Research Article

Development of a Three-Stage Hybrid Model by Utilizing a Two-Stage Signal Decomposition Methodology and Machine Learning Approach to Predict Monthly Runoff at Swat River Basin, Pakistan

Muhammad Sibtain,[1] Xianshan Li,[1] Ghulam Nabi,[2] Muhammad Imran Azam [ID],[3] **and Hassan Bashir[4]**

[1]Laboratory for Operation and Control of Cascaded Hydropower Station, China Three Gorges University, Yichang 44302, China
[2]Center of Excellence in Water Resources Engineering, University of Engineering and Technology, Lahore 54000, Pakistan
[3]College of Hydraulic & Environmental Engineering, China Three Gorges University, Yichang 44302, China
[4]China College of Environmental Science and Engineering, Hunan University, Changsha 410082, China

Correspondence should be addressed to Muhammad Imran Azam; imranazam@ctgu.edu.cn

Precise and reliable hydrological runoff prediction plays a significant role in the optimal management of hydropower resources. Nevertheless, the hydrological runoff practically possesses a nonlinear dynamics, and constructing appropriate runoff prediction models to deal with the nonlinearity is a challenging task. To overcome this difficulty, this paper proposes a three-stage novel hybrid model, namely, CVS (CEEMDAN-VMD-SVM), by coupling the support vector machine (SVM) with a two-stage signal decomposition methodology, combining complete ensemble empirical decomposition with additive noise (CEEMDAN) and variational mode decomposition (VMD), to obtain inclusive information of the runoff time series. Hydrological runoff data of the Swat River, Pakistan, from 1961 to 2015 were taken for prediction. CEEMDAN decomposes the runoff time series into subcomponents, and VMD performs further decomposition of the high-frequency component obtained after CEEMDAN decomposition to improve the prediction activity. Afterward, the SVM algorithm was applied to the decomposed subcomponents for the prediction purpose. Finally, four statistical indices are utilized to measure the performance of the CVS model compared with other hybrid models including CEEMDAN-VMD-MLP (multilayer perceptron), CEEMDAN-SVM, VMD-SVM, CEEMDAN-MLP, VMD-MLP, SVM, and MLP. The CVS model performs better during the training period by reducing RMSE by 71.28% and 40.06% compared with MLP and CEEDMAD-VMD-SVM models, respectively. However, during the testing period, the error reductions include RMSE by 68.37% and 35.33% compared with MLP and CEEDMAD-VMD-SVM models, respectively. The results highlight that the CVS model outperforms other models in terms of accuracy and error reduction. The research also highlights the superiority of other hybrid models over standalone in predicting the hydrological runoff. Therefore, the proposed hybrid model is applicable for the nonlinear features of runoff time series with feasibility for future planning and management of water resources.

1. Introduction

Managing water resources is crucial from several aspects such as the development of future water bodies, efficient exploitation of hydropower for power generation and irrigation purposes, to prevent disputes, and for the protection of existing water bodies from overexploitation and pollution [1, 2]. Hydrological runoff prediction is an important area in managing water resources, and accurate runoff planning and prediction can prove to be a feasible measure. Runoff and rainfall prediction depend on nonlinear factors including precipitation, uneven flow, topography, anthropic activities,

and evaporation, which make the runoff prediction task a great challenge [3, 4].

Process-driven and data-driven are two approaches for runoff prediction. The data-driven approaches are becoming more popular for extracting accurate predictions with their speedy growth, increasing power of computation, and fewer information needs than the process-driven approaches [5]. Conversely, these artificial intelligence- (AI-) based models exhibit drawbacks and shortcomings like the sensitivity of SVM to parameter selection and overfitting problems faced by artificial neural networks (ANNs) [6]. Furthermore, the preprocessing of input and/or output data is a requirement to make these models to handle nonstationary data [7]. The direct use of the input signal for modeling in AI-based models may not deliver acceptable results; however, the model performance can be improved by applying a pre-processing technique [8]. Appropriate data preprocessing techniques are required to eliminate noise and extract trend from hydrological time series [9].

To overcome the inadequacies of the data-driven models and for more reliable and accurate predictions, hybrid models have been proposed to apply for the prediction of the hydrological time series [10, 11]. Hybrid techniques by employing the data-driven approaches can obtain inclusive and correct information of the different parameters with an additional benefit of improvement in the prediction accuracy. Moreover, these techniques can detect periodicity, volatility, and the random nature of runoff processes [12].

Statistical models are widely employed for modeling of runoff time series [13]. The main disadvantages of these models are the requirements of stationarity and the linearity of the runoff data. They also require a specific time series data length for a robust prediction result [14]. Therefore, the modeling competency of statistical models is limited due to their linear nature for prediction of runoff time series, which exhibits a highly nonlinear and nonstationary nature [15, 16]. Machine learning (ML) models are appropriate for non-stationarity and nonlinearity exhibited by the runoff time series, overcome the constraints of the statistical time series models, and achieve better performance and accuracy than the conventional statistical models for the time series [17, 18]. The comprehensive evaluation of machine learning techniques is provided in [19, 20] and [21]. SVM and MLP are the popular ML techniques in the field of runoff prediction. SVM is a very efficient and robust algorithm with applications in numerous runoff prediction studies due to its better performance. Moreover, SVM offers excellent simplification ability and promising results compared with other machine learning methods for hydrological runoff prediction [22–25].

SVM does not experience the problem of localized minimization and requires less time for the computation compared to ANN; therefore, there are fewer chances of overfitting and poor prediction results compared to ANN [26–28]. SVM obtains the best cooperation between learning capability and the model complexity, based on limited model information, to obtain the best result [29, 30]. Furthermore, global optimization can be used to improve the parameters of SVM which results in better prediction performance than ANN [31]. The MLP represents an advanced version of ANN

and is popular among hydrologists as compared to other ANNs [32]. Many research works exist which used MLP for runoff prediction [33–36]. The previous research as mentioned above highlights the superior performance of SVM for runoff prediction; therefore, in the present study, SVM has been chosen as the final stage to accomplish the task of runoff prediction in the present study.

In the field of runoff prediction, hybrid methods coupled with ML techniques were proposed to improve the prediction accuracy and to obtain better management of data [37, 38]. Hybrid ML methods in the runoff prediction field offer advantages of automated and timely performance evaluation and management of the ensemble algorithms [32]. The research in [39, 40] presents a review of the applications of the hybrid ML methods for runoff-rainfall prediction.

Decomposition techniques can be applied as data preprocessing tools to study the nonlinear and nonstationary characteristics of runoff series, such as ensemble empirical mode decomposition (EEMD) and variational mode decomposition (VMD) [17, 41, 42]. Time series decomposition finds successful implementation in improving the performance of ML methods used for runoff modeling. The decomposition method decomposes the original time series into several individual components; afterward, the ML models are employed for prediction purposes [17]. The components obtained as a result of employing an effective decomposition method are much easier to evaluate than an original time series [43].

The hydrological time series has been analyzed by many researchers by employing wavelet transform (WT) due to its excellent performance in conditions with multiple resolutions in frequency and time domains [44, 45]. The WT represents a Fourier transform with an adjustable window requiring a stable signal in the WT window. Consequently, WT is prone to the restrictions of the Fourier transform. Although WT provides high resolution in both the time domain and the frequency domain, some false harmonic waves are produced during WT due to certain limitations of this method. Therefore, the selection of WT basis functions is crucial due to its significant influence on the process of the wavelet decomposition. Empirical mode decomposition (EMD) was proposed to overcome the limitations of the WT [46]. EMD decomposes the trend components or the multiscale fluctuation in the signal to smoothen the signal and generates intrinsic mode functions along with a residual. EMD technique reflects a more accurate representation of the nonlinearity and nonstationary in the original series compared to the WT technique. Therefore, EMD is considered as a more efficient way to process complex signals than the WT. The hydrological time series in classical hydrology can be considered as a combination of random, periodic, and trend components. The high-frequency and low-frequency components along with the residual obtained through EMD technique in the case of perfect decomposition can be approximated as random and periodic components along with the trend [47, 48]. EMD finds successful implementations in the hydrological research [49, 50], but EMD exhibits problems such as the mode mixing of IMFs and the orthogonality effect which affects the precision of prediction and the performance of EMD. Therefore, EEMD

was developed to resolve the issues and to lessen the impact of mode mixing as faced by EMD [51, 52]. Nevertheless, the limitations of mode mixing of some signals and the end effect still exist in EEMD-based techniques [53]. Complete ensemble empirical mode decomposition with additive noise (CEEMDAN) is an advanced technique that overcomes the issues faced by EMD and EEMD like mode mixing and computational complexity, respectively. It is possible to achieve the reconstruction error close to zero by utilizing the CEEMDAN technique and by requiring fewer integration times, with the addition of adaptive white noise at each step [43]. However, CEEMDAN is also unable to completely resolve the issue faced by EEMD, such as the presence of residual noise in the modes and appearance of signal information later than in EEMD with specious modes in the initial decomposition stages [54].

VMD is another adaptive and nonrecursive signal analysis technique which, unlike empirical mode techniques, decomposes the original series into multiple modes and updates them [55]. VMD is more robust to noise and sampling, with outstanding performance in frequency search and separation. VMD can improve the mode mixing problem and extract the time-frequency features precisely by yielding narrow-banded modes [56]. The VMD is a comparatively new technique for hydrological application [17], and relatively, a few research works exist regarding applications of VMD for runoff prediction.

A single-layer hybrid model consisting of a decomposition technique and machine learning method is one of the most frequently employed methods to analyze the time series. These hybrid models consisted of a single-layer decomposition technique that can enhance the predictive performance of nonlinear time series to some extent but unable to completely predict the nonlinearity and nonstationarity of the original signals. Consequently, the hybrid model based on two-layer decomposition methodology is employed to overcome the limitations of the single-layer decomposition technique [43]. Therefore, this study proposes a three-stage hybrid model based on CEEMDAN, VMD, and SVM and its applicability to the runoff time series. The first decomposition stage employs the CEEMDAN technique and decomposes the runoff series into random, periodic, and trend components intending to improve the prediction of nonlinear and nonstationary monthly runoff series. The VMD is proposed as an additional decomposition technique to diminish stochastic behaviors, noise, and trends of the data. Finally, the SVM algorithm predicts the monthly runoff data series.

The main objectives of the present research are as follows:

(1) The development of ML and the signal decomposition-based hybrid model by taking the hydrological runoff data of the Swat River, Pakistan.

(2) Applicability of the hybrid model for the runoff prediction.

(3) Verifying the performance and accuracy of the proposed model by comparing results with the similar models used to predict the runoff time series.

The rest of the paper is arranged as follows. The modeling techniques and the proposed approach are described in Section 2. The results and discussion are presented in Section 3, while Section 4 concludes the research work. The research will be useful for prediction and planning purposes and will provide new directions in the field of hydrology.

2. Materials and Methods

2.1. Decomposition Techniques

2.1.1. CEEMDAN. In CEEMDAN, the information regarding noise is shared between all workers as opposed to EEMD to efficiently solve the mode mixing issue of EMD [57]. The CEEMDAN technique enables us to get near to zero reconstruction error by adding a finite number of adaptive white noises at every stage through a lesser average number of integration times. This enables the CEEMDAN to avoid the mode mixing and the computational complexity issues [43]. The CEEMDAN process proceeds as follows:

Step 1: create the original time series with added noise:

$$x^i(t) = x(t) + \delta_0 \gamma^i(t). \tag{1}$$

Step 2: use CEEMDAN to get the first IMF for each $x^i(t)$ and take the average:

$$\text{IMF1}(t) = \frac{1}{M} \sum_{i=1}^{M} \text{IMF1}^i(t). \tag{2}$$

The first residual is $r_1(t) = x(t) - \text{IMF1}(t)$. Step 2 is similar to EMD.

