ARTICLE

Wavelet Decomposition Impacts on Traditional Forecasting Time Series Models

W. A. Shaikh[1,2,*], S. F. Shah[2], S. M. Pandhiani[3] and M. A. Solangi[2]

[1]Department of Mathematics and Statistics, Quaid-e-Awam University of Engineering, Science & Technology, Nawabshah, Sindh, Pakistan

[2]Department of Basic Sciences & Related Studies, Mehran University of Engineering & Technology, Jamshoro, Sindh, Pakistan

[3]Department of General Studies, Jubail University College, Al Jubail, Saudi Arabia

[*]Corresponding Author: W. A. Shaikh. Email: wajid@quest.edu.pk

ABSTRACT

This investigative study is focused on the impact of wavelet on traditional forecasting time-series models, which significantly shows the usage of wavelet algorithms. Wavelet Decomposition (WD) algorithm has been combined with various traditional forecasting time-series models, such as Least Square Support Vector Machine (LSSVM), Artificial Neural Network (ANN) and Multivariate Adaptive Regression Splines (MARS) and their effects are examined in terms of the statistical estimations. The WD has been used as a mathematical application in traditional forecast modelling to collect periodically measured parameters, which has yielded tremendous constructive outcomes. Further, it is observed that the wavelet combined models are classy compared to the various time series models in terms of performance basis. Therefore, combining wavelet forecasting models has yielded much better results.

KEYWORDS

Impact; wavelet decomposition; combined; traditional forecasting models; statistical analysis

1 Introduction

Due to the predictive importance, researchers have developed various forecasting models. As better environmental forecasting arts can be used to make appropriate management decisions, researchers are continually striving to improve the effectiveness and efficiency of the models. For decades the term wavelet has been used for the exploration of signal processing and geophysics. Therefore, this article looks at the WD combined with various traditional forecast time-series models. For decades the term wavelet has been used for the exploration of signal processing and geophysics. The last decade has shown vast interest in wavelets; it is a subject area that can be appropriate applicable and coalesced in various fields such as applied mathematics, physics, electrical engineering, etc.

Consequently, the WD has significantly impacted various fields, such as image processing, differential equations, statistics, and chemical signal processing [1,2]. WD algorithm is used as a mathematical approach for extracting nontrivial and potentially useful information from different data types such as historical data information, re-analysis, local or global climate model simulations, etc. Usually, WD based models are required to analyse the datasets and decomposing them thoroughly. Its hypothesis is to get better re-construct datasets with minimal losses. The wavelet transform technique uses the signal frequency from the time domain to the wavelet domain. Therefore, the newly acquired domain has contained more complex basis functions called wavelets, mother wavelets, or analysing wavelets [3].

Wavelet-based models are a noteworthy edge in de-noising the datasets to develop an efficient model. It has made it easy to analyse streamflow processes on different parameters without eliminating the effects of the time-frequency accompanied by conventional bandpass filters. The WD tool can let on information within the signal in both frequencies, time and scale domains [4,5]. The WD application controls the time-frequency or signal scale content and judges the temporal variation spectrum [6]. In contrast, the Fourier transforms interpret a quite different perspective that allows estimating the signal frequency but is not suitable to estimate the time-frequency dependence. Therefore, the wavelet transform has its origins in the Fourier transform. Thus, the WD has allowed for tracking the time evolution of processes at various scales in the signal as it has both time and scale measures of localisation. The WD tool signals can be classified as high limited frequency events or a significant number of scale-variable methods because it provides explicit information for rate forecasting classification [7]. In a review of the applications of the wavelet transform in hydrological dataset modelling describing the multifaceted information that can be obtained from such an analysis and recognition of seasonality, streamflow trends, and data de-noising.

The performance and accuracy of the traditional time series forecasting models continuously may be improved. Therefore, it can be inspired by the researchers to intend an improved version of the models [8]. This study describes the performance impact of the WD as an optimization in the traditional time series models in which the optimal response is continuously exchanged during the simulation and can be approved. The effectiveness of the models has been tested on two different streamflow datasets, including Indus and Chenab Rivers.

2 Study Areas and Data Utilization

To endorse the discussed TS forecast models and forecast the rivers, streamflow of the rate of the rivers have been collected 484 and 550 months, respectively, from two renowned Indus and Chenab Rivers of Pakistan (Figs. 1–2).
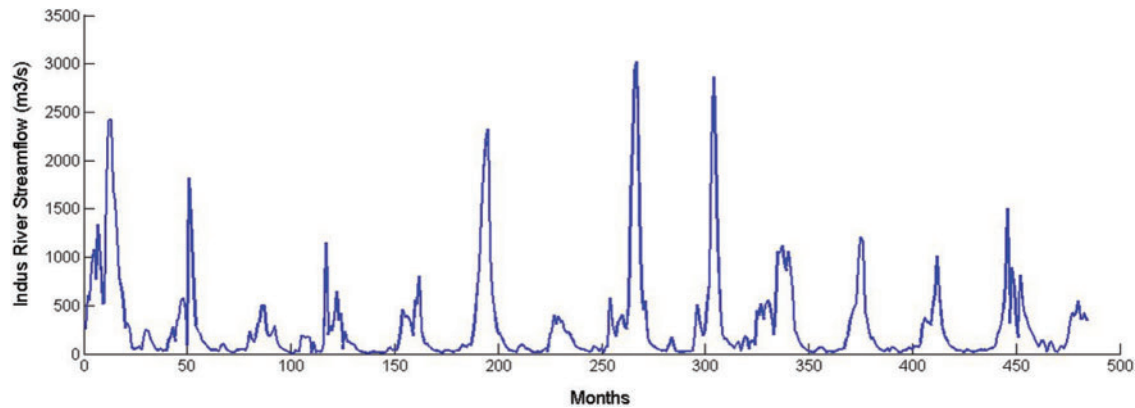
Figure 1: 484 months streamflow rate of Indus River
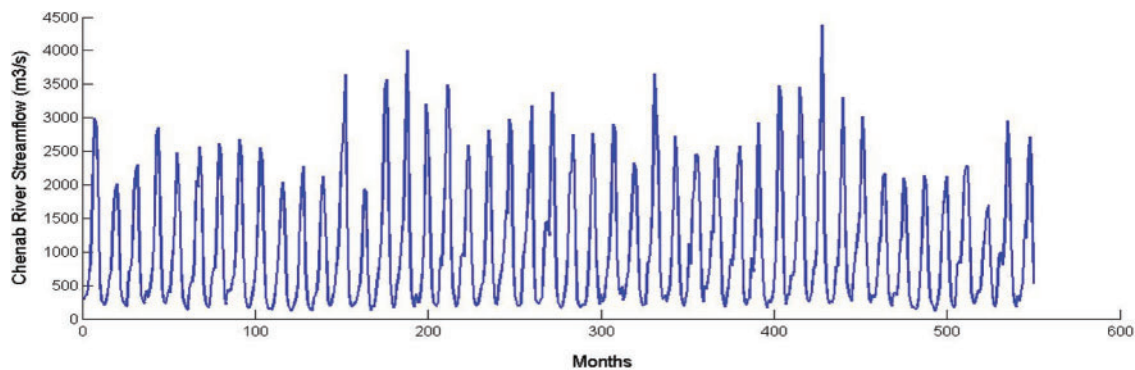
Figure 2: 550 months streamflow rate of Chenab River

