Check for updates

Monthly Runoff Forecasting Using Variational Mode Decomposition Coupled with Gray Wolf Optimizer-Based Long Short-term Memory Neural Networks

Bao-Jian Li[1,2] ⓘ · **Guo-Liang Sun**[1] · **Yan Liu**[3] · **Wen-Chuan Wang**[1,2] · **Xu-Dong Huang**[1,2]

Abstract

Accurate and reliable monthly runoff forecasting plays an important role in making full use of water resources. In recent years, long short-term memory neural networks (LSTM), as a deep learning technology, has been successfully applied in forecasting monthly runoff. However, the hyperparameters of LSTM are predetermined, which has a significant influence on model performance. In this study, given that the decomposition of monthly runoff series may provide a more accurate prediction, as revealed by many previous studies, a hybrid model, namely, VMD-GWO-LSTM, is proposed for monthly runoff forecasting. The proposed hybrid model comprises two main components, namely, variational mode decomposition (VMD) coupled with the gray wolf optimizer (GWO)-based LSTM. First, VMD is utilized to decompose raw monthly runoff series into several subsequences. Second, GWO is implemented to optimize the hyperparameters of the LSTM for each subsequence on the condition that the inputs are determined. Finally, the total output of all subsequences is aggregated as the final forecast result. Four quantitative indices are employed to evaluate the model performance. The proposed model is demonstrated using 73 and 62 years of monthly runoff series data derived from the Xinfengjiang and Guangzhao Reservoirs in China's Pearl River system, respectively. To identify the feasibility and superiority of the proposed model, backpropagation neural networks (BPNN), support vector machine (SVM), LSTM, EMD-LSTM, VMD-LSTM and GWO-LSTM are also utilized for comparison. The results indicate that the proposed hybrid model can yield best forecast accuracy among these models, making it a promising new method for monthly runoff forecasting.

✉ Bao-Jian Li
libaojian@ncwu.edu.cn

1   College of Water Resources, North China University of Water Resources and Electric Power, Zhengzhou 450046, China

2   Henan Key Laboratory of Water Resources Conservation and Intensive Utilization in the Yellow River Basin, Zhengzhou 450046, China

3   College of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou 450066, China

1 Introduction

Accurate and reliable monthly runoff forecasting plays an important role in water resources management, such as water supply (Şen 2021), hydroelectric generation and ecological restoration. Generally, existing methods can be approximately partitioned into data-driven (Chu et al. 2021; Feng et al. 2021; Liao et al. 2020; Riahi-Madvar et al. 2021) and physical-based models (Abuzied and Mansour 2019; Abuzied and Pradhan 2020; Abuzied et al. 2016; Bournas and Baltas 2021; Budamala and Mahindrakar 2020; El Harraki et al. 2021; Liao et al. 2016). Data-based models can simulate the relationship between input and output without regard to complex mechanisms of runoff generation (Niu et al. 2019). In contrast, physical-based models take into account specific physical process and demand mass data, such as underlying surface conditions, human activity influences and climate change, which are not easily collected (Feng and Niu 2021). Unlike the physical-based models, the data-based models demand less data and can offer satisfactory forecast results. As a typical representative of data-based models, artificial neural networks (ANN) have been widely and successfully utilized in hydrology-related areas, for instance, precipitation forecasting (Nourani et al. 2009), runoff forecasting (Shu et al. 2021; Noorbeh et al. 2020), and water level forecasting (Seo et al. 2015). In recent decades, numerous ANN architectures and algorithms have been investigated in hydrological time series forecasting (ASCE-Task-Committee 2000).