Step 3: CEEMDAN gets second and the remaining IMFs by decomposing the residual with the added noise as shown in the following:

$$\text{IMF}_2 = \frac{1}{M} \sum_{i=1}^{M} E_1\left(r_1(t) + \delta_1 E_1\left(\gamma^i(t)\right)\right), \tag{3}$$

$$r_2(t) = r_1(t) - \text{IMF}_2, \tag{4}$$

where $E_1(\cdot)$ represents the first IMF decomposed from the original signal. Similarly, the k-th IMF and the residual can be calculated as

$$\text{IMF}_k = \frac{1}{M} \sum_{i=1}^{M} E_1\left(r_{k-1}(t) + \delta_{k-1} E_{k-1}\left(\gamma^i(t)\right)\right), \tag{5}$$

$$r_k(t) = r_{k-1}(t) - \text{IMF}_k. \tag{6}$$

Step 4: CEEMDAN obtains numerous IMFs and computes the residual as shown in the following:

$$R(t) = x(t) - \sum_{k=1}^{K} \text{IMF}_k. \tag{7}$$

Based on VMD, this study introduces a second decomposition of IMF1, owing to the unpredictability and the highest frequency of IMF1.

2.1.2. Variational Mode Decomposition.

VMD is a quasi-orthogonal and adaptive decomposition method, where the modes are obtained nonrecursively [55]. It approximates the corresponding modes concurrently and determines the relevant band adaptively [53]. The VMD can be expressed as [55]

$$f = \min_{\{X_k\}, \{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\tau t} \right) \times X_k(t) \right] e^{-j\omega_k(t)} \right\|_2^2 \right\}$$

$$\text{subject to} \quad \sum_k X_k, \tag{8}$$

where $\{X_k\}$: $\{X_1, X_2, \ldots, X_k\}$ and $\{\omega_k\}$: $\{\omega_1, \omega_2, \ldots, \omega_k\}$ represent expressions related to all modes along with their central frequency.

Furthermore, $\delta(t)$ and $\times$ represent Dirac distribution and convolution, respectively. The term of the quadratic penalty and Lagrangian multipliers are used to convert this constrained optimization problem to an unconstrained one [58]:

$$L(\{X_k\}, \{\omega_k\}, \lambda) = \alpha \left\{ \sum_k \left\| \partial_t \left[\left(\delta_t + \frac{j}{\pi t} \right) \times X_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\}$$

$$+ \left\| f(t) - \sum_k X(t) \right\|_2^2$$

$$+ \langle \lambda(t), \quad f(t) - \sum_k X_k(t) \rangle. \tag{9}$$

The above equation can be resolved using different approaches, and the two stages of the equation are given as follows:

(i) X_k minimization:

$$X_{fk}^{n+1} = \frac{f_f(\omega) - \sum_{i \neq k} X_{fi}(\omega) + \lambda_f(\omega)/2}{1 + 2\alpha(\omega - \omega_k)^2}. \tag{10}$$

(ii) ω_k minimization:

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |X_{fk(\omega)}|^2 d\omega}{\int_0^\infty |X_{fk(\omega)}|^2 d\omega}, \tag{11}$$

where n denotes the number of iterations and $X_{fk}^{n+1}, \lambda_f(\omega), f_f(\omega), X_{fi}(\omega)$ show the Fourier transform of $X_k^{n+1}, \lambda(\omega), f(\omega), X_i(\omega)$, respectively.

VMD technique relies on the three fundamental concepts including wiener filtering, frequency mixing and heterodyne demodulation, and analytic signal and Hilbert transform. The original signal is decomposed into IMFs that reproduce the original signal with different sparsity features. In contrast to the original decomposition techniques, VMD relies on the alternate direction method of multipliers (ADMM) for reconstruction process [53]. VMD utilizes the principle of the variational mode to obtain the IMFs, thereby minimizing the sum of estimated bandwidth of each IMF, which makes this technique different from EMD. The bandwidth and center frequency of the IMFs are revised in the course of solving the variational model. The frequency domain of the signal results in adaptive segmentation of the signal band, and additionally, the IMF obtained has a narrow band [31].

The number of intrinsic modes defines correctly resolved data for an acceptable prediction model of a time series; therefore, the determination of the number of intrinsic modes is vital in the VMD process. The specification of the original time series dataset is impossible to be given if less intrinsic mode components than the required one are chosen. On the contrary, the intrinsic modes in excess may result in the poor performance that causes error accumulation by each prediction unit in the accumulation stage [59]. Nevertheless, the IMFs generated by the VMD process are usually evener compared to the mode functions obtained by other techniques like EEMD and wavelet transforms [60]. This lessens the accumulation of error over time. Another important aspect of VMD is the selection of several parameters, which requires trial and error methods [3].

2.2. Machine Learning Techniques

2.2.1. Support Vector Machine.

SVM is a nonlinear search algorithm [61] used to minimize the expected errors in ML and to reduce the issue of overfitting [62]. Based on the training, by considering the past data, SVM predicts a forward quantity in time [32]. In the case of properly determined kernel filters and support vectors, SVM performs more efficiently than ANNs [24]. SVM works by constructing a hyperplane to enable the maximum sorting between the samples and to minimize the sample to the hyperplane distance [31].

SVMs were developed for the binary classification but are also applicable for the regression problems by introducing a loss function. The SVM algorithm only deals with linear problems. In the case of a nonlinear system, a nonlinear mapping is used to map the input vector x into the high-dimensional feature space z; afterward, the linear regression is performed in this space. In case of the radial basis function [63],

$$K(x, x_j) = \exp \left\{ \left\| x_j - x^2 \right\| \delta^2 \right\}, \tag{12}$$

Support vector regression (SVR) is used to apply SVM [24]. SVR based on the structural risk minimization theory and Vapnik–Chervonenkis dimension model is a feasible method to deal with prediction problems [26, 64]. Equation (13) gives the standard form of the SVR model:

$$y = f(x) = (w \cdot \varnothing(x)) + b. \tag{13}$$

The coefficients w and b are predictable by decreasing the risk function $R(C)$.

Three parameters dominate the accuracy of the SVR network when the quality and span of the training samples are fixed: ε is the epsilon and controls the width of the epsilon tube in the training loss function, σ controls the width of kernel Gaussian function, and C is the regularization parameter and controls the risk degree of SVR empirically [65–67].

2.2.2. Multilayer Perceptron.

The MLP is the most widely used type of ANN for modeling the hydrological runoff data [68]. MLP belongs to feedforward neural networks. MLP can approximate both integrable and continuous functions. MLP consists of neurons arranged in layers in the form of groups. The input nodes in MLP are all in one layer, while the hidden portion has one or more hidden layers. The selection of layers is dependent on the problem being considered, and there are no specific rules for the selection of these layers [69]. Many algorithms [70–73] have been proposed for finding the optimum structure of the network, but the optimal solution of parameters is not guaranteed by any of these methods. Figure 1 illustrates a simple structure of the MLP network.

In MLP, the nodes of the input layer denote the length of the input data, while the neurons of the output layer show the length of the output data [74]. The calculations in the MLP network are performed successively from the input layer to the output layer. The calculation of the node is performed at the same time, which is present at the same level, and there is no interference of nodes during the process with each other [75]. The weighted sum of all nodes in the preceding layer is equal to the value of each node. The following formula can be used to calculate the value of each node in MLP:

$$x_{ij} = f\left(W_i X_{i-1} + b_{i-1}\right), \tag{14}$$

where f is the activation function and W_i denotes the weight vector. The value vector of all neurons in the $i-1$ layer is X_{i-1}. x_{ij} shows the j neuron value in the i layer, and bias of layer $i-1$ is represented by b_{i-1}.

The linear and nonlinear functions are the most widely used activation functions. MLP is essentially a single-layer perceptron in the case of a linear activation function. The most commonly used nonlinear activation is the sigmoid function [74]. Equation (3) describes the sigmoid activations as follows:

$$f(x) = \text{sigmoid}(x) = \frac{1}{1 + e^x}. \tag{15}$$
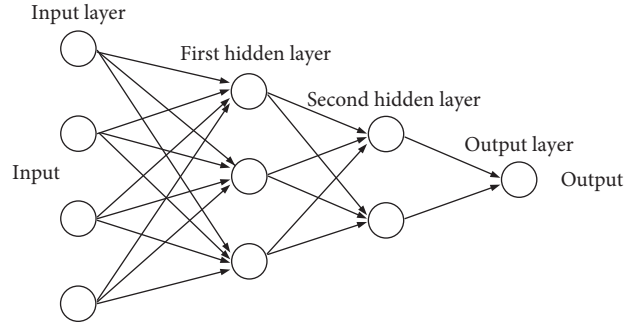

Figure 1: Simple structure of a three-layer MLP model having one input, two hidden, and one output layer.

The expression of loss function of the actual value and the ideal output can be given as

$$J(W, b; x, z) = \frac{1}{2}\left\|h_{W,b(x)} - z\right\|^2, \tag{16}$$

where z denotes the actual value, the output value is given by h, and distance norm is shown by $\|\cdot\|$

The backpropagation (BP) algorithm is usually used to adjust the parameters of MLP and serves to minimize the loss function. The gradient descent (GD) algorithm is the simplest and commonly employed parameter adjustment algorithm. The stochastic gradient descent (SGD) algorithm is another useful algorithm to adjust the parameters of MLP [74]. The SGD algorithm performs well for the optimization process; however, it exhibits a slow convergence rate. Additionally, the chance exists for the gradient descent to experience the issue of loss of the function's saddle point [76]. Several alternative approaches exist to address these issues and for updating the parameters of the neural network. These adaptive approaches diagonally scale the gradient through an approximation of the curvature of function [77]. The most widely used optimizer in deep learning is Adam (adaptive moment estimation) which can be chosen as the best optimizer for nonstationary objectives without the need of other optimizers [78].

2.3. Quantitative Performance Indices.

In this research, several statistical indices are used to assess the performance of the observed and the predicted runoff data. The root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), mean squared error (MSE), and the coefficient of determination (R^2) were applied to evaluate the reliability of the predicted model.

2.3.1. RMSE.

RMSE is used to measure the deviation between the predicted and the observed value, and to represent the extent of dispersion of a dataset, the smaller value represents the better performance of the algorithm:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(N_i - O_i\right)^2}. \tag{17}$$

2.3.2. MAE. MAE reflects the actual condition of the predicted error, and the smaller value represents the better performance of the algorithm:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |N_i - O_i|. \tag{18}$$

2.3.3. MAPE. MAPE is the measure of the model's prediction outcome, and the smaller value represents the superior performance of the algorithm. The MAPE is useful for qualitative assessment of the model's accuracy and depends on the complexity of the application [31]:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{N_i - O_i}{O_i} \right|. \tag{19}$$

2.3.4. MSE. It represents an average of squares of the difference between the predicted and the observed value, and the smaller value represents the better performance of the algorithm:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (N_i - O_i)^2. \tag{20}$$

2.3.5. Coefficient of Determination (R^2). R^2 summarizes the error by evaluating the linear correlation between the observed and the predicted data, with values ranging between 0 and 100%:

$$R^2 = \left[\frac{\sum_{i=1}^{n} (O_i - \overline{O}_i)(N_i - \overline{N}_i)}{\sqrt{1/n \sum_{i=1}^{n} (O_i - O_{im})^2 \sum_{i=1}^{n} (N_i - N_{im})^2}} \right]^2. \tag{21}$$

In expressions (11)–(18), N_i and O_i denote the ith predicted and actual values of runoff, respectively, whereas n shows the total number of predictions.