3 Methodology

The use of wavelet application with the various traditional forecasting models such as LSSVM, ANN and MARS has improved the efficiency of the models and found excellent outcomes. These tractable combined models have been implemented as efficient tools on streamflow datasets to forecast phenomena that provide comprehensive signals information. The developed combined wavelet with AI models implements the following two-step protocol for forecasting activities.

(1) The WD methodology has been used as a preprocessor of input datasets. As a result, it has a time-frequency signal analysis at distinct intervals in the time-domain and considerable detail about input datasets.

(2) After obtaining the input signal by WD, it has been used for further processes as AI input in various traditional forecasting models.

Initially, the forecasting time-series datasets have been decomposed into a sub-time-series $\{W_1, W_2, \cdots, W_p, C_p\}$ by the WD algorithm, where, $W_1, W_2, \cdots, W_p$ and C_p have been described as detailed time-series and background time-series, respectively. Commonly, these have different roles in the forecasting time-series datasets and each sub-time-series have different behaviour. Therefore, the attribute and influence on the forecasting time-series dataset have different from each other. The developed LSSVM, ANN and MARS forecast models in which t are the input

sub-time-series of the models, and the forecasting time-series datasets at $t+T$ time have the output of models. Where T is described as the measurement of forecast time-frequency. Finally, wavelet network models (WNM) were created, in which certain weights have been learned using specific techniques [9]. The key object of forecasting time-series datasets to the WNM algorithm is to construct combined models with the WD algorithm.

3.1 Wavelet Decomposition (WD)

The algorithm of WD ability to de-noise non-stationary signals into sub-signals at different levels has a suitable resource for improved streamflow elucidation [10]. The novel developed forecasting models have been individual capacities that comprised WD and some conventional AI model techniques. The combined wavelet with various AI forecasting models has been an appropriate methodology [11]. According to forecasting time series models, the wavelet technique is becoming an ever more effective and essential tool used in models. The primary reason for WD has analysed the time series datasets in the time and frequency domain of the valuable decomposition of the original time series by taking useful information in various frequency levels using wavelet functions. The main advantage of using the WD is its robustness, as it does not include any potentially erroneous assumption or parametric testing procedure. The following WD mathematical structure is defined as a continuous time-series $x(t)$, $t \in [-\infty, \infty]$:

$$\psi(t, s) = \frac{1}{\sqrt{s}} \psi\left(\frac{t - \tau}{2}\right) \tag{1}$$

where defined t, τ and $s \in [0, \infty]$ for time, for time step that iterates the window function and wavelet scale respectively, whereas $\psi(t)$ is known as mother wavelet, it can be defined as follows:

$$\int_{-\infty}^{\infty} \psi(t)dt = 0 \tag{2}$$

Therefore, the continuous wavelet transform (CWT) can be defined as follows [12];

$$W(\tau, s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} x(t)\bar{\psi}\left(\frac{t - \tau}{s}\right) dt \tag{3}$$

where, $\bar{\psi}(t)$ and $W(\tau, s)$ defined for complex conjugation of $\psi(t)$ and presents the sum overall time of the time series multiplied by scale and shifted version of wavelet function $\psi(t)$, respectively. The usage of WD for forecasting is not feasible since it is time-consuming to measure the wavelet coefficient on any conceivable scale and produces many results. WD is preferred in most forecasting problems because of its simplicity and ability to compute with less time. The WD involves choosing scales and positions on powers of 2, so-called dyadic scales and translations; then, the analysis is much more efficient as well as more accurate. The main advantage of using the WD is its robustness, as it does not include any potentially erroneous assumption or parametric testing procedure. The following mathematical structure is known as WD:

$$\psi_{m,n}\left(\frac{t - \tau}{s}\right) = \frac{1}{\sqrt{s_0^{m/2}}} \psi\left(\frac{t - n\tau_0 s_0^m}{s_0^m}\right) \tag{4}$$

where $m, n \in Z$ that control the scale and length of the time-series datasets respectively, s_0 is stated, fixed dilation step greater than 1, τ_0 is the location parameter, which must be greater than zero. The parameters $s_0 = 2$ and $\tau_0 = 1$ are the most common preference. For a discrete time-series $x(t)$ that occur at t discrete time, then WD formation is defined as follows:

$$W_{m,n} = 2^{-m/2} \sum_{t=0}^{N-1} \psi(2^{-m}t - n)x(t) \tag{5}$$

where, $W_{m,n}$ is the wavelet coefficient of the WD at defined scale $s = 2^m$ and $\tau = 2^m n$, $x(t)$ is discrete time-series (i.e., $t = 1, 2, \ldots, N-1$) and $N \in Z$ to the power of $2 (N = 2^m)$; n is the time translation parameter that changes in the ranges $0 < n < 2M - m$ and $1 < m < M$, whereas M has defined for the decomposition level. The analysed datasets have decomposed into several wavelet input components that depend on the particular decomposition level. Decisive the advantage-able decomposition level of the datasets in WD plays a mark-able role in conserved the figures and reduces the curvature of the datasets. However, there is no prevailing theory to inform how many decomposition levels have required for any time series dataset. By the following formula is estimated the decomposition level:

$$M = \log(n) \tag{6}$$

Decomposition levels $M = 2.6848$ and $M = 2.7403$ estimated when $n = 484$ and $n = 550$ monthly streamflow datasets (Figs. 1–2) have been used for Indus and Chenab Rivers, respectively, and approximately $M = 3$ round-off decomposition level has used for both said rivers in this research work. By the hypothesis of Mallat's, the original discrete time-series $x(t)$ decomposed into a series of linearity independent estimation and detail signals by the inverse WD and defined as follows:

$$x(t) = T + \sum_{m=1}^{M} \sum_{t=0}^{2^{M-m-1}} W_{m,n} 2^{-m/2} \psi(2^{-m}t - n) \tag{7}$$

or inverse WD also in a simple format defined as follows:

$$x(t) = A_M(t) + \sum_{m=1}^{M} D_m(t) \tag{8}$$

where, $A_M(t)$ is known as the residual term at levels M and $D_m(t) (m = 1, 2, \ldots, M)$ are the detailed sub-series of information capable of catching specific functions of data interpretation value. The following various forms of transfer functions (TF) have been emerged and vastly used in the model [13]:

$$\left.\begin{array}{ll} \textit{Polynomial Function} & y = z \\ \textit{Sigmoid Function} & y = \frac{1}{1+\exp(-z)} \\ \textit{Radial Basis Function} & y = \exp(-z^2) \\ \textit{Hyperbolic Tangent Function} & y = \tanh(z) = \frac{2}{1+\exp(-2z)} - 1 \end{array}\right\} \tag{9}$$