Long short-term memory neural networks (LSTM) proposed by Hochreiter and Schmidhuber (1997) are a special kind of recurrent neural network (RNN) and have the merits of fast convergence and good nonlinear predictive capability. To avoid the problems of training long sequences and vanishing gradients faced by the traditional RNN, LSTM implement constant error flow via constant error carrousels within special memory cells. Referring to LSTM, many studies have been conducted on hydrological time series forecasting (Lv et al. 2020; Ni et al. 2020; Wang et al. 2021). Nevertheless, the hyperparameters of LSTM are predetermined, which has a certain impact on forecast accuracy. In general, there are two main methods to improve the forecast accuracy in previous studies. The first is to combine decomposition algorithms to decompose original time series data into several subcomponents, employ LSTM to simulate each subcomponent, and aggregate the results of each subcomponent as the final result (Lv et al. 2020). Zuo et al. (2020), for example, proposed single-model forecasting based on VMD and LSTM to predict daily streamflow 1–7 days ahead and investigated the robustness and efficiency of the proposed model for forecasting highly nonstationary and nonlinear streamflow. The second is to utilize optimization algorithms to optimize the hyperparameters of the LSTM (ElSaid et al. 2018). Yuan et al. (2018), for example, used the ant lion optimizer (ALO) to calibrate the parameters of the LSTM, and verified its effectiveness with the historical monthly runoff of the Astor River Basin. At present, there are several commonly used decomposition algorithms (Colominas et al. 2014; Roushangar et al. 2021; Shahid et al. 2020); for instance, wavelet decomposition, empirical mode decomposition (EMD) and VMD. Optimization algorithms, such as particle swarm optimization and ant colony optimization, can be seen in the literature as optimizing the parameters of neural networks (Wan et al. 2017; Yu et al. 2008). In this study, both methods are considered.

Variational mode decomposition (VMD) (Dragomiretskiy and Zosso 2014) is an entirely nonrecursive variational model that can extract modes concurrently. Via VMD, a signal can be decomposed into a sequence of subcomponents with different frequency bands and time resolutions (Fang et al. 2019). Compared to empirical mode decomposition

(EMD), VMD is capable of separating tones of similar frequencies. VMD has been widely applied in many research fields, such as fault diagnosis (Zhang et al. 2017), signal processing (Wang et al. 2017), wind speed monitoring (Liu et al. 2018) and hydrological time series forecasting (Feng et al. 2020; Li et al. 2021; Sibtain et al. 2021). In this study, VMD was selected as a data preprocessing tool to decompose monthly runoff series. In recent years, an emerging swarm intelligence algorithm called the gray wolf optimizer (GWO) has been proposed, which imitates the social hierarchy and hunting behavior of gray wolves (Mirjalili et al. 2014). With its strong robustness and searching ability in solving optimization problems, the GWO has been widely and successfully applied in many fields, such as model parameter calibration (Tikhamarine et al. 2020), reservoir operation (Niu et al. 2021) and optimal power dispatch (Nuaekaew et al. 2017). Hence, in view of its strong robustness and searching ability, the GWO can be adopted to optimize the hyperparameters of LSTM.

In this paper, a hybrid model, referred to as the VMD-GWO-LSTM, is proposed for monthly runoff forecasting. According to the monthly runoff series of two real-world hydropower reservoirs in China, the proposed method is certified to be feasible. The innovation of this study can be stated as follows. (1) To decrease modeling difficulty, VMD is adopted to decompose monthly runoff series into several simple subcomponents. (2) For each subcomponent, the input–output relationships are identified by the LSTM, and the GWO method is employed to optimize the hyperparameters of the LSTM. (3) The results of the case study indicate that, compared to several traditional models, the proposed hybrid VMD-GWO-LSTM method can yield better forecast accuracy. To our knowledge, there have been few studies combining VMD, LSTM, and GWO to forecast monthly runoff, demonstrating that this study has the potential to fill this gap.

The rest of this work is organized as follows: Sect. 2 describes the details of the proposed approach; in Sect. 3, the proposed method is utilized to forecast the monthly runoff of two reservoirs; and finally, the conclusions are summarized.