2.4. Proposed Hybrid Modelling. As explained above, hydrological runoff exhibits nonstationary and nonlinear characteristics [79, 80]. These properties of runoff result in the undesirable performance of many prediction models, along with poor generalization due to the requirements of many pseudo-variations, which also affects the accurate knowledge of data variations [81]. Therefore, this paper proposes a three-stage hybrid model by coupling the ML method with signal decomposition techniques for a reliable and accurate runoff prediction.

The flowcharts representing the main steps of the proposed methodology and seven other models developed for comparison with the proposed CVS model are given in Figure 2. Furthermore, the main steps of the CVS model are explained as follows:

> Step 1: the Pearson correlation coefficient method was applied to the original runoff time series and its lagged values to determine the appropriate input variable, and the time-lagged series having the highest value of the

correlation coefficient with the original runoff series was chosen as an initial input for the decomposition model

Step 2: CEEMDAN technique was applied to the lagged runoff time series obtained as a result of Step 1, which decomposes the series into subcomponents (IMFs and residual) having a different frequency

Step 3: the component of high pass (IMF1) produced by CEEMDAN was further decomposed by VMD

Step 4: SVM algorithm was applied to construct a prediction model for the whole dataset containing extracted IMF components and the runoff data signal to make a prediction for each component accordingly

Step 5: to produce a collective output, the predicted results of all extracted IMFs obtained by the SVM algorithm were reconstructed to produce the final prediction result of the original runoff series

Step 6: finally, the statistical performance metrics evaluated the results in the training and testing periods

3. Results and Discussion

3.1. Case Study. The present research considers the runoff data of the Swat River basin collected from the Water and Power Development Authority (WAPDA), Pakistan, for the prediction purpose. The Swat River is a perennial river located in the northern part of Khyber-Paktunkhwa Province, Pakistan (Figure 3). It originates from Hindu Kush mountains and flows through the Kalam valley to Madyan and lower areas of Swat valley up to Chakdara. The river outflows into the Kabul river and has a total length of 240 km. The Swat River serves the purpose of irrigation, power generation, and a natural habitat for fishes and birds. The catchment area of the river is generally hilly, with altitudes ranging from 360 m to 4,500 m approximately, from south to north. The location of the catchment of the Swat River basin is between longitude 70°59′ east to 72°47′ east and latitude of 34°00′ north to 35°56′ north [82].

3.2. Data Selection. The monthly runoff data of the Swat River from 1961 to 2015 were taken at Chakdara hydrological station in the Swat River catchment. The data are available on a daily basis, and to obtain monthly data, the average monthly data were calculated from the daily data. The monthly runoff data series is shown in Figure 4 and was selected for prediction.

For developing the CVS hybrid model, runoff data are divided into training (approximately 80% of the whole dataset) and testing datasets (approximately 20% of the whole dataset) to predict a 1-month-ahead runoff. To compare the effectiveness of the proposed CVS model, seven other models were used for evaluation: CEEMDAN-VMD-MLP, CEEMDAN-SVM, VMD-SVM, SVM, CEEMDAN-MLP, VMD-MLP, and MLP. Afterward, four statistical indexes are employed, including RMSE, MAE, MAPE, and MSE and R^2 to compare the performance of the proposed model with other models.
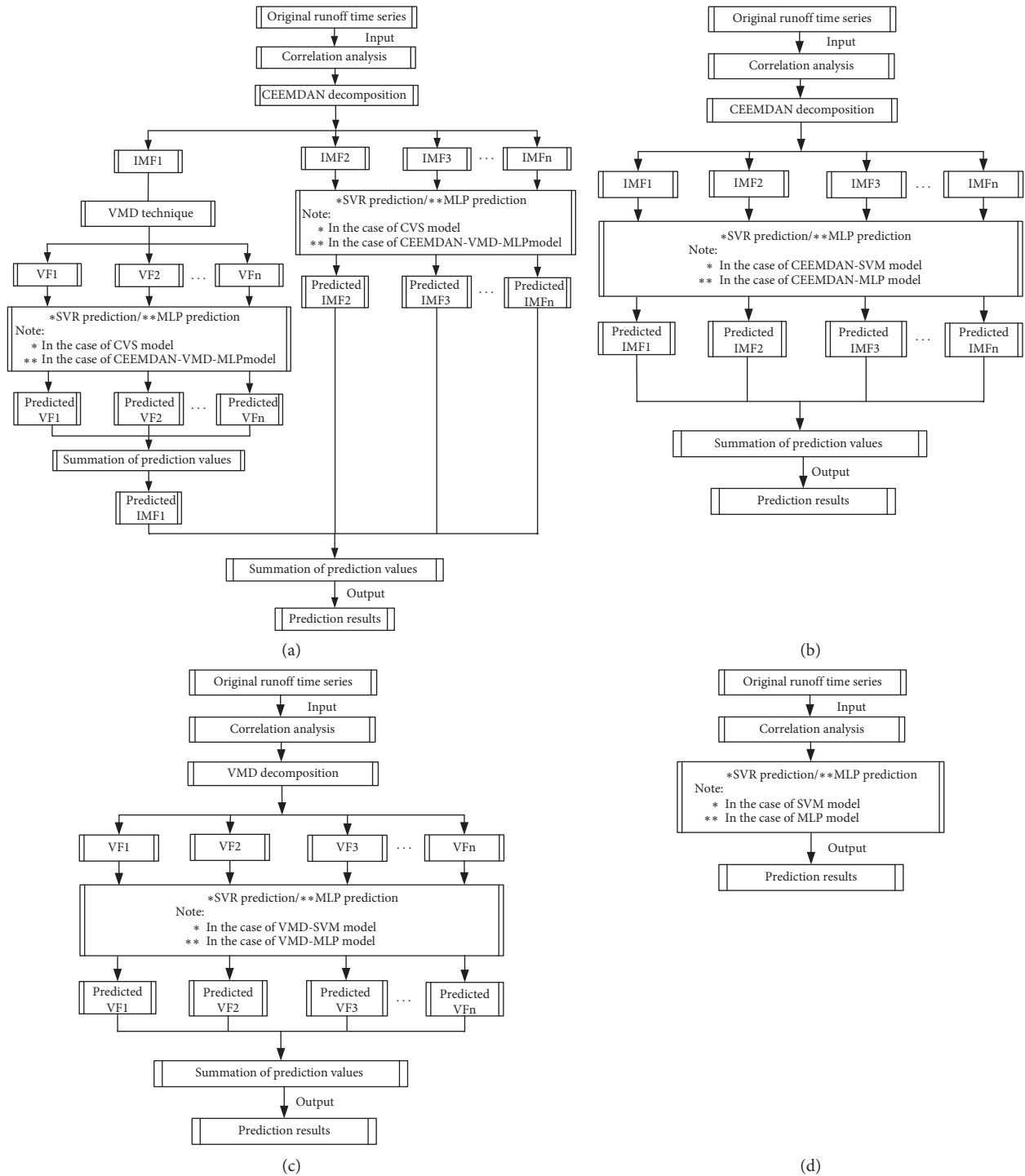
Figure 2: Flowcharts of different models. (a) CVS/CEEMDAN-VMD-VLP model. (b) CEEMDAN-SVM/CEEMDAN-MLP model. (c) VMD-SVM/VMD-VLP model. (d) SVM/MLP model.

The research work was carried out using a 64-bit Windows 10 operating system on a 3.70 GHz, Intel (*R*) Core i7-10510U CPU with 16 GB memory. The analyses were performed using Matlab R2015a software and Python 3.6 relying on pandas and NumPy packages. The optimal parameters were selected after different trials and errors, considering the best results. SVM and MLP networks were developed with Keras using Google Tensorflow backend. MLP network in all

MLP-based models was developed with two hidden layers having 64 and 32 hidden neurons, respectively, with sigmoid activation function, while the output layer has 1 neuron to predict runoff. Moreover, different learning rates were selected for each MLP-based model. Due to the nonstationary and noisy nature of the runoff time series, we applied the adaptive moment estimation (Adam) optimizer for efficient stochastic optimization [83].
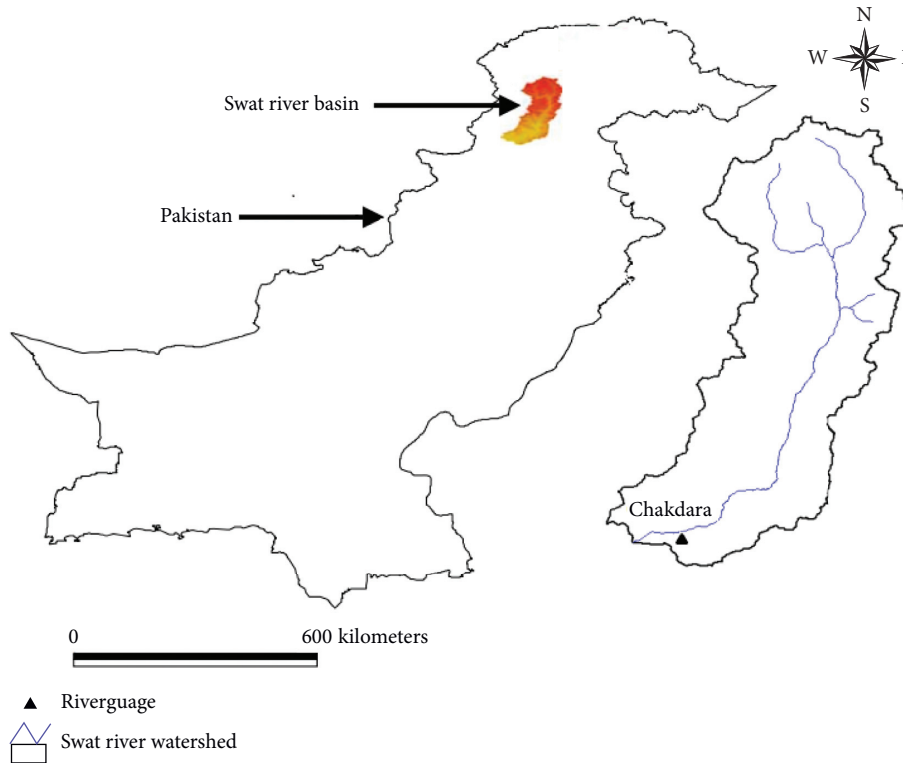
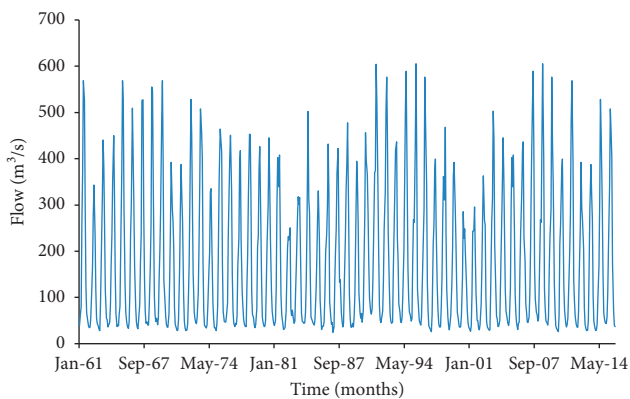FIGURE 3: The location of Swat River basin in Pakistan.