3.2 Wavelet Artificial Neural Networks (WANN)

The human brain's functioning principle influenced an artificial neural network (ANN) as a forecasting model. Several architectures in the literature are available to forecast the streamflow and many other applications, one of which is the ANN algorithm mostly used. It is comprised of a network system with many interconnected nodes called neurons. The number of layers of an ANN is used to classify it, and layer(s) exists between an input and an output layer. Therefore, a single-layer feed-forward (SLFF) neural network is an architecture with just one layer for establishing connection among the nodes of the input, middle, and output layers. This type of system is characterised as a multi-layer feed-forward (MLFF) neural network built by more than one middle layer [14]. ANN has the merits of fault-tolerance, an efficient nonlinear developing capability that provides a well-organised model for streamflow forecasting. Though the model has some drawbacks, such as slow optimization processing, model complexities, and the approximation of the applications could not be overlooked. Therefore, this forecasting architectural model needed to be improved. The WD has been compounded to the ANN model for this purpose. The WANN model has been combined with the strengths of the WD and ANN applications and to achieve nonlinear strong approximation capability. WANN model architectural design has based on multi-layer perceptron (MLP). WANN is quite corresponding to that of $(1 + 1/2)$ layer neural network and has contained the following three layers [15]:

Layer-i. input layer, which is introduced to the network and takes one or more than one inputs.

Layer-ii. the hidden layer, where data is manipulating procedure with a feed-forward neural network accompanied with orthonormal WD basis by activation functions developed.

Layer-iii. output layer, which contains one or more linear combiners and the corresponding estimations are consistent with the given inputs.

The training process has acquired the weights of the network connections. The WANN model can have various TF of different nodes in identical or different layers. The TF such as sigmoid, hyperbolic tangent functions are used for hidden layers, and there is no appropriation for the output layer. The WANN model has been successfully used for forecasting estimations. Two key approaches to developing the WANN model technique are described as following:

- The WD technique and the ANN model processing are used separately. Firstly decomposed, the input signal employs various WD basis functions (Eq. (1)) with neurons in the hidden layer. Then, the wavelet coefficients have one or more output activities in their input weights have adjusted according to the certain learning algorithm.
- In this case, two structures, WD mathematical and ANN artificial intelligence algorithms have been combined and performed. The transferal and dilation of the WD accomplished weights that have been adjusted according to a certain learning algorithm.

Only dyadic dilations and translations of the WD have developed the wavelet basis function whenever the first approach occurs. Therefore, this objective approach of WANN has often been known as a wavenet.

3.3 Wavelet Least-Squares Support Vector Machine (WLSSVM)

The following reprocessed support vector machine (SVM) classifier governs the application of minimisation [16]

$$Min\ J(\omega,\ b,\ e) = \frac{\mu}{2}\omega^T\omega + \frac{\zeta}{2}\sum_{i=1}^{N}e_i^2 \tag{10}$$

Under the following equality condition

$$y_i[\omega^T\phi(x_i) + b] = 1 - e_i \qquad (i = 1,\ 2,\ \ldots,\ N) \tag{11}$$

The usage of the LSSVM classifier is implicitly compatible with the definition of regression using binary conditions $y_i = \pm 1 \ \Rightarrow y_i^2 = 1$ and obtained

$$\sum_{i=1}^{N}e_i^2 = \sum_{i=1}^{N}(y_ie_i)^2 = \sum_{i=1}^{N}[y_i - (\omega^T\phi(x_i) + b)]^2 \tag{12}$$

Therefore, developing a sense for least square (LS) data fitting is equivalent to LSSVM classifier development

$$J(\omega,\ b,\ e) = \mu E_w + \zeta E_D$$
$$\left(\because E_w = \frac{1}{2}\omega^T\omega \quad and \quad E_D = \frac{1}{2}\sum_{i=1}^{N}[y_i - (\omega^T\phi(x_i) + b)]^2\right) \tag{13}$$

where, μ and ζ are known to be hyper-parameters to modify the regularisation number *vs.* the sum square error. The original formulation ratio $\gamma = \zeta/\mu$ has provided the solution as a tuning parameter. Apply parameters μ and ζ to LSSVM using a Bayesian description. The solution of the LSSVM model has been obtained after the development of the following Lagrangian function:

$$\{L(\omega,\ b,\ e,\ \alpha) = J(\omega,\ e) - \sum_{i=1}^{N}\alpha_i\left[\left\{\omega^T\phi(x_i) + b\right\} + e_i - y_i\right] = \frac{\omega^T\omega}{2} + \frac{\gamma}{2}\sum_{i=1}^{N}e_i^2$$
$$- \sum_{i=1}^{N}\alpha_i\left[\left\{\omega^T\phi(x_i) + b\right\} + e_i - y_i\right] \tag{14}$$

Since α_i are Lagrange multipliers $\forall\alpha_i \in \mathbb{R}$. For the LSSVM model, the following structures are appropriate:

$$\left. \begin{array}{l} \dfrac{\partial L}{\partial\omega} = 0 \quad \rightarrow \quad \omega = \sum_{i=1}^{N}\alpha_i\phi(x_i); \\[2ex] \dfrac{\partial L}{\partial b} = 0 \quad \rightarrow \quad \sum_{i=1}^{N}\alpha_i = 0; \\[2ex] \dfrac{\partial L}{\partial e_i} = 0 \quad \rightarrow \quad \alpha_i = \gamma e_i; \quad (\because i = 1, 2, \cdots, N) \\[2ex] \dfrac{\partial L}{\partial\alpha_i} = 0 \quad \rightarrow \quad y_i = \omega^T\phi(x_i) + b + e_i; \quad (\because i = 1, 2, \cdots, N) \end{array} \right\} \tag{15}$$

After eliminating ω and e, in place of quadratic programming (QP) problem, following linear system obtained:

$$\begin{bmatrix} 0 & 1_N^T \\ 1_N & \Omega + \gamma^{-1}I_N \end{bmatrix}\begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} o \\ Y \end{bmatrix} \tag{16}$$

Since $1_N = [1, \ldots, 1]^T$. I_N is N^{th} order identity matrix; $\Omega \in \mathbb{R}^{N \times N}$ is kernel matrix and defined as $\Omega_{ij} = K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$; and $Y = [y_1, \ldots, y_N]$.