2 Methodology

2.1 Variational Mode Decomposition

VMD is a novel variational method that can nonrecursively decompose a nonstationary signal into a given number of mode functions, and each individual mode is compact around its center frequency (Dragomiretskiy and Zosso 2014). To obtain each mode and its center frequency, a constrained variational problem can be expressed as follows:

$$
\begin{cases}
\min\limits_{\{u_k(t)\},\{\omega_k(t)\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \dfrac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\
s.t. \ \sum_k u_k(t) = f(t)
\end{cases}
\tag{1}
$$

where t is the time step; $u_k(t)$ and $\omega_k(t)$ denote the k-th mode and its corresponding center frequency, respectively; $\delta(t)$ is the Dirac distribution, * denotes the convolution calculation; and $f(t)$ denotes the t-th data of the input signal.

To facilitate the solution, the quadratic penalty factor α and the Lagrangian multiplier λ are introduced to transform the constrained variational problem into an unconstrained variational problem. Hence, the augmented Lagrangian structure can be expressed as follows:

$$
L\left(\{u_k\}, \{\omega_k\}, \lambda\right) = \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2
$$
$$
+ \left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle
$$

(2)

where $\langle \cdot \rangle$ represents the inner product operation.

Equation (2) can then be solved by the alternating direction method of multipliers (ADMM) to obtain the saddle point of the augmented Lagrangian function. In the ADMM, the variables ($\hat{u}_k^{n+1}, \omega_k^{n+1}$ and $\hat{\lambda}^{n+1}$) are continuously updated to optimize each modal component.

2.2 Long Short-term Memory Neural Networks

As a type of deep learning neural networks, the LSTM was proposed to overcome the vanishing/exploding gradient problem faced by traditional RNN (Hochreiter and Schmidhuber 1997). The LSTM takes the place of the conventional hidden unit with a memory cell and contains multiple memory blocks, each of which includes three gates: input gate, forget gate and output gate and at least one memory cell. By using the LSTM, information from the three gates can be added or deleted to the memory cell state. Based on the previous state, current memory and current input, the LSTM has the ability to decide which cells are restrained and promoted and on the basis of the three gates what information is saved and forgotten during the training process (Altan et al. 2021). The structure of the LSTM is shown in Fig. 1. For the three gates, the multiplicative input gate unit is employed to recognize new information that can be gathered in the cell; the multiplicative output gate unit is utilized to compute the information that can be propagated to the network; and the multiplicative forget gate unit is used to decide whether the last status of the cell can be forgotten (Li et al. 2018).

The calculation of the three gates and cell state can be generally expressed as follows:

Fig. 1 Schematic diagram of long short-term memory neural networks
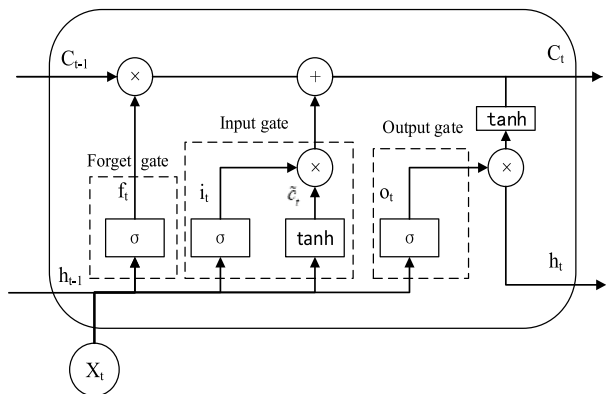
$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right)$$
$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right)$$
$$\tilde{c}_t = \tanh\left(W_c \cdot [h_{t-1}, x_t] + b_c\right)$$
$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t$$
$$o_t = \sigma\left(W_o \odot [h_{t-1}, x_t] + b_0\right)$$
$$h_t = o_t \cdot \tanh\left(c_t\right)$$

$$(3)$$

where f_t, i_t, o_t denote the output of the forget gate, input gate and output gate, respectively; $\tilde{c}_t$ is the potential cell state; c_t and h_t denote the cell state and cell output at time t, respectively; W_f, W_i, W_c W_o and b_f, b_i, b_c, b_0 denote weight matrices and the corresponding bias vectors, respectively; x_t is the input at time t; σ is the sigmoid function; and $\odot$ denotes matrix multiplication.