FIGURE 4: The monthly runoff series of Swat River at Chakdara.

For SVR-based models, the radial basis function (RBF) was selected as a kernel for all models with different values of C and σ for each model. In the case of CEEMDAN, the standard deviation of noise was selected as 0.2, the number of realizations allowed was chosen as 500, while the allowed maximum number of sifting iterations was taken as 5000. The values of the different parameters of CEEMDAN are taken from [57], and the same reference explains the detailed procedure of the parameter selection for CEEMDAN. The selected parameters for VMD include moderate bandwidth constraint, alpha = 2000; uniform initialization of omegas, init = 1; criterion for the tolerance of convergence, tol = 1e-7; and noise tolerance, tau = 0, while the value for the number of modes, K was chosen through correlation analysis of the frequency modes generated by CEEMDAN.

3.3. Analysis.
In developing ML-based hydrological models, the selection of suitable input variables is one of the most important steps [84]. The autocorrelation function (ACF) determines an appropriate input dataset for the model corresponding to the runoff at the output by applying a lag time to the original runoff time series [3, 17, 85]. Therefore, to determine a suitable input dataset for the hybrid model in the present study, an ACF was applied to the runoff time series by applying a monthly time lag for a year. As evident from Figure 5, Q_{12} shows the highest value of correlation; therefore, the Q_{12} dataset was selected as an input for runoff prediction.

By employing the CEEMDAN as a preprocessing technique, the selected runoff data series after a time lag (input signal) was decomposed into a sequence of eight independent IMFs and a residual, i.e., eight quasi-stable components and one trend component are obtained due to the decomposition of a nonstationary runoff data series (Figure 6). The denoising process of the time series is not required since CEEMDAN has good antinoise features [43]. It is evident that the IMF1 component has the highest frequency and shows strong nonlinearity and significant fluctuations. However, the remaining IMFs (IMF2~IMF8) and the residual indicate a stable and regular fluctuation which shows a gradual reduction in the frequency with an increase in the wavelength.

The secondary decomposition of IMF1 was carried out by VMD due to the presence of high oscillatory fluctuations in IMF1. The trial and error method was used to select several parameters in the VMD technique [3]. The value of the K parameter can also be determined in ensemble decomposition techniques by correlation analysis [31]. In the
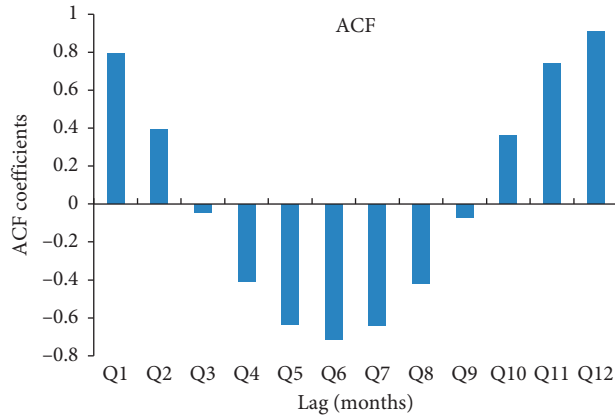
FIGURE 5: The autocorrelation function (ACF) coefficient graph of monthly runoff.
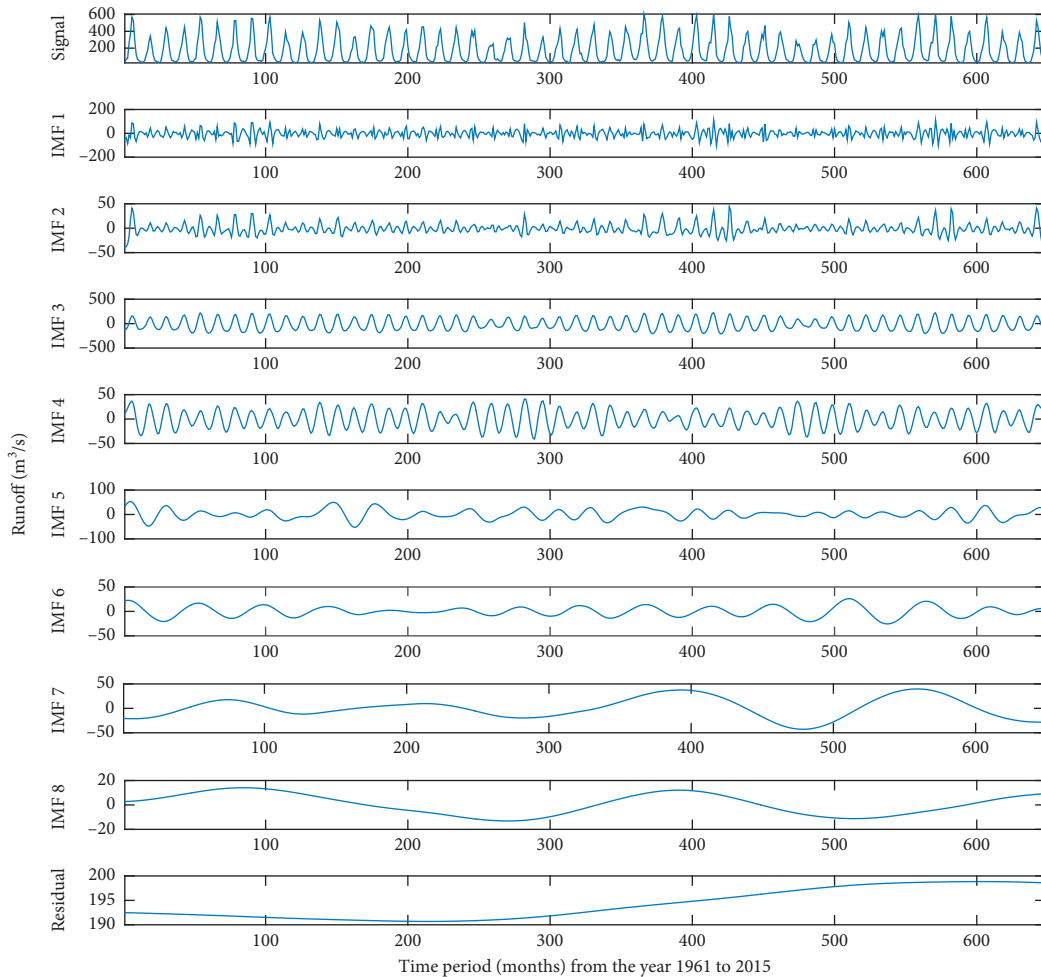

FIGURE 6: Decomposition of runoff time series using CEEMDAN.

present study, we will also obtain K value by correlation analysis of the intrinsic modes produced by CEEMDAN. The correlation coefficient between the input signal and the IMFs including the residual was calculated (Table 1).

The third IMF shows a strong correlation with the input signal, and IMF3 was considered as a borderline for the selection of IMFs as values of K in VMD. IMFs 1 and 2 showed less correlation than IMF3 and were considered as one value of K for VMD decomposition, while the remaining IMFs 3–9 including the residual were taken as seven values of K. Hence, we obtained the value of $K = 8$ for VMD. Figure 7 shows the decomposition results of IMF1 by VMD.

TABLE 1: The correlation coefficient between the input signal and the IMFs of CEEMDAN.

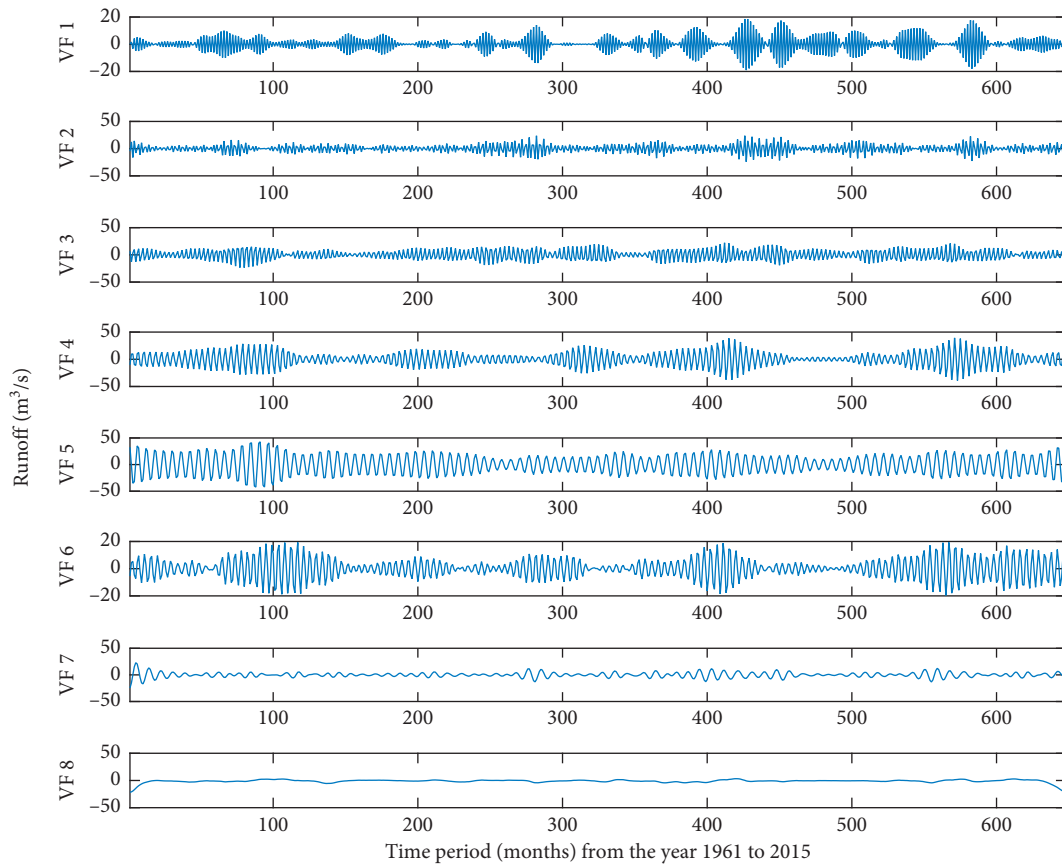| IMF/residual	Correlation coefficient
IMF1	0.3162
IMF2	0.6750
IMF3	0.9369
IMF4	0.8131
IMF5	0.1534
IMF6	0.1023
IMF7	0.1337
IMF8	0.0823
Residual	0.0177

FIGURE 7: VMD decomposition of IMF1 (obtained after CEEMDAN decomposition).

The VMD produces smoother intrinsic modes compared with other decomposition techniques [60] which is also verified by the decomposition result of IMF1 (Figure 7). The decomposed time series components obtained after applying CEEMDAN and VMD along with the runoff data series were applied as an input to SVR for training and validation of data. SVR was used to predict VF1-VF8; afterward, the prediction results of IMF1 were combined with IMFs (IMF2–IMF8 and residual) produced by CEEMDAN to obtain the prediction results of the runoff time series of Swat River. The performance of the CVS model during training and testing periods is evaluated and compared with CEEMDAN-VMD-MLP, CEEMDAN-SVM, VMD-SVM, CEEMDAN-MLP, VMD-MLP, VMD, and MLP models to verify the effectiveness of the proposed model. The results are presented in Figures 8–14 and Tables 2 and 3.