Here, pick $K(x_i, x_j) = e^{\frac{-||x-x_i||^2}{\sigma^2}}$ as a kernel function, $\sigma \in R^+$ is a scale parameter, and evaluate the inputs scaling in the RBF kernel [17,18].

The wavelet least-squares support vector machine (WLSSVM) model has been developed with the potential worth of the WD algorithm and LSSVM processing and obtained optimum nonlinear approximation ability. WLSSVM model has been consists of an input layer, hidden layer, and output layer and the model successfully has been used for forecasting approximations [16].

3.4 Multivariate Adaptive Regression Splines (MARS)

The MARS model schemes discoveries to forecasting continuous numeric outcomes. Appropriate, the MARS model scheme has been implemented in two stages containing forward-backward stepwise techniques. The stepwise forward technique has observed a large set of input variables (basis function) with a different knot; though, this stepwise technique might be developing complexity and a multi-layered model [19]. Determination of such type of model has weak forecasting presentation. For increasing forecast accuracy, the backward stepwise technique eradicates the pointless variables amongst the before chosen set, which may have fewer effects on the approximation procedure pruned by the MARS. For the projection of x the input variable to a novel y output variable based on appropriation, named basis functions that defines point of inflection along with the input range [20]:

$$y = \begin{cases} \max(0, x - c) \\ \max(0, c - x) \end{cases} \tag{17}$$

In these y functions, x treat as input and c chosen is a threshold value is said to be a knot. The function is useful in forward-backward stepwise techniques for each input unknown to classify the position of knots, where the value of the function changes. These y basis functions are called Spline functions, which is a $c - knot$ reflected pair. The following description is the standard form of the MARS model [21].

$$y = f(x) = c_0 + \sum_{i=1}^{M} c_i B_i(x) \tag{18}$$

where the output variable y estimated by the MARS model, c_0 is constant, c_i is the ith basis function coefficient determined by minimising the Root Mean Squared Errors (RMSE), and $B_i(x)$

is the ith basis function. The optimal MARS model scheme is designated based on the smallest value of the Generalised Cross-Validation (GCV) principle. The GCV is defined as follows [17]:

$$GCV(M) = \frac{\sum_{i=1}^{n} [y_i - f(x_i)]^2}{n\left(1 - \frac{C(M)}{n}\right)^2} \tag{19}$$

where, y_i is the objective of output, $f(x_i)$ is the projected output, n is the number of inputs, and $C(M)$ is a penalty, expressed as follows:

$$C(M) = d \times M + M + 1 \tag{20}$$

where d is the penalty for each basis function consisted in the model, M is the number of basis functions. Wavelet Multivariate Adaptive Regression Splines (WMARS) model has combined the prospective techniques of WD and MARS for achieving robust nonlinear estimation potentiality. WMARS model binding designed is found on a multi-layer perception (MLP). The WMARS model is used for forecast estimations and has been dependent on an input layer, hidden layer, and output layer [22].

3.5 Statistical Parameters Assessment

The statistical parameters are used to demonstrate the effectiveness in terms of forecastability of the models assessed by comparing the actual and forecasted values. Usually, the Mean Absolute Error (MAE), the Root Mean Square Error (RMSE), and the correlation coefficients (CC), are used to determine the efficiency of the models and outcomes fitted to the best fit line [23,24]

$$MAE = \frac{1}{m} \sum_{t=1}^{m} |y_t - \hat{y}_t| \tag{21}$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{t=1}^{m} (y_t - \hat{y}_t)^2} \tag{22}$$

$$CC = \frac{\frac{1}{m} \sum_{t=1}^{m} (y_t - \bar{y}_t)(\hat{y}_t - \bar{\hat{y}}_t)}{\sqrt{\frac{1}{m} \sum_{t=1}^{m} (y_t - \bar{y}_t)^2} \sqrt{\frac{1}{m} \sum_{t=1}^{m} (\hat{y}_t - \bar{\hat{y}}_t)^2}} \tag{23}$$

where y_t, $\hat{y}_t$, $\bar{y}_t$ and $\bar{\hat{y}}_t$ are the actual (streamflow observed), forecasted, average streamflow, and average forecast values at a time t, respectively, whereas m is the total number of observations.

The statistical parameters (21)–(23) have been used to assess the impact of discussed different models. MAE statistics provide an appropriate picture of the actual position in terms of the projected value error, whereas the RMSE statistics are the deviation of the models between the observed and projected values, the least values of these statistics are evaluation criteria for the best model. Similarly, the degree of the linear correlation coefficient is measured by the CC and observed the best flow effect with its high value [25].

4 Results and Discussion

This article examines the wavelet impact on traditional forecasting models by fitting input hydrological time-series datasets collected from Indus and Chenab Rivers. The computational code of the conversed forecasting models has been written in the MATLAB application, including the wavelet toolbox.

The six (M1–M6) appropriate input data specimens have been prepared and used in the training and testing phases for traditional and combination of WD forecast models as shown in Table 1. M1–M6 and WM1–WM6 input combinations signify the number of variables based on earlier analysis for monthly river streamflow rates in Figs. 1–2.

Table 1: Model structure with different input combinations for forecasting streamflow

Model	Original datasets	Model	WD datasets
M1	y_{t-1}	WM1	WD_{t-1}
M2	y_{t-1}, y_{t-2}	WM2	WD_{t-1}, WD_{t-2}
M3	$y_{t-1}, y_{t-2}, y_{t-3}$	WM3	$WD_{t-1}, WD_{t-2}, WD_{t-3}$
M4	$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}$	WM4	$WD_{t-1}, WD_{t-2}, WD_{t-3}, WD_{t-4}$
M5	$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}$	WM5	$WD_{t-1}, WD_{t-2}, WD_{t-3}, WD_{t-4}, WD_{t-5}$
M6	$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}, y_{t-6}$	WM6	$WD_{t-1}, WD_{t-2}, WD_{t-3}, WD_{t-4}, WD_{t-5}, WD_{t-6}$

The training dataset of the models is described for approximated parameters and the testing dataset has characterised by choosing the best combination model amongst every number of hidden layers considered. A trial-and-error technique has estimated the optimum complexity of conversed models. The statistical approaches, such as the Correlation Coefficient (CC), the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE), respectively, have estimated the outcomes. The estimates of both streamflow datasets are described in Tables 2–7 regarding accuracy and errors of the forecasting time-series models, such as LSSVM, WLSSVM, ANN, WANN, MARS and WMARS, respectively.

