journal of Hydrology*, Vol. 318, No. 1-4, 2006, pp. 7-16.
- [11] J. Y. Lin, C. T. Cheng and K. W. Chau, "Using Support Vector Machines for Long-Term Discharge Prediction," *Hydrological Sciences Journal*, Vol. 51, No. 4, 2006, pp. 599-612. doi:10.1623/hysj.51.4.599
- [12] W. C. Wang, K. W. Chau, C. T. Cheng and L. Qiu, "A Comparison of Performance of Several Artificial Intelligence Methods for Forecasting Monthly Discharge Time Series," *Journal of Hydrology*, Vol. 374, No. 3-4, 2009, pp. 294-306. doi:10.1016/j.jhydrol.2009.06.019
- [13] J. A. K. Suykens and J. Vandewalle, "Least Squares Support Vector Machine Classifiers," *Neural Processing Letters*, Vol. 9, No. 3, 1999, pp. 293-300. doi:10.1023/A:1018628609742
- [14] D. Hanbay, "An Expert System Based on Least Square Support Vector Machines for Diagnosis of Valvular Heart Disease," *Expert Systems with Applications*, Vol. 36, No. 4, 2009, pp. 8368-8374.
- [15] Y. W. Kang, J. Li, C. Y. Guang, H.-Y. Tu, J. Li and J. Yang, "Dynamic Temperature Modeling of an SOFC Using Least Square Support Vector Machines," *Journal of Power Sources*, Vol. 179, No. 2, 2008, pp. 683-692. doi:10.1016/j.jpowsour.2008.01.022
- [16] B. Krishna, Y. R. Satyaji Rao and P. C. Nayak, "Time Series Modeling of River Flow Using Wavelet Neutral Networks," *Journal of Water Resources and Protection*, Vol. 3, No. 1, 2011, pp. 50-59. doi:10.4236/jwarp.2011.31006
- [17] L. C. Simith, D. L. Turcotte and B. Isacks, "Streamflow Characterization and Feature Detection Using a Discrete Wavelet Transform," *Hydrological Processes*, Vol. 12, No. 2, 1998, pp. 233-249.
 <u>doi:10.1002/(SICI)1099-1085(199802)12:2<233::AID-H</u> YP573>3.0.CO;2-3
- [18] D. Wang and J. Ding, "Wavelet Network Model and Its Application to the Prediction of Hydrology," *Nature and Science*, Vol. 1, No. 1, 2003, pp. 67-71.
- [19] D. Labat, R. Ababou and A. Mangin, "Rainfall-Runoff Relations for Karastic Springs: Part II. Continuous Wavelet and Discrete Orthogonal Multiresolution Analysis," *Journal of Hydrology*, Vol. 238, No. 3-4, pp. 2000, pp. 149-178. doi:10.1016/S0022-1694(00)00322-X
- [20] S. G. Mallat, "A Theory for Multi Resolution Signal Decomposition: The Wavelet Representation," *IEEE Trans-*

actions on Pattern Analysis and Machine Intelligence, Vol. 11, No. 7, 1998, pp. 674-693.

- [21] A. Grossman and J. Morlet, "Decomposition of Harley Functions into Square Integral Wavelets of Constant Shape," *SIAM Journal on Mathematical Analysis*, Vol. 15, No. 4, 1984, pp. 723-736. doi:10.1137/0515056
- [22] I. Daubechies, "Orthogonal Bases of Compactly Supported Wavelets," *Communications on Pure and Applied Mathematics*, Vol. 41, No. 7, 1988, pp. 909-996. <u>doi:10.1002/cpa.3160410705</u>
- [23] E. Foufoula-Georgiou and P. E Kumar, "Wavelets in Geophysics," Academic, San Diego and London, 1994.
- [24] R. M. Rao and A. S. Bopardikar, "Wavelet Transforms: Introduction to Theory and Applications," Addison Wesley Longman, Inc., Reading, 1998, 310 p.
- [25] M. Kucuk and N. Agiralioğlu, "Wavelet Regression Technique for Streamflow Prediction," *Journal of Applied Statistics*, Vol. 33, No. 9, 2006, pp. 943-960. doi:10.1080/02664760600744298
- [26] O. Kisi, "Wavelet Regression Model as an Alternative to Neural Networks for Monthly Streamflow Forecasting," *Hydrological Processes*, Vol. 23, No. 25, 2009, pp. 3583-3597. doi:10.1002/hyp.7461
- [27] O. Kisi, "Wavelet Regression Model for Short-Term Streamflow Forecasting," *Journal of Hydrology*, Vol. 389, No. 3-4, 2010, pp. 344-353. doi:10.1016/j.jhydrol.2010.06.013
- [28] P. Y. Ma, "A Fresh Engineering Approach for the Forecast of Financial Index Volatility and Hedging Strategies," PhD thesis, Quebec University, Montreal, 2006.
- [29] M. Firat, "Comparison of Artificial Intelligence Techniques for River Flow Forecasting," *Hydrology and Earth System Sciences*, Vol. 12, No. 1, 2008, pp. 123-139.
- [30] R. Samsudin, S. Ismail and A. Shabri, "A Hybrid Model of Self-Organizing Maps (SOM) and Least Square Support Vector Machine (LSSVM) for Time-Series Forecasting," *Expert Systems with Applications*, Vol. 38, No. 8, 2011, pp. 10574-10578.
- [31] M. T. Gencoglu and M. Uyar, "Prediction of Flashover Voltage of Insulators Using Least Square Support Vector Machines," *Expert Systems with Applications*, Vol. 36, No. 7, 2009, pp. 10789-10798. doi:10.1016/j.eswa.2009.02.021
- [32] L. Liu and W. Wang, "Exchange Rates Forecasting with Least Squares Support Vector Machines," *International Conference on Computer Science and Software Engineering*, Wuhan, 12-14 December 2008, pp. 1017-1019.