Boxplots (Figures 12 and 13) indicate the range of quartile-based predicted and original (observed) runoff, while whiskers show the variability from the exterior of the 25th to 75th percentiles. The testing phase indicates more skewness and dispersion in prediction compared with the training phase. From the above figures showing training and testing results, it is evident that the CVS model simulates well with stable behavior than all the other models, indicating the superior capability of the CVS model in nonlinear runoff modeling. Moreover, the CVS model can mimic the runoff well than the other models in both training and testing phases, and overall, the hybrid approach
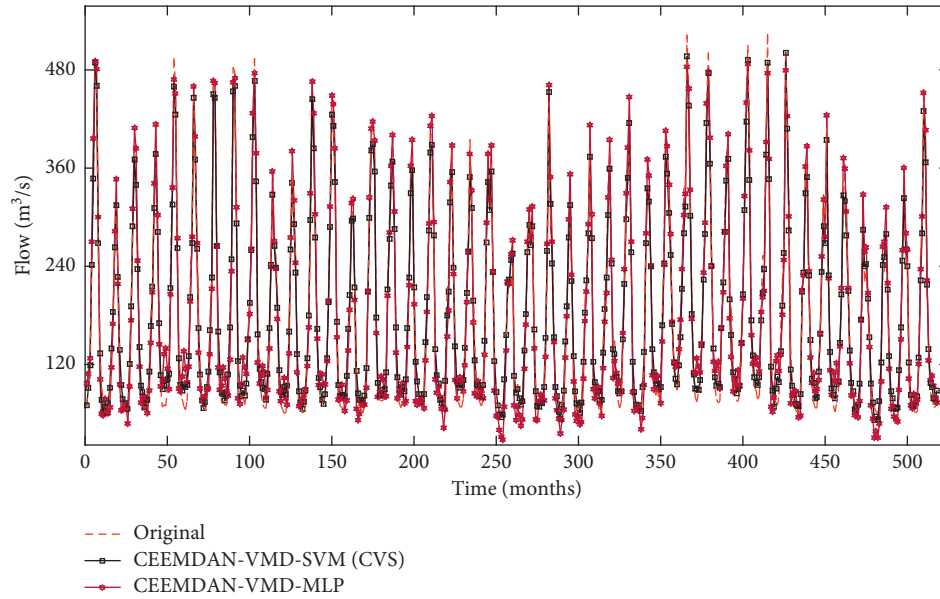
FIGURE 8: Observed and predicted runoff during the model training period using CVS and CEEMDAN-VMD-MLP models.
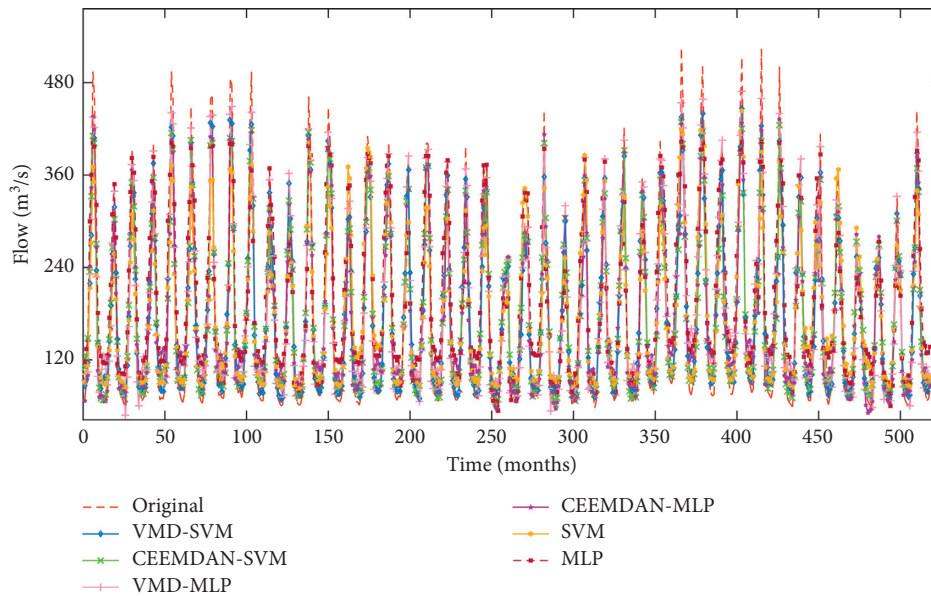

FIGURE 9: Observed and predicted runoff during the model training period using VMD-SVM, CEEMDAN-SVM, VMD-MLP, and CEEMDAN-MLP models.

performance is better than the individual models. Furthermore, the CVS model shows better prediction in the training period compared with the testing period.

As per the results (Tables 2 and 3) during the training period, the CVS model showed the lowest error in terms of lowest RMSE (0.1185), MAE (0.0941), MAPE (0.2398), and MSE (0.0140), whereas the MLP model showed the least performance than all other models with highest RMSE (0.4126), MAE (0.3957), MAPE (1.1065), and MSE (0.1702). During the testing period, MLP again revealed the greatest error with RMSE (0.4578), MAE (0.4282), MAPE (1.1896), and MSE (0.2096), while the CVS model outperformed all

the other models having lowest error with lowest RMSE (0.1448), MAE (0.1192), MAPE (0.26276), and MSE (0.0209). To elaborate on the performance of the CVS model for the runoff prediction, a comparison of R^2 values for different models is provided in Figure 14.

The correlations amongst the original and the predicted runoffs for standalone models (SVM and MLP) are lowest than the hybrid models (Figure 14). The CVS model shows the highest correlation for training ($R^2 = 0.9856$) and testing ($R^2 = 0.9804$) periods, while MLP showed the lowest correlation during training ($R^2 = 0.8263$) and testing ($R^2 = 0.8050$) periods. The performance of standalone
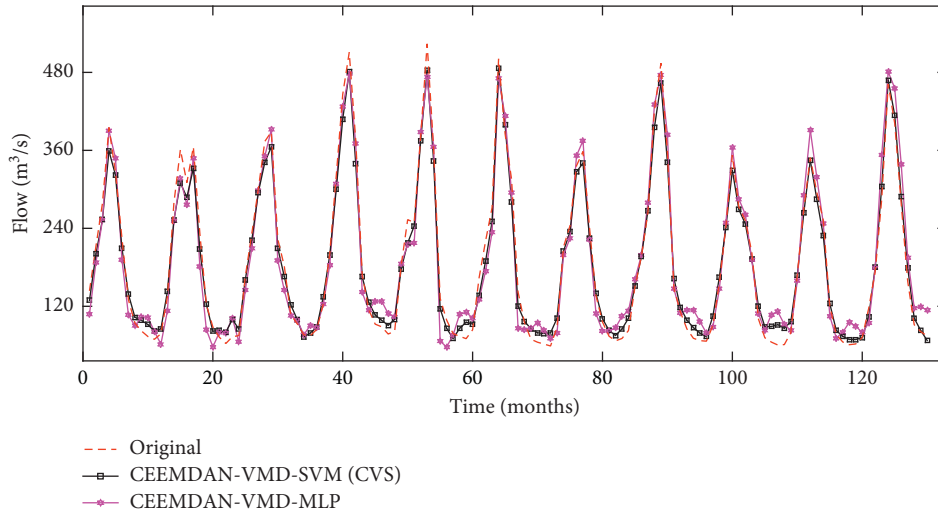
FIGURE 10: Observed and predicted runoff during the model testing period using CVS and CEEMDAN-VMD-MLP models.
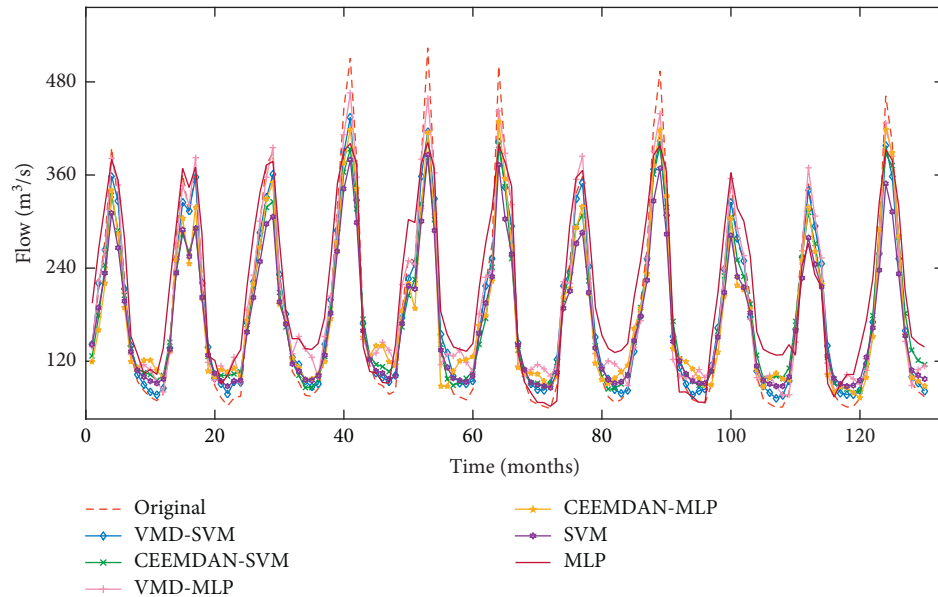

FIGURE 11: Observed and predicted runoff during the model testing period using VMD-SVM, CEEMDAN-SVM, VMD-MLP, and CEEMDAN-MLP models.

models was significantly improved in hybrid combinations, which reflects the significance of the hybrid models.

Based on the calculation method in [81], the CVS models perform better during the training period by reducing RMSE by 71.28%, MAE by 76.22%, MAPE by 80.14%, and MSE by 91.77% compared with the MLP model; RMSE by 40.06%, MAE by 41.15%, MAPE by 22.50%, and MSE by 64.10% compared with the CEEDMAD-VMD-SVM model. However, during the testing period, the error reductions include RMSE by 68.37%, MAE by 72.16%, MAPE by 77.91%, and MSE by 90.03% compared with the MLP model; RMSE by 35.33%, MAE by 35.43%, MAPE by 36.02%, and MSE by 52.28% compared with the CEEDMAD-VMD-SVM model.