In Table 2, the testing phase outcomes of the LSSVM model are observed that the M6 input combination performed well for Indus and Chenab Rivers. Likewise, Table 3 examines that WM2 and WM6 input combinations of the WLSSVM model have been estimated good results for both Indus and Chenab Rivers, respectively. After observing both tables, Table 3 yielded more inspired outcomes using the wavelet transform combined model compared to Table 2.

Table 2: Forecast outcomes of LSSVM model

	INDUS RIVER						CHENAB RIVER					
	Training phase			Testing phase			Training phase			Testing phase		
	CC	MAE	RMSE	CC	MAE	RMSE	CC	MAE	RMSE	CC	MAE	RMSE
M1	0.9026	0.0086	0.00025	0.868	0.0318	0.0109	0.9172	0.0421	0.0027	0.9041	0.0437	0.0042
M2	0.9032	0.0086	0.00026	0.7357	0.0279	0.0097	0.9266	0.0401	0.004	0.9126	0.0445	0.004
M3	0.9049	0.0086	0.00025	0.9161	0.0313	0.0109	0.9323	0.0393	0.0039	0.9196	0.043	0.0036
M4	0.907	0.0084	0.00027	0.893	0.0323	0.0115	0.9342	0.036	0.0036	0.9296	0.0414	0.0034
M5	0.9057	0.0085	0.00024	0.9217	0.0336	0.0102	0.9432	0.0345	0.0031	0.9408	0.0382	0.0027
M6	0.9145	0.0078	0.00019	**0.9335**	**0.0247**	**0.0019**	0.9418	0.036	0.0032	**0.9467**	**0.0327**	**0.0021**

Table 3: Forecast outcomes of WLSSVM model

	INDUS RIVER						CHENAB RIVER					
	Training phase			Testing phase			Training phase			Testing phase		
	CC	MAE	RMSE	CC	MAE	RMSE	CC	MAE	RMSE	CC	MAE	RMSE
WM1	0.8994	0.0098	0.0014	0.9137	0.0188	0.0026	0.8689	0.015	0.0019	0.9071	0.0279	0.0203
WM2	0.9224	0.0065	0.0012	**0.9398**	**0.0178**	**0.002**	0.8728	0.0159	0.0015	0.9183	0.0255	0.0019
WM3	0.9063	0.0055	0.0003	0.905	0.023	0.0053	0.9006	0.0157	0.0013	0.9299	0.0258	0.0018
WM4	0.816	0.0041	0.0015	0.7872	0.0251	0.0068	0.927	0.015	0.001	0.9467	0.0247	0.0017
WM5	0.8836	0.0047	0.0021	0.832	0.0246	0.0076	0.9245	0.0141	0.0007	0.9502	0.0239	0.0012
WM6	0.8926	0.0032	0.0023	0.8052	0.0227	0.0067	0.943	0.0135	0.0005	**0.9511**	**0.0233**	**0.0016**

Additionally, in Table 4 is noticed that $M2$, $M5$ and $M1$, $M3$ input combinations of the ANN model are better for Indus and Chenab Rivers. Similarly, $WM5$ and $WM1$ input combinations of the WANN model have given improved outcomes in the analysis of Table 5 for both Indus and Chenab Rivers. Afterward, perceiving both tables, the estimations of Table 5 are more attractive outcomes using the combined model paralleled to Table 4.

Table 4: Forecast outcomes of ANN model

	INDUS RIVER						CHENAB RIVER					
	Training phase			Testing phase			Training phase			Testing phase		
	CC	MAE	RMSE	CC	MAE	RMSE	CC	MAE	RMSE	CC	MAE	RMSE
M1	0.8678	0.0114	0.022	0.8193	0.0301	0.1035	0.9147	0.0173	0.0367	**0.8144**	0.0783	**0.1008**
M2	0.8769	0.0124	0.0213	**0.9125**	**0.0288**	0.0922	0.9278	0.0219	0.0301	0.7956	0.0889	0.2934
M3	0.8536	0.0125	0.0231	0.7094	0.0487	0.165	0.9026	0.0261	0.0346	0.7842	**0.0298**	0.372
M4	0.7428	0.019	0.0297	0.6683	0.04	0.1059	0.8713	0.0271	0.0374	0.6945	0.2685	0.5161
M5	0.7216	0.0207	0.0307	0.8643	0.0318	**0.0824**	0.8102	0.0199	0.0307	0.6141	0.1515	0.1973
M6	0.8454	0.014	0.0237	0.8572	0.0306	0.0892	0.7009	0.0263	0.0332	0.5314	0.1259	0.2392

Table 5: Forecast outcomes of WANN model

	INDUS RIVER						CHENAB RIVER					
	Training phase			Testing phase			Training phase			Phase		
	CC	MAE	RMSE	CC	MAE	RMSE	CC	MAE	RMSE	CC	MAE	RMSE
WM1	0.855	0.0357	0.0027	0.8293	0.0419	0.0029	0.8649	0.0132	0.0275	**0.9286**	**0.0271**	**0.0539**
WM2	0.8678	0.029	0.0019	0.8995	0.036	0.0025	0.8287	0.0155	0.0254	0.8763	0.0719	0.2728
WM3	0.8732	0.0263	0.0016	0.9015	0.0335	0.002	0.8007	0.0231	0.0353	0.8124	0.0991	0.3215
WM4	0.8796	0.0223	0.0012	0.9398	0.0317	0.0019	0.7732	0.0195	0.0295	0.8013	0.1978	0.5081
WM5	0.8871	0.0182	0.0007	**0.9597**	**0.029**	**0.0016**	0.7411	0.0176	0.0263	0.7646	0.207	0.1908
WM6	0.8799	0.0253	0.0015	0.9475	0.0324	0.0017	0.729	0.0154	0.0259	0.7116	0.1079	0.2271

Furthermore, in Table 6 is analysed that $M3$ and $M6$ input combinations of the MARS model are obtained appropriate results for Indus and Chenab Rivers. As well in Table 7 is observed that $WM4$ and $WM5$ input combinations for the WMARS model have been found improved estimations for both Indus and Chenab Rivers. Therefore, Table 7 valuations are improved due to the usage of the combined model than in Table 6.