It is worth noting that the LSTM relies heavily on a set of hyperparameters to achieve good performance, which usually requires a certain amount of practical experience to manually select and optimize the hyperparameters. Therefore, for convenience, automatic algorithmic approaches with the ability to converge faster and gain an optimal/near optimal solution within an acceptable time can be employed to enhance the performance of the LSTM (Nakisa et al. 2018).

2.3 Gray Wolf Optimizer

The GWO algorithm is a novel swarm intelligent optimization algorithm that simulates the leadership hierarchy and predation strategy of gray wolves (Mirjalili et al. 2014). Gray wolves possess a very strict social dominant hierarchy, which can be divided into four categories: alpha wolf (α), beta wolf (β), delta wolf (δ) and omega wolf (ω). The alpha dominates the whole wolf pack and is responsible for making decisions. Beta wolves are subordinate to the alpha in the hierarchy but command delta and omega wolves as well. The hunting process of gray wolves can be divided into three stages: (i) tracking, chasing and approaching prey; (ii) hunting, surrounding and cornering the prey until it stops moving; and (iii) attacking the prey (Mirjalili et al. 2014). The GWO algorithm can be generally described as follows.

First, encircling prey is carried out by the gray wolves before the hunting process, which can be defined as follows:

$$\vec{D} = \left| \vec{C} \cdot \vec{X}_p(t) - \vec{X}(t) \right| \tag{4}$$

$$\vec{X}(t + 1) = \vec{X}(t) - \vec{A} \cdot \vec{D} \tag{5}$$

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \tag{6}$$

$$\vec{C} = 2 \cdot \vec{r}_2 \tag{7}$$

where t is the current iteration; $\vec{D}$ is the distance between the gray wolf and the prey; $\vec{X}_p(t)$ is the position vector of the t-th prey; $\vec{X}(t)$ is the position vector of the t-th gray wolf; $\vec{A}$ and

$\vec{C}$ are coefficient vectors; $\vec{r}_1$ and $\vec{r}_2$ are random numbers in [0,1]; and $\vec{a}$ is a transition parameter that is linearly reduced from 2 to 0 during the iterative computation.

Then, the hunting process is implemented. After recognizing the position of the prey and encircling it, the wolves will hunt the prey, which is guided by the alpha and the beta and delta will occasionally participate. The formulas are given as follows, in which other wolves should be obeyed to update their positions:

$$\begin{cases} \vec{D}_\alpha = \left| \vec{C}_1 \cdot \vec{X}_\alpha - \vec{X} \right| \\ \vec{D}_\beta = \left| \vec{C}_2 \cdot \vec{X}_\beta - \vec{X} \right| \\ \vec{D}_\delta = \left| \vec{C}_3 \cdot \vec{X}_\delta - \vec{X} \right| \end{cases} \tag{8}$$

$$\begin{cases} \vec{X}_1 = \vec{X}_\alpha - A_1 \cdot \vec{D}_\alpha \\ \vec{X}_2 = \vec{X}_\beta - A_2 \cdot \vec{D}_\beta \\ \vec{X}_3 = \vec{X}_\delta - A_3 \cdot \vec{D}_\delta \end{cases} \tag{9}$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \tag{10}$$

Finally, attacking prey is executed. To simulate being close to the prey, the value of $\vec{a}$ is decreased linearly, and correspondingly, the fluctuation range of $\vec{A}$ is also decreased within the interval of [-2a, 2a]. When $\vec{A}$ ranges in [-1, 1], the next position of a gray wolf in any position is between its current position and the position of the prey (Mirjalili et al. 2014). Thus, the attack on the prey can be realized.