The results of figures and tables highlight that the three-layer CEEMDAN-VMD-MLP model also performs better than two-layer models (CEEMDAN-MLP and VMD-MLP); however, its performance is inferior to the two-stage VMD-SVM hybrid model in error reduction and accuracy. Furthermore, all the hybrid models show better performance than the standalone models with direct prediction. The results highlight that the decomposition-based ensemble models are better than standalone ML models since the decomposition approach decomposes the complex input signal into simple-to-study subcomponents, which are favorable to predict and analyze. It can also be concluded (Tables 2 and 3) that the VMD technique is superior to the CEEMDAN technique since the relevant VMD-based
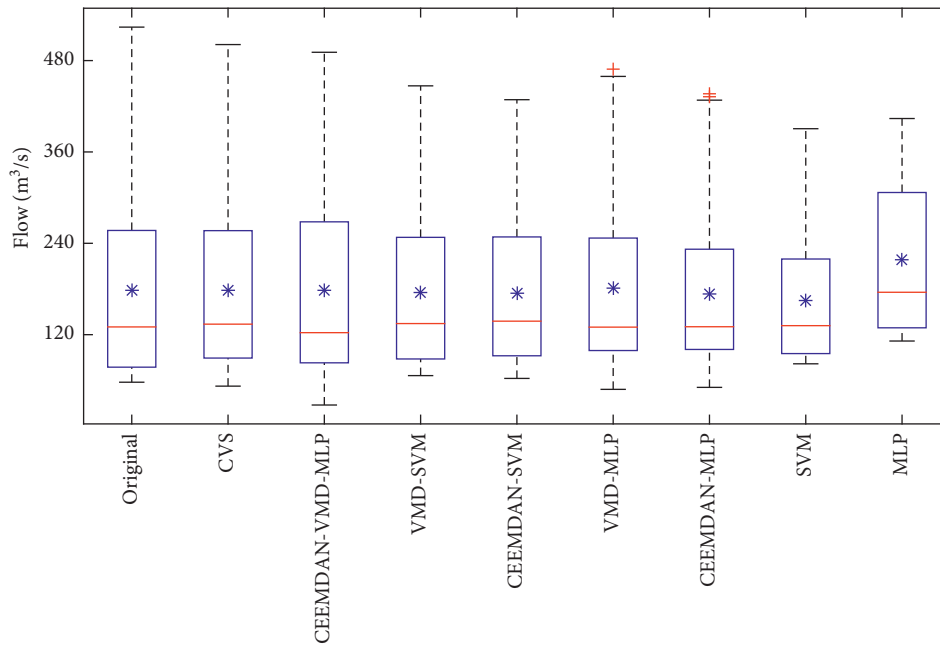
FIGURE 12: Boxplot of runoff during the training period for hybrid models and the benchmark model. The red line marks the median in addition to the mean (blue star).
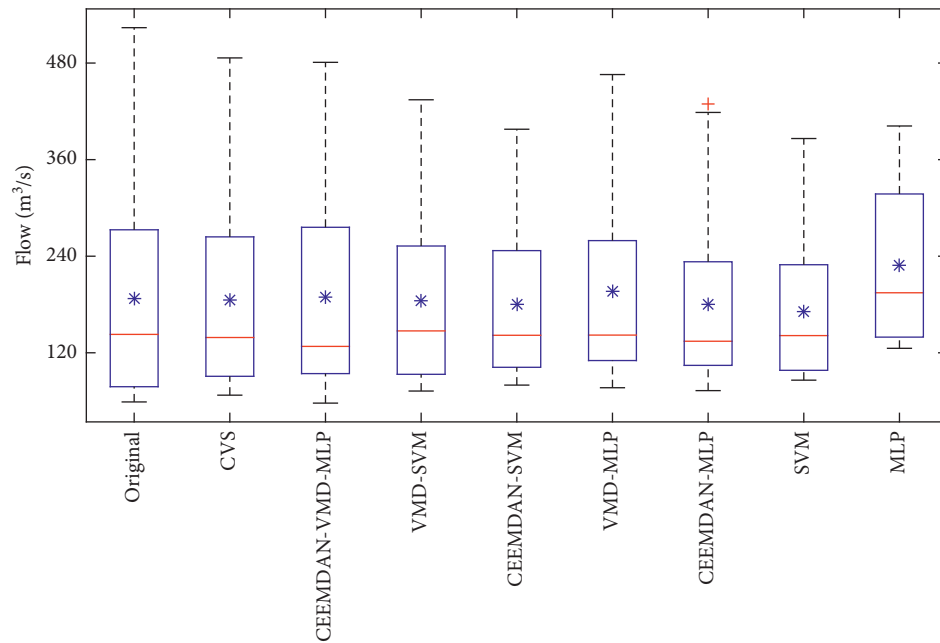

FIGURE 13: Boxplot of runoff during the testing period for hybrid models and the benchmark model. The red line marks the median in addition to the mean (blue star).

models (VMD-MLP and VMD-SVM) perform better than CEEMDAN-based models (CEEMDAN-MLP and CEEM-DAN-SVM).

3.3.1. Extreme Value Analysis. Figure 15 shows the performance of different models in predicting the extreme values of observed runoff during the training and testing periods. The three-stage hybrid models show a superior capability to predict the extreme values of runoff, compared with all the other models, during the training and testing periods. The CVS model shows the best prediction results, while the MLP model in predicting the extreme values of runoff shows the poorest results. Furthermore, the two-stage hybrid models also perform relatively better than the standalone models. All the models show better performance
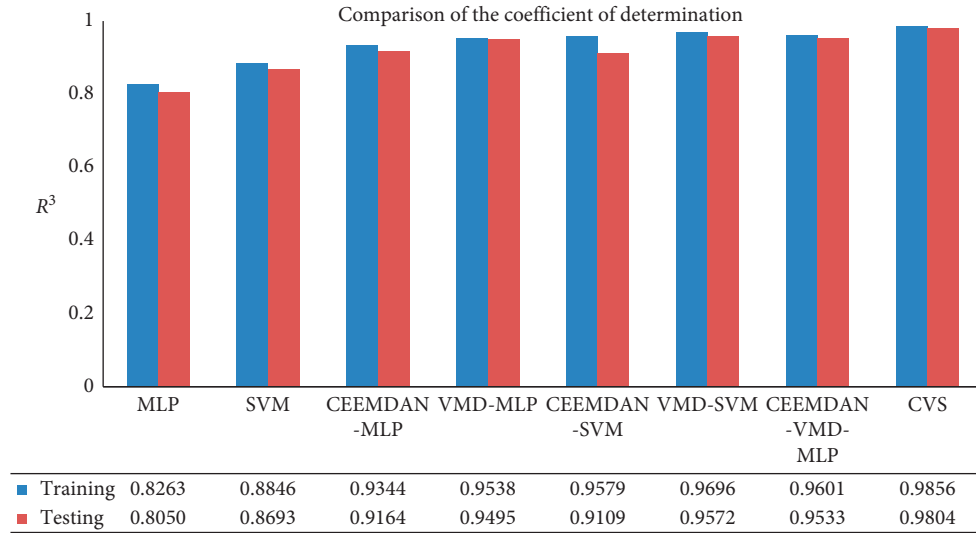
Figure 14: Comparison of the coefficient of determination of CVS with other models.

	MLP	SVM	CEEMDAN-MLP	VMD-MLP	CEEMDAN-SVM	VMD-SVM	CEEMDAN-VMD-MLP	CVS
Training	0.8263	0.8846	0.9344	0.9538	0.9579	0.9696	0.9601	0.9856
Testing	0.8050	0.8693	0.9164	0.9495	0.9109	0.9572	0.9533	0.9804

Table 2: Comparison of the evaluation metrics of the CVS model with other models in the training period.

Model	RMSE	MAE	MAPE	MSE
MLP	0.4126	0.3957	1.1065	0.1702
SVM	0.3363	0.2516	0.4146	0.1131
CEEMDAN-MLP	0.2535	0.2029	0.4105	0.0643
VMD-MLP	0.2127	0.1739	0.3968	0.0452
CEEMDAN-SVM	0.2031	0.1514	0.3469	0.0413
VMD-SVM	0.1726	0.1252	0.2383	0.0297
CEEMDAN-VMD-MLP	0.1977	0.1599	0.2836	0.0390
CVS	0.1185	0.0941	0.2198	0.0140

Table 3: Comparison of the evaluation metrics of the CVS model with other models in the testing period.

Model	RMSE	MAE	MAPE	MSE
MLP	0.4578	0.4282	1.1896	0.2096
SVM	0.3748	0.2830	0.4281	0.1404
CEEMDAN-MLP	0.2997	0.2531	0.4251	0.0898
VMD-MLP	0.2330	0.1868	0.4249	0.0542
CEEMDAN-SVM	0.2594	0.2392	0.4132	0.0957
VMD-SVM	0.2145	0.1508	0.2049	0.0460
CEEMDAN-VMD-MLP	0.2239	0.1846	0.4107	0.0501
CVS	0.1448	0.1192	0.2628	0.0209

in predicting the extreme values of the maximum and minimum runoff in the training period compared with the testing period. Taking the example of the maximum runoff during the training period, the CVS model, CEEMDAN-VMD-MLP, VMD-SVM, CEEMDAN-SVM, VMD-MLP, CEEMDAN-MLP, SVM, and MLP model show an error of 4.47%, 5.45%, 19.67%, 24.03%, 13.78%, 23.45%, 27.29%, and 28.46%, respectively in predicting the observed runoff; however, these models show an error of 6.89%, 7.19%, 27.45%, 33.53, 16.22%, 24.70%, 29.96%, and 31.40%, respectively, during the testing period. Therefore, the proposed model shows a comparatively satisfactory performance in tracking the extreme values of runoff compared with all the other models.

Albeit the runoff process is a complex task for prediction, all the hybrid models generally performed well in all simulations. The results prove the findings of [86–88] according to which, it is impossible practically for a single model to predict precisely the complex hydrological runoff due to the effects of the external factors. The superiority of the proposed hybrid approach proves the viability of decomposition and ML-coupled hybrid approach for hydrological prediction and can also provide a feasible practical reference for similar prediction tasks. The CVS model can identify the intricate nonlinear connection between the original runoff data and the prediction with the best accuracy and performance. Nevertheless, the performance of the model is highly dependent on the reliability of hydrological time
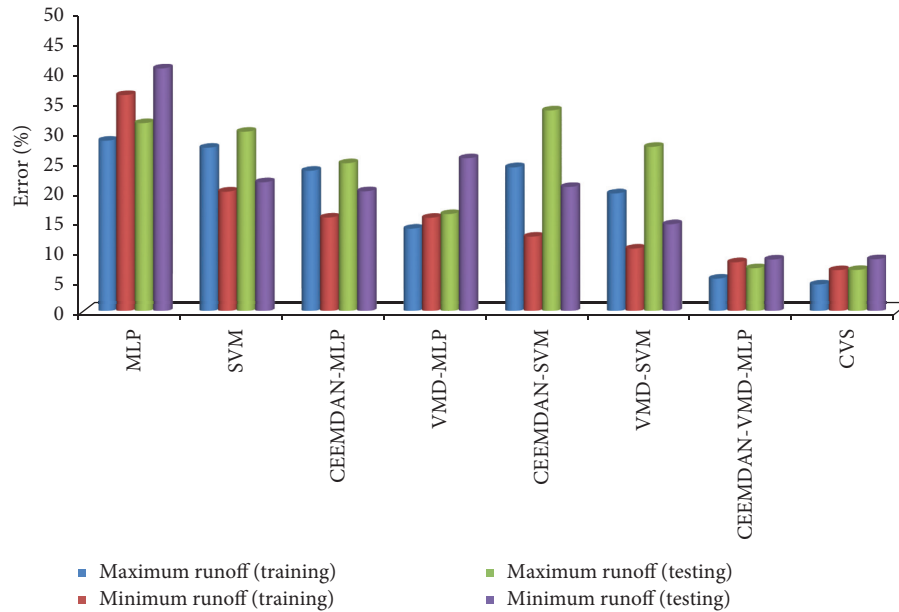
FIGURE 15: Models' performance in predicting the extreme values of observed runoff during the training and testing periods.

series data, mode selection by VMD, and hyperparameter selection by the SVM algorithm. This study deals with monthly runoff prediction by utilizing decomposition-based ML models. However, it is also essential to explore the performances of the CEEMDAN-VMD-based ML models on a daily, weekly, and annual basis for effective management of river basin, reservoir operation and planning, and allocation of water resources. Furthermore, the segregation of hydrological data in normal, drought, and wet periods and maneuver over the performance of models in each period also provides an effective approach for runoff prediction. Moreover, this study has the limitation that it considers only runoff as a predictor for runoff modeling without considering the runoff factors (groundwater flow, surface, and subsurface factors), hydrophysical factors (infiltration, evaporation, etc.), and factors due to humans. Consequently, the authors suggest the implementation of advanced techniques in the future to deal with the limitations of the existing study in a more useful way for reliable hydrological runoff studies. It is expected that this research will provide new directions to study hydrological time series prediction, which will be useful for scientific and technical communities.