Table 6: Forecast outcomes of MARS model

| | INDUS RIVER | | | | | | CHENAB RIVER | | | | | |
| | Training phase | | | Testing phase | | | Training phase | | | Testing phase | | |
	CC	MAE	RMSE	CC	MAE	RMSE	CC	MAE	RMSE	CC	MAE	RMSE
M1	0.8715	0.0112	0.0006	0.7991	0.025	0.008	0.884	0.0863	0.0129	0.7869	0.118	0.0271
M2	0.9061	0.0094	0.0003	0.8561	0.0266	0.0087	0.8976	0.0354	0.0029	0.8174	0.1152	0.0297
M3	0.9034	0.0097	0.0003	**0.8968**	**0.0267**	**0.0085**	0.9069	0.0313	0.0025	0.7869	0.1213	0.0352
M4	0.9054	0.0096	0.0003	0.8958	0.0265	0.0083	0.8895	0.03	0.0023	0.7684	0.1243	0.0396
M5	0.9042	0.0098	0.0003	0.8825	0.0275	0.0089	0.9004	0.0281	0.0021	0.7336	0.1352	0.0458
M6	0.9056	0.0096	0.0003	0.8774	0.0274	0.0089	0.9397	0.0344	0.0031	**0.8471**	**0.1112**	**0.0278**

Table 7: Forecast outcomes of WMARS model

| | INDUS RIVER | | | | | | CHENAB RIVER | | | | | |
| | Training phase | | | Testing phase | | | Training phase | | | Testing phase | | |
	CC	MAE	RMSE	CC	MAE	RMSE	CC	MAE	RMSE	CC	MAE	RMSE
WM1	0.9001	0.0112	0.0006	0.8089	0.0315	0.0079	0.8899	0.0693	0.0049	0.8169	0.1075	0.0279
WM2	0.9018	0.0094	0.0003	0.8479	0.0271	0.0081	0.8808	0.0545	0.0037	0.8074	0.1059	0.0265
WM3	0.9048	0.0097	0.0003	0.8491	0.0268	0.0078	0.8809	0.0419	0.0032	0.8091	0.1108	0.0271
WM4	0.9067	0.0095	0.0003	**0.9012**	**0.0259**	**0.0077**	0.8895	0.0379	0.0028	0.8561	0.1047	0.0277
WM5	0.9059	0.0098	0.0003	0.8355	0.0278	0.0084	0.9014	0.0301	0.0027	**0.8599**	**0.0993**	**0.026**
WM6	0.9064	0.0096	0.0003	0.8532	0.0269	0.0082	0.9009	0.0316	0.0031	0.8591	0.1019	0.0269

Figs. 3–4 are described the predicted linear trend line scatter displaying the LSSVM, WLSSVM, ANN, WANN, MARS and WMARS models employing the testing phase and show the performance as relatively to observed data on Indus and Chenab banks. The scatter lines are described in the equation of regression line $(y = a + bx)$ by MATLAB software for each model.