2.4 Hybrid Model for Monthly Runoff Forecasting

To improve the forecast accuracy of monthly runoff forecasting, a hybrid model shortened to the VMD-GWO-LSTM is proposed and illustrated in Fig. 2. The main procedure can be described as follows:

Step 1: Data preprocessing. VMD is utilized to decompose the original runoff sequence to obtain K subsequences with different frequencies. All subsequences that are divided into calibration and validation data are normalized to [-1, 1].
Step 2: Input determination. The partial autocorrelation function (PACF) is utilized to determine the input variables of each subsequence for the LSTM model.
Step 3: Hyperparameter optimization. For each LSTM model, the optimal parameters of the number of hidden layer neurons, the number of epochs and the learning rate are searched by the GWO, and root-mean-squared error (RMSE) is selected as the optimization criterion.
Step 4: Aggregation. The forecast results of all subsequences are arithmetically aggregated as the final forecast results.
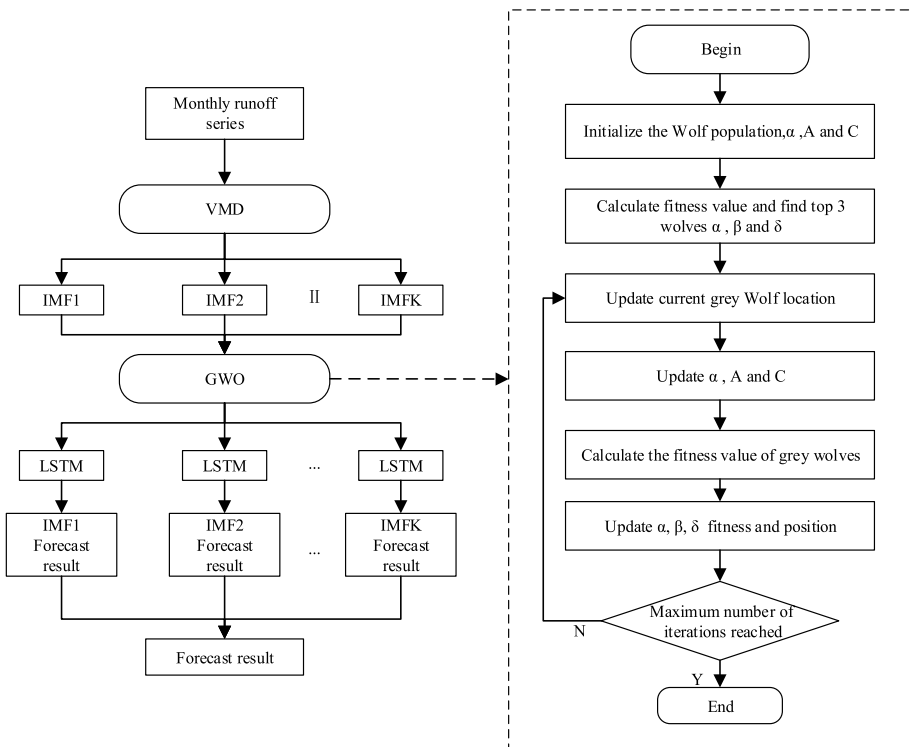
Fig. 2 The flowchart of VMD-GWO-LSTM for monthly runoff forecasting

2.5 Evaluation Index

In this section, four evaluation indices namely, RMSE, mean absolute percentage error (MAPE), coefficient of correlation (R) and Nash–Sutcliffe efficiency coefficient (CE), are employed. Generally, the smaller RMSE and MAPE are and the higher R and CE are, the better the model performance. These indices are listed below:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Q_i - \hat{Q}_i)^2} \qquad (11)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{Q_i - \hat{Q}_i}{Q_i} \right| \times 100 \qquad (12)$$

$$R = \frac{\sum_{i=1}^{n} (Q_i - \overline{Q}_i)(\hat{Q}_i - \overline{\hat{Q}}_i)}{\sqrt{\sum_{i=1}^{n} (Q_i - \overline{Q}_i)^2 \sum_{i=1}^{n} (\hat{Q}_i - \overline{\hat{Q}})^2}} \qquad (13)$$

$$CE = 1 - \frac{\sum\limits_{i=1}^{n}(Q_i - \overline{Q}_i)^2}{\sum\limits_{i=1}^{n}(Q_i - \overline{Q})^2} \tag{14}$$

where n is the number of observed data; Q_i and $\hat{Q}_i$ are the observed and forecasted values, respectively; and $\overline{Q}_i$ and $\hat{Q}_i$ are the averages of all observed and forecasted values, respectively.