4. Conclusions

This paper proposes a three-stage hybrid prediction model, linking the robustness of CEEMDAN-VMD with the SVM algorithm to enhance prediction accuracy with the lowest prediction error for the hydrological runoff time series. Five hybrid models and two standalone models were also used as a benchmark comparison. The models were developed by taking the runoff data of the Swat River, Pakistan. Four statistical performance assessment measures were employed to assess the performance of various models. Considering the results of the prediction accuracy and the error reduction, the following can be concluded from this research work, regarding runoff time series prediction:

(i) Three-stage hybrid models (CVS and CEEMDAN-VMD-MLP) coupling a two-stage signal decomposition methodology (CEEMDAN-VMD) with ML techniques (MLP and SVM) perform better than the two-stage hybrid (CEEMDAN-SVM, VMD-SVM, CEEMDAN-MLP, and VMD-MLP) and standalone models (MLP and SVM).

(ii) CVS model showed superior performance than all the other models in training and testing periods. The suspicions and projecting inaccuracies associated with the proposed CVS model were relatively less than the other models, which endorse the significance of the proposed model for the hydrological runoff prediction.

(iii) Two-stage hybrid models combining single-stage signal decomposition methodology with ML techniques exhibit superior performance than standalone models.

(iv) ML techniques (SVM and MLP) are applicable to predict the runoff time series, and the SVM algorithm is superior to MLP.

(v) Both signal decomposition techniques (VMD and CEEMDAN) significantly improved the prediction results, showing that both techniques apply to the complex, noisy, and nonstationary runoff time series. VMD has shown better performance than CEEMDAN in all cases.

(vi) Limitations: the quantity and quality of the available data play a significant role in the prediction task, and it is not easy to meet this requirement. ML techniques are sensitive to parameter and

hyperparameter selection. Furthermore, ML techniques lack physical relations and concepts, which add complexity in the structuring of ML models.

(vii) Significance and future directions: the research is vital to manage the study area with higher-order trends and noises. The error criteria determine the results of the performance evaluation, and the study judged the performance by employing the well-known performance measures; the superior results indicate the suitability of the CVS model for prediction purposes. Moreover, all the hybrid models showed better performance than standalone models. Therefore, hybrid models combining decomposition techniques with ML methods can play a role in the forthcoming prediction studies. The financial, societal, and ecological benefits of the precise runoff prediction sound for further enhancements in the runoff prediction; therefore, future research will consider new approaches based on deep learning models to study the nonlinear connections among runoff, temperature, climate condition, and precipitation.

Data Availability

The runoff data of Swat River, Pakistan, used to support the findings of this study are included within the article. The data are also available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] C. M. Bürger, O. Kolditz, H. J. Fowler, and S. Blenkinsop, "Future climate scenarios and rainfall-runoff modelling in the Upper Gallego catchment (Spain)," *Environmental Pollution*, vol. 148, no. 3, pp. 842–854, 2007.

[2] C. Zhou, N. Sun, L. Chen et al., "Optimal operation of cascade reservoirs for flood control of multiple areas downstream: a case study in the upper yangtze river basin," *Water*, vol. 10, no. 9, p. 1250, 2018.

[3] M. Rezaie-Balf, S. Fani Nowbandegani, S. Z. Samadi, H. Fallah, and S. Alaghmand, "An ensemble decomposition-based artificial intelligence approach for daily streamflow prediction," *Water*, vol. 11, no. 4, p. 709, 2019.

[4] H. Singh and A. Sankarasubramanian, "Systematic uncertainty reduction strategies for developing streamflow forecasts utilizing multiple climate models and hydrologic models," *Water Resources Research*, vol. 50, no. 2, pp. 1288–1307, 2014.

[5] J. Adamowski and K. Sun, "Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds," *Journal of Hydrology*, vol. 390, no. 1-2, pp. 85–91, 2010.

[6] C. Di, X. Yang, and X. Wang, "A four-stage hybrid model for hydrological time series forecasting," *PLoS One*, vol. 9, no. 8, 2014.

[7] M. Tayyab, J. Zhou, R. Adnan, C. Meng, and A. Zahra, "Streamflow prediction by applying generalized regression network with time series decomposition method," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 4, no. 3, pp. 611–616, 2016.

[8] C. Wu and K. W. Chau, "Rainfall–runoff modeling using artificial neural network coupled with singular spectrum analysis," vol. 399, no. 3-4, pp. 394–409, 2011.

[9] C. Wu, K. Chau, and C. Fan, "Prediction of rainfall time series using modular artificial neural networks coupled with data-preprocessing techniques," vol. 389, no. 1-2, pp. 146–167, 2010.

[10] S. A. Akrami, A. El-Shafie, and O. Jaafar, "Improving rainfall forecasting efficiency using modified adaptive neuro-fuzzy inference system (MANFIS)," *Water Resources Management*, vol. 27, no. 9, pp. 3507–3523, 2013.

[11] G. B. Sahoo, S. G. Schladow, and J. E. Reuter, "Forecasting stream water temperature using regression analysis, artificial neural network, and chaotic non-linear dynamic models," *Journal of Hydrology*, vol. 378, no. 34, pp. 325–342, 2009.

[12] H. Chu, J. Wei, J. Li, Z. Qiao, and J. Cao, "Improved medium- and long-term runoff forecasting using a multimodel approach in the yellow river headwaters region based on large-scale and local-scale climate information," *Water*, vol. 9, no. 8, p. 608, 2017.

[13] X.-h. Zhao and X. Chen, "Auto regressive and ensemble empirical mode decomposition hybrid model for annual runoff forecasting," *Water Resources Management*, vol. 29, no. 8, pp. 2913–2926, 2015.

[14] X. Zhang, Q. Zhang, G. Zhang, Z. Nie, and Z. Gui, "A hybrid model for annual runoff time series forecasting using elman neural network with ensemble empirical mode decomposition," *Water*, vol. 10, no. 4, p. 416, 2018.

[15] O. Kisi, "Streamflow forecasting and estimation using least square support vector regression and adaptive neuro-fuzzy embedded fuzzy c-means clustering," *Water Resources Management*, vol. 29, no. 14, pp. 5109–5127, 2015.

[16] A. Shalamu, *Monthly and Seasonal Streamflow Forecasting in the Rio Grande Basin*, New Mexico State University, Las Cruces, New Mexico, 2009.

[17] Y. Seo, S. Kim, and V. Singh, "Machine learning models coupled with variational mode decomposition: a new approach for modeling daily rainfall-runoff," *Atmosphere*, vol. 9, no. 7, p. 251, 2018.

[18] J. Abbot, J. Marohasy, and A. Research, "Input selection and optimisation for monthly rainfall forecasting in Queensland, Australia, using artificial neural networks," *Atmospheric Research*, vol. 138, pp. 166–178, 2014.

[19] Z. M. Yaseen, A. El-Shafie, O. Jaafar, H. A. Afan, and K. N. Sayl, "Artificial intelligence based models for stream-

flow forecasting: 2000-2015," *Journal of Hydrology*, vol. 530, pp. 829–844, 2015.

[20] F. Fahimi, Z. M. Yaseen, and A. El-shafie, "Application of soft computing based hybrid models in hydrological variables modeling: a comprehensive review," *Theoretical and Applied Climatology*, vol. 128, no. 3-4, pp. 875–903, 2017.

[21] F. Fotovatikhah, M. Herrera, S. Shamshirband, K.-w. Chau, S. Faizollahzadeh Ardabili, and M. J. Piran, "Survey of computational intelligence as basis to big flood management: challenges, research directions and future work," *Engineering Applications of Computational Fluid Mechanics*, vol. 12, no. 1, pp. 411–437, 2018.

[22] Y. Gong, Y. Zhang, S. Lan, and H. Wang, "A comparative study of artificial neural networks, support vector machines and adaptive neuro fuzzy inference system for forecasting groundwater levels near lake okeechobee, Florida," *Water Resources Management*, vol. 30, no. 1, pp. 375–391, Florida, 2016.

[23] M. Dehghani, B. Saghafian, F. Nasiri Saleh, A. Farokhnia, and R. Noori, "Uncertainty analysis of streamflow drought forecast using artificial neural networks and Monte-Carlo simulation," *International Journal of Climatology*, vol. 34, no. 4, pp. 1169–1180, 2014.

[24] F. Granata, R. Gargano, and G. de Marinis, "Support vector regression for rainfall-runoff modeling in urban drainage: a comparison with the EPA's storm water management model," *Water*, vol. 8, no. 3, p. 69, 2016.

[25] M. S. Gizaw and T. Y. Gan, "Regional flood frequency analysis using support vector regression under historical and future climate," *Journal of Hydrology*, vol. 538, pp. 387–398, 2016.

[26] W.-C. Hong, "Chaotic particle swarm optimization algorithm in a support vector regression electric load forecasting model," *Energy Conversion and Management*, vol. 50, no. 1, pp. 105–117, 2009.

[27] C. Voyant, M. Muselli, C. Paoli, and M.-L. Nivet, "Numerical weather prediction (NWP) and hybrid ARMA/ANN model to predict global radiation," *Energy*, vol. 39, no. 1, pp. 341–355, 2012.

[28] B. Z. Shen, H. F. Shen, B. C. Liu et al., "Water modelling of level fluctuation in thin slab continuous casting mould," *Ironmaking & Steelmaking*, vol. 36, no. 1, pp. 33–38, 2009.

[29] H. Nie, G. Liu, X. Liu, and Y. Wang, "Hybrid of ARIMA and SVMs for short-term load forecasting," *Energy Procedia*, vol. 16, pp. 1455–1460, 2012.

[30] G.-F. Fan, L.-L. Peng, W.-C. Hong, and F. Sun, "Electric load forecasting by the SVR model with differential empirical mode decomposition and auto regression," *Neurocomputing*, vol. 173, pp. 958–970, 2016.

[31] W. Su, Z. Lei, L. Yang, and Q. Hu, "Mold-level prediction for continuous casting using VMD-SVR," *Metals*, vol. 9, no. 4, p. 458, 2019.

[32] A. Mosavi, P. Ozturk, and K.-w. Chau, "Flood prediction using machine learning models: literature review," *Water*, vol. 10, no. 11, p. 1536, 2018.

[33] K. Doycheva, G. Horn, C. Koch, A. Schumann, and M. König, "Assessment and weighting of meteorological ensemble forecast members based on supervised machine learning with application to runoff simulations and flood warning," *Advanced Engineering Informatics*, vol. 33, pp. 427–439, 2017.

[34] S. Kim and V. P. Singh, "Flood forecasting using neural computing techniques and conceptual class segregation," *JAWRA Journal of the American Water Resources Association*, vol. 49, no. 6, pp. 1421–1435, 2013.