Clearly, Fig. 3 shows that the forecasting model with wavelet transforms performs remarkably because scatter graphs are more appropriate than traditional models.
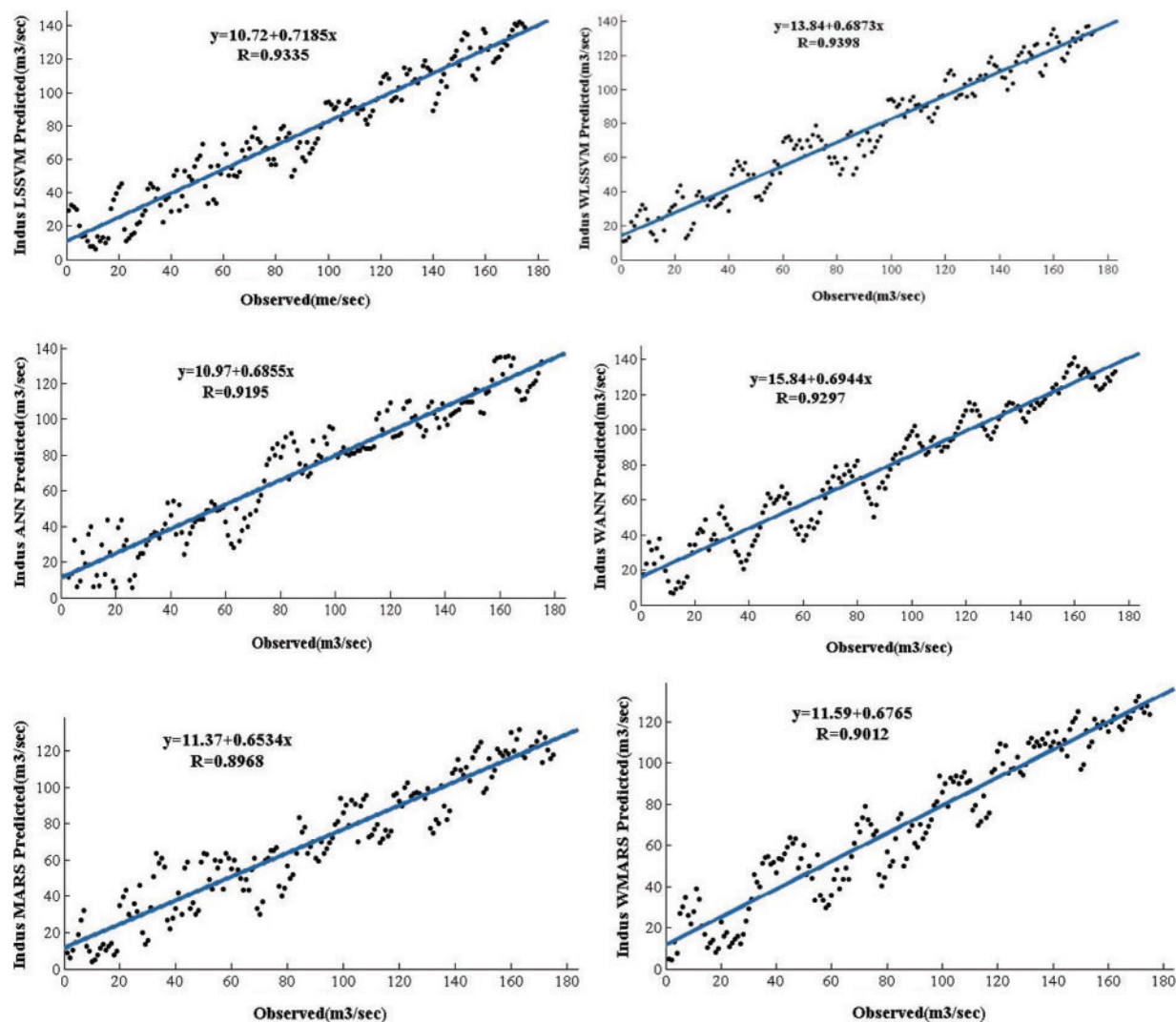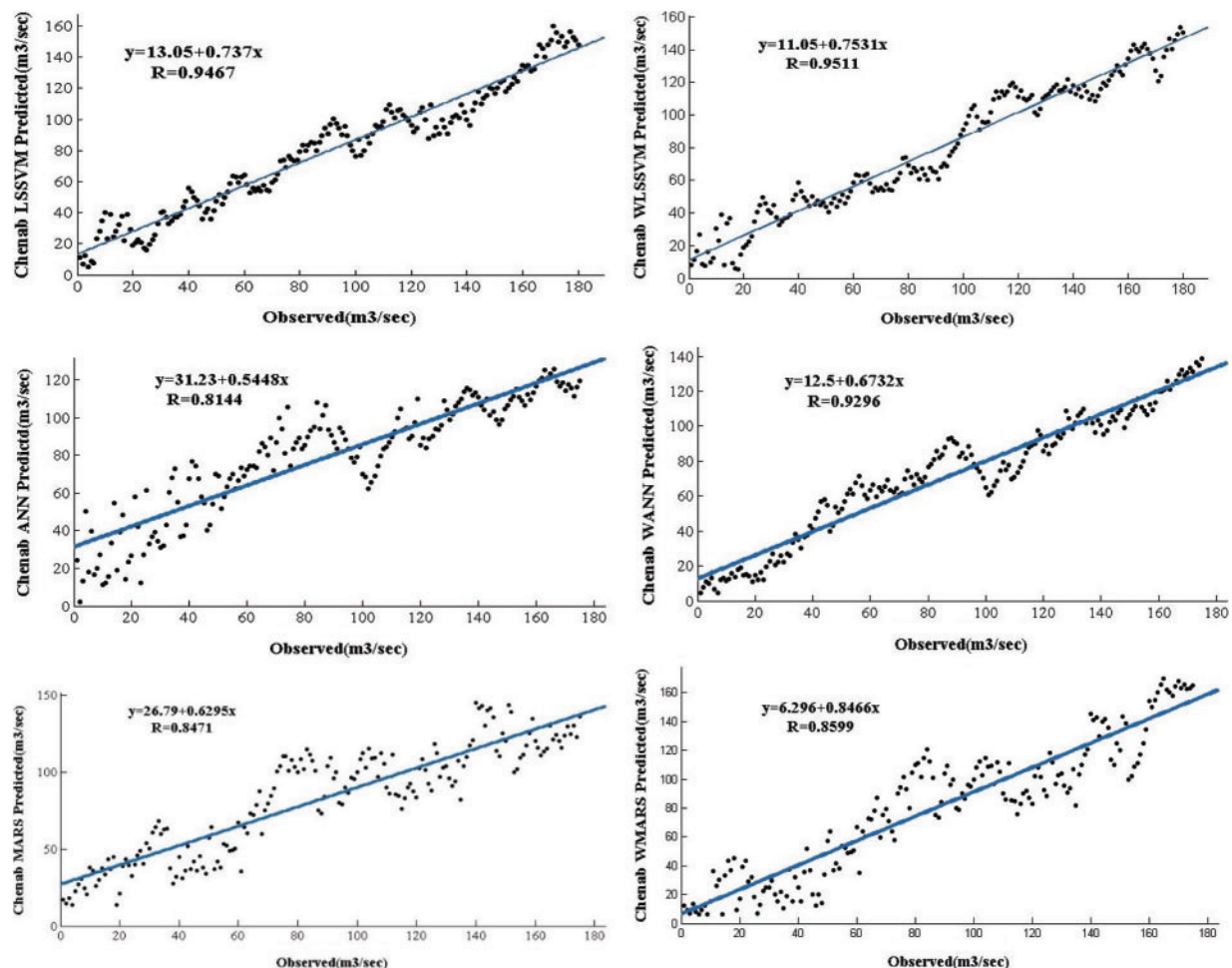
Figure 3: Linear trend line scatter graphs of predicted and observed for Indus streamflow through testing phase

Similarly, Fig. 4 shows that forecasting models with wavelet transform have better forecasting scatter graphs than traditional models.

The approximations of each model are shown in Table 8, scrutinised by various statistical parameters with small errors and large coefficient values with the evidence that models with wavelet decomposition are more efficient than the traditional models.

The CC-values of WD combined models are close to 100% for Indus and Chenab datasets compared to traditional models. Therefore, the WD algorithm combined with traditional models has made a tremendous impact and performs the role of a gadget to deliver improved estimations of both streamflow rivers. Consequently, the combined WLSSVM, WANN, and WMARS forecasting methodologies have been used as the second type models and provide excellent results instead of the first type traditional models LSSVM, ANN and MARS.

Figure 4: Linear trend line scatter graphs of predicted and observed for Chenab streamflow through the testing phase

Table 8: Comparison of forecasting results

Dataset	Model	CC	MAE	RMSE
Indus River	LSSVM	0.9335	0.0247	0.0019
	WLSSVM	0.9398	0.0178	0.002
	ANN	0.9125	0.0288	0.0824
	WANN	0.9597	0.029	0.0016
	MARS	0.8968	0.0267	0.0085
	WMARS	0.9012	0.0259	0.0077
Chenab River	LSSVM	0.9467	0.0327	0.0021
	WLSSVM	0.9511	0.0233	0.0016
	ANN	0.8144	0.0298	0.1008
	WANN	0.9286	0.0271	0.0539
	MARS	0.8471	0.1112	0.0278
	WMARS	0.8599	0.0993	0.026

5 Conclusions

It is concluded that the use of wavelet application with the addressed forecast time-series models has improved the efficiency and yielded tremendous results. The traditional forecasting time-series models have been prescribed by utilising the impact of the wavelet algorithm. The significance of wavelet information is to improve the efficiency of the models that determines the appropriate outcomes for time-series models. The performance of the wavelet combined models mapping with their associated resampling outcomes. Filtrations of the streamflow time-series datasets have been interpreted from the WD application and these features have not been observed in traditional models. The nonlinear input combination models have been constructed with the WD application and used as input estimators with traditional models that improve the forecast efficiency of the combined models. Therefore, the WD application has become an efficient and interesting valuable tool to analysed simulations of time-series datasets models in various domains.

Thus researchers have a good argument in the future for the extensive usage of the wavelet algorithm to build up the novelty in the model or improve the existing models by suitable transform other than wavelet. The WD algorithm provides the optimum ability to pick the appropriate input and logically improves the output of the traditional forecasting models.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. AL W. S., Ismail, M. T., Karim, S. A. A. (2010). Forecasting financial time series data base on wavelet transforms and neural network model. *International Conference on Mathematical Applications in Engineering* , pp. 3–5. Kuala Lumpur, Malaysia.
2. Shao, X. G., Leung, A. K. M., Chau, F. T. (2003). Wavelet: A new trend in chemistry. *Accounts of Chemical Research, 36(4),* 276–283. DOI 10.1021/ar990163w.
3. Kisi, O. (2010). Wavelet regression model for short-term streamflow forecasting. *Journal of Hydrology, 389(3–4),* 344–353. DOI 10.1016/j.jhydrol.2010.06.013.
4. Kaboudan, M. (2004). *Wavelets in forecasting. School of Business,* University of Redlands, Redlands.
5. Nourani, V., Baghanam, A. H., Adamowski, J., Kisi, O. (2014). Applications of hybrid wavelet–artificial intelligence models in hydrology: A review. *Journal of Hydrology, 514,* 358–377. DOI 10.1016/j.jhydrol.2014.03.057.
6. Lee, H. Y., Beh, W. L., Lem, K. H. (2020). Wavelet as a viable alternative for time series forecasting. *Austrian Journal of Statistics, 49(3),* 38–47. DOI 10.17713/ajs.v49i3.1030.
7. Esfetanaj, N. N., Nojavan, S. (2018). The use of hybrid neural networks, wavelet transform and heuristic algorithm of WIPSO in smart grids to improve short-term prediction of load, solar power, and wind energy. *Operation of Distributed Energy Resources in Smart Distribution Networks,* 75–100. DOI 10.1016/B978-0-12-814891-4.00004-7.
8. Pritpal, S. (2021). FQTSFM: A fuzzy-quantum time series forecasting model. *Information Sciences, 566,* 57–79. DOI 10.1016/j.ins.2021.02.024.

9. Wang, W., Ding, J. (2003). Wavelet network model and its application to the prediction of hydrology. *Nature and Science, 1(1),* 67–71.

10. Sang, Y., Wang, D. (2008). Wavelets selection method in hydrologic series wavelet analysis. *Journal of Hydraulic Engineering, 39(3),* 295–300.

11. Mirbagheri, S. A., Nourani, V., Rajaee, T., Alikhani, A. (2010). Neuro-fuzzy models employing wavelet analysis for suspended sediment concentration prediction in rivers. *Hydrological Sciences, 55(7),* 1175–1189. DOI 10.1080/02626667.2010.508871.

12. Pandhiani, S. M., Shabri, A. B. (2015). Time series forecasting by using hybrid models for monthly streamflow data. *Applied Mathematical Sciences, 9(57),* 2809–2829. DOI 10.12988/ams.2015.52164.

13. Senkal, S., Ozgonenel, O. (2013). Performance analysis of artificial and wavelet neural networks for short term wind speed prediction. *8th International Conference on Electrical and Electronics Engineering*, pp. 196–198. Antalya, Turkey.

14. Pritpal, S., Huang, Y. P. (2019). A high-order neutrosophic-neuro-gradient descent algorithm-based expert system for time series forecasting. *International Journal of Fuzzy Systems, 21(7),* 2245–2257. DOI 10.1007/s40815-019-00690-2.

15. Yonaba, H., Anctil, F., Fortin, V. (2010). Comparing sigmoid transfer functions for neural network multistep ahead streamflow forecasting. *Journal of Hydrologic Engineering, 15(4),* 275–283. DOI 10.1061/(ASCE)HE.1943-5584.0000188.

16. Pandhiani, S. M., Shabri, A. B. (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. DOI 10.4236/ojs.2013.33021.

17. Shaikh, W., Shah, S., Solangi, M., Pandhiani, S. (2019). Forecasting analysis of GMDH model with LSSVM and MARS models for hydrological datasets (case study). *Indian Journal of Science and Technology, 12,* 1–6. DOI 10.17485/ijst/2019/v12i39/147941.

18. Ismail, S., Shabri, A., Samsudin, R. (2012). A hybrid model of self organising maps and least square support vector machine for river flow forecasting. *Hydrology and Earth System Sciences, 16(11),* 4417–4433. DOI 10.5194/hess-16-4417-2012.

19. Bansal, P., Salling, J. (2013). Multivariate Adaptive Regression Splines (MARS). http://www.ideal.ece.utexas.edu/courses/ee380l_ese/2013/mars.pdf.

20. Adamowski, J., Chan, H. F., Prasher, S. O., Sharda, V. N. (2012). Comparison of multivariate adaptive regression splines with coupled wavelet transform artificial neural networks for runoff forecasting in himalayan micro-watersheds with limited data. *Journal of Hydroinformatics, 14(3),* 731–744. DOI 10.2166/hydro.2011.044.

21. Fathian, F., Mehdizadeh, S., Kozekalani Sales, A., Safari, M. J. S. (2019). Hybrid models to improve the monthly river flow prediction: Integrating artificial intelligence and non-linear time series models. *Journal of Hydrology, 575,* 1200–1213. DOI 10.1016/j.jhydrol.2019.06.025.

22. Kao, L. J., Chiu, C. C., Lu, C. J., Chang, C. H. (2013). A hybrid approach by integrating wavelet-based feature extraction with MARS and SVR for stock index forecasting. *Decision Support Systems, 54(3),* 1228–1244. DOI 10.1016/j.dss.2012.11.012.

23. Sahoo, B. B., Jha, R., Singh, A., Kumar, D. (2019). Application of support vector regression for modeling low flow time series. *KSCE Journal of Civil Engineering, 23(2),* 923–934. DOI 10.1007/s12205-018-0128-1.

24. Pritpal, S. (2020). A novel hybrid time series forecasting model based on neutrosophic–PSO approach. *International Journal of Machine Learning and Cybernetics, 11,* 1643–1658. DOI 10.1007/s13042-020-01064-z.

25. Sun, Y., Niu, J., Sivakumar, B. (2019). A comparative study of models for short-term streamflow forecasting with emphasis on wavelet-based approach. *Stochastic Environmental Research and Risk Assessment, 33,* 1875–1891. DOI 10.1007/s00477-019-01734-7.