[35] M. Rezaeian-Zadeh, H. Tabari, and H. Abghari, "Prediction of monthly discharge volume by different artificial neural network algorithms in semi-arid regions," *Arabian Journal of Geosciences*, vol. 6, no. 7, pp. 2529–2537, 2013.

[36] S. Araghinejad, M. Azmi, and M. Kholghi, "Application of artificial neural network ensembles in probabilistic hydrological forecasting," vol. 407, no. 1–4, pp. 94–104, 2011.

[37] T. G. Dietterich, "Ensemble methods in machine learning," in *International Workshop on Multiple Classifier Systems*, Springer, Berlin, Germany, 2000.

[38] F. Sajedi-Hosseini, A. Malekian, B. Choubin et al., "A novel machine learning-based approach for the risk assessment of nitrate groundwater contamination," *Science of The Total Environment*, vol. 644, pp. 954–962, 2018.

[39] J. Zhang, G. Hou, B. Ma, and W. Hua, "Operating characteristic information extraction of flood discharge structure based on complete ensemble empirical mode decomposition with adaptive noise and permutation entropy," *Journal of Vibration and Control*, vol. 24, no. 22, pp. 5291–5301, 2018.

[40] Q. Ouyang, W. Lu, X. Xin, Y. Zhang, W. Cheng, and T. Yu, "Monthly rainfall forecasting using EEMD-SVR based on phase-space reconstruction," *Water Resources Management*, vol. 30, no. 7, pp. 2311–2325, 2016.

[41] M. Rezaie-Balf, O. Kisi, and L. H. C. Chua, "Application of ensemble empirical mode decomposition based on machine learning methodologies in forecasting monthly pan evaporation," *Hydrology Research*, vol. 50, no. 2, pp. 498–516, 2019.

[42] Y. Seo, S. Kim, V. Singh et al., "Comparison of different heuristic and decomposition techniques for river stage modeling," vol. 190, no. 7, p. 392, 2018.

[43] X. Xu and W. Ren, "A hybrid model based on a two-layer decomposition approach and an optimized neural network for chaotic time series prediction," *Symmetry*, vol. 11, no. 5, p. 610, 2019.

[44] Y.-F. Sang, Z. Wang, and C. Liu, "Discrete wavelet-based trend identification in hydrologic time series," *Hydrological Processes*, vol. 27, no. 14, pp. 2021–2031, 2013.

[45] W. Wang and J. J. N. Ding, "Wavelet network model and its application to the prediction of hydrology," *Nature and Science of Sleep*, vol. 1, no. 1, pp. 67–71, 2003.

[46] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, pp. 903–995, 1971.

[47] Y.-F. Sang, Z. Wang, and C. Liu, "Period identification in hydrologic time series using empirical mode decomposition and maximum entropy spectral analysis," *Journal of Hydrology*, vol. 424-425, pp. 154–164, 2012.

[48] Y.-F. Sang, Z. Wang, and C. Liu, "Comparison of the MK test and EMD method for trend identification in hydrological time series," *Journal of Hydrology*, vol. 510, pp. 293–298, 2014.

[49] Y. Huang, F. G. Schmitt, Z. Lu, and Y. Liu, "Analysis of daily river flow fluctuations using empirical mode decomposition and arbitrary order Hilbert spectral analysis," *Journal of Hydrology*, vol. 373, no. 1-2, pp. 103–111, 2009.

[50] X. Li and Z. Ding, "Emd method for multiple time-scale analysis on fluctuation characteristic of natural annual runoff time series of fen river," *Water Resources and Power*, vol. 26, no. 1, pp. 30–32, 2008.

[51] Z. Wu and N. E. Huang, "A study of the characteristics of white noise using the empirical mode decomposition method," *Proceedings of the Royal Society of London. Series A:*

Mathematical, Physical and Engineering Sciences, vol. 460, no. 2046, pp. 1597–1611, 2004.

[52] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 1, pp. 1–41, 2009.

[53] M. Firdaus Isham, M. Salman Leong, M. Hee Lim et al., "Variational mode decomposition: mode determination method for rotating machinery diagnosis," *Journal of Vibroengineering*, vol. 20, no. 7, 2018.

[54] H. S. Lee, "Improvement of decomposing results of empirical mode decomposition and its variations for sea-level records analysis," *Journal of Coastal Research*, vol. 85, no. sp1, pp. 526–530, 2018.

[55] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2013.

[56] P. Shi and W. Yang, "Precise feature extraction from wind turbine condition monitoring signals by using optimised variational mode decomposition," *IET Renewable Power Generation*, vol. 11, no. 3, pp. 245–252, 2016.

[57] M. E. Torres, M. A. Colominas, G. Schlotthauer et al., "A complete ensemble empirical mode decomposition with adaptive noise," in *Proceedings of the 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, Czech Republic, May 2011.

[58] Y. Wang, R. Markert, J. Xiang, and W. Zheng, "Research on variational mode decomposition and its application in detecting rub-impact fault of the rotor system," *Mechanical Systems and Signal Processing*, vol. 60-61, pp. 243–251, 2015.

[59] X. Zhang, Q. Zhang, G. Zhang, Z. Nie, Z. Gui, and H. Que, "A novel hybrid data-driven model for daily land surface temperature forecasting using long short-term memory neural network based on ensemble empirical mode decomposition," *International Journal of Environmental Research and Public Health*, vol. 15, no. 5, p. 1032, 2018.

[60] M. Niu, Y. Hu, S. Sun, and Y. Liu, "A novel hybrid decomposition-ensemble model based on VMD and HGWO for container throughput forecasting," *Applied Mathematical Modelling*, vol. 57, pp. 163–178, 2018.

[61] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intelligent Systems and Their Applications*, vol. 13, no. 4, pp. 18–28, 1998.

[62] V. Vapnik and S. Mukherjee, "Support vector method for multivariate density estimation," in *Advances in Neural Information Processing Systems*, MIT Press, Cambridge, MA, USA, 2000.

[63] Q. Wen, W. Mu, L. Sun et al., "Daily sales forecasting for grapes by support vector machine," in *Proceedings of the International Conference on Computer and Computing Technologies in Agriculture*, Springer, Beijing, China, September 2013.

[64] P.-S. Yu, S.-T. Chen, and I. Chang, "Support vector regression for real-time flood stage forecasting," vol. 328, no. 3-4, pp. 704–716, 2006.

[65] P. Kang, D. Kim, and S. Cho, "Semi-supervised support vector regression based on self-training with label uncertainty: an application to virtual metrology in semiconductor manufacturing," *Expert Systems with Applications*, vol. 51, pp. 85–106, 2016.

[66] S. M. Clarke, J. H. Griebsch, and T. W. Simpson, "Analysis of support vector regression for approximation of complex engineering analyses," *Journal of Mechanical Design*, vol. 127, no. 6, pp. 1077–1087, 2005.

[67] C.-H. Wu, G.-H. Tzeng, and R.-H. Lin, "A Novel hybrid genetic algorithm for kernel function and parameter optimization in support vector regression," *Expert Systems with Applications*, vol. 36, no. 3, pp. 4725–4735, 2009.

[68] W. Wang, P. H. Van Gelder, J. Vrijling et al., "Forecasting daily streamflow using hybrid ANN models," vol. 324, no. 1–4, pp. 383–399, 2006.

[69] Z. Ali, I. Hussain, M. Faisal et al., "Forecasting drought using multilayer perceptron artificial neural network model," *Advances in Meteorology*, vol. 2, 2017.

[70] J. Sietsma, "Neural net pruning-why and how," in *Proceedings of the International Conference on Neural Networks*, San Diego, CA, USA, 1988.

[71] A. Roy, L. S. Kim, and S. Mukhopadhyay, "A polynomial time algorithm for the construction and training of a class of multilayer perceptrons," *Neural Networks*, vol. 6, no. 4, pp. 535–545, 1993.

[72] N. Murata, S. Yoshizawa, and S. Amari, "Network information criterion-determining the number of hidden units for an artificial neural network model," *IEEE Transactions on Neural Networks*, vol. 5, no. 6, pp. 865–872, 1994.

[73] Z. Wang, C. Di Massimo, M. T. Tham, and A. Julian Morris, "A procedure for determining the topology of multilayer feedforward neural networks," *Neural Networks*, vol. 7, no. 2, pp. 291–300, 1994.

[74] P. Zhang, Y. Jia, J. Gao et al., "Short-term rainfall forecasting using multi-layer perceptron," *IEEE Transactions on Big Data*, vol. 6, no. 1, 2018.

[75] K. Gehmlich, S. Lange, A. S. Lun et al., "MLP and CARP are linked to chronic PKCα signalling in dilated cardiomyopathy," *Nature Communications*, vol. 7, no. 1, 2016.

[76] Y. N. Dauphin, R. Pascanu, C. Gulcehre et al., "Identifying and attacking the saddle point problem in high-dimensional non-convex optimization," in *Advances in Neural Information Processing Systems*, MIT Press, Cambridge, MA, USA, 2014.

[77] F. M. Bianchi, E. Maiorino, M. C. Kampffmeyer et al., "An overview and comparative analysis of recurrent neural networks for short term load forecasting," 2017, http://arxiv.org/abs/1705.04378.

[78] J. Moolayil and S. John, *Learn Keras for Deep Neural Networks*, Springer, Berlin, Germany, 2019.

[79] Z.-k. Feng, W.-j. Niu, C.-t. Cheng et al., "Optimizing hydropower reservoirs operation via an orthogonal progressive optimality algorithm," *Journal of Water Resources Planning and Management*, vol. 144, no. 3, Article ID 04018001, 2018.

[80] B. Xu, P.-A. Zhong, R. C. Zambon, Y. Zhao, and W. W.-G. Yeh, "Scenario tree reduction in stochastic programming with recourse for hydropower operations," *Water Resources Research*, vol. 51, no. 8, pp. 6359–6380, 2015.

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

[82] I. Ahmad, D. Tang, T. Wang et al., "Precipitation trends over time using Mann-Kendall and Spearman's rho tests in Swat River Basin, Pakistan," *Advances in Meteorology*, vol. 2015, Article ID 431860, 15 pages, 2015.

[83] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, 2014.

[84] C. W. Dawson and R. L. Wilby, "Hydrological modelling using artificial neural networks," *Progress in Physical Geography: Earth and Environment*, vol. 25, no. 1, pp. 80–108, 2001.

[85] D. P. Kingma and Jimmy Ba, "Adam: a method for stochastic optimization," 2014, https://arxiv.org/abs/1412.6980.

[86] Z.-k. Feng, W.-j. Niu, W.-c. Wang, J.-z. Zhou, and C.-t. Cheng, "A mixed integer linear programming model for unit commitment of thermal plants with peak shaving operation aspect in regional power grid lack of flexible hydropower energy," *Energy*, vol. 175, pp. 618–629, 2019.

[87] T. Zhao, J. Zhao, P. Liu, and X. Lei, "Evaluating the marginal utility principle for long-term hydropower scheduling," *Energy Conversion and Management*, vol. 106, pp. 213–223, 2015.

[88] Z.-k. Feng, W.-j. Niu, C.-t. Cheng, and X.-y. Wu, "Optimization of hydropower system operation by uniform dynamic programming for dimensionality reduction," *Energy*, vol. 134, pp. 718–730, 2017.