3 Case Studies

3.1 Study Area and Data

Two multipurpose reservoirs, the Xinfengjiang and Guangzhao Reservoirs located in China, were selected as case studies. The Xinfengjiang Reservoir is located on the Xinfeng River, which is the largest tributary of the Dongjiang River and a second-level tributary of the Pearl River. Located in Heyuan city, Guangdong Province, the Xinfeng River Basin has a subtropical monsoon climate. With obvious interannual and seasonal variations, the annual average precipitation is 1742.0 mm, of which approximately 76% is accounted for from April to September. The total drainage area of the Xinfeng River Basin is 5813 km^2 and the upstream area of the Xinfengjiang Reservoir is 5740 km^2. With an average gradient of 1.29%, the length of the river is 163 km. Mainly constituted of hills and mountains, the terrain of the Xinfeng River Basin is low in the west and high in the east. For the Xinfeng River Baisn, the mountainous area accounts for 33.6%, and the hilly area accounts for 63.5%. With 336.1 MW of installed capacity and 13.896 billion m^3 of storage volume, the Xinfengjiang Reservoir is the largest reservoir in southern China. For the Xinfengjiang Reservoir, the primary goal is power generation. The Guangzhao Reservoir is located on the middle reaches of the Beipan River, which is a tributary of the Xijiang River and a second-level tributary of the Pearl River. Located on the slope of the Yunnan-Guizhou Plateau, the Beipan River Basin is connected to the hilly basin of central Guizhou Province in the east and has a subtropical plateau monsoon climate. With an obvious seasonal variation, the annual average precipitation is 1178.0 mm, of which approximately 80% is accounted for from May to September. The total drainage area of the Beipan River Basin is 26557 km^2 and the upstream area of the Guangzhao Reservoir is 13548 km^2. With an average gradient of 0.437%, the length of the river is 441.9 km. The terrain of the Beipan River Basin is high in the northwest and low in the southeast. For the Beipan River basin, the mountainous area accounts for 85%, and the hilly area accounts for 10%. With 1040 MW of installed capacity and 3.245 billion m^3 of storage volume, the primary goal of the Guangzhao Reservoir is power generation. Hence, accurate monthly runoff forecasting is vital for these two reservoirs.

Monthly runoff series data from the Xinfengjiang and Guangzhao Reservoirs were retrieved to validate the proposed method. The monthly runoff data for the Xinfengjiang Reservoir cover 1943 to 2015 and the data for the Guangzhao Reservoir cover 1956 to 2017. For these two reservoirs, approximately 70% of the data were used for calibration, and the remaining data were used for validation.

3.2 Decomposition Results

According to VMD, the key parameter of the number of modes should be predefined, which affects the decomposed results (Wen et al. 2019). To obtain satisfactory performance, the traditional EMD method was employed to ascertain the number of subsequences. The decomposed results for the Xinfengjiang Reservoir utilizing VMD and EMD are shown in Fig. 3. There were significant differences in the acquired subcomponents for the two reservoirs, which indicate the variability of VMD and EMD in extracting intrinsic information from the original monthly runoff series.

3.3 Input Determination

The selection of input variables that directly affect the forecast results, should be predetermined. As a statistical method, the partial autocorrelation function (PACF) can be employed to analyze and determine the input variables (Feng et al. 2020; He et al. 2019). In practice, the input variables are often determined by means of the PACF values in which the previous values are selected as inputs when all PACF values fall into the confidence interval. It is worth mentioning that if the number of input variables is too small, the forecast accuracy of the model will be low. Hence, the determination of the input variables is also needed based on experience or other methods. In this study, the input variables were determined by the fact that if the number of input variables was equal to or less than 2 for the first time, PACF values falling back into the confidence interval in the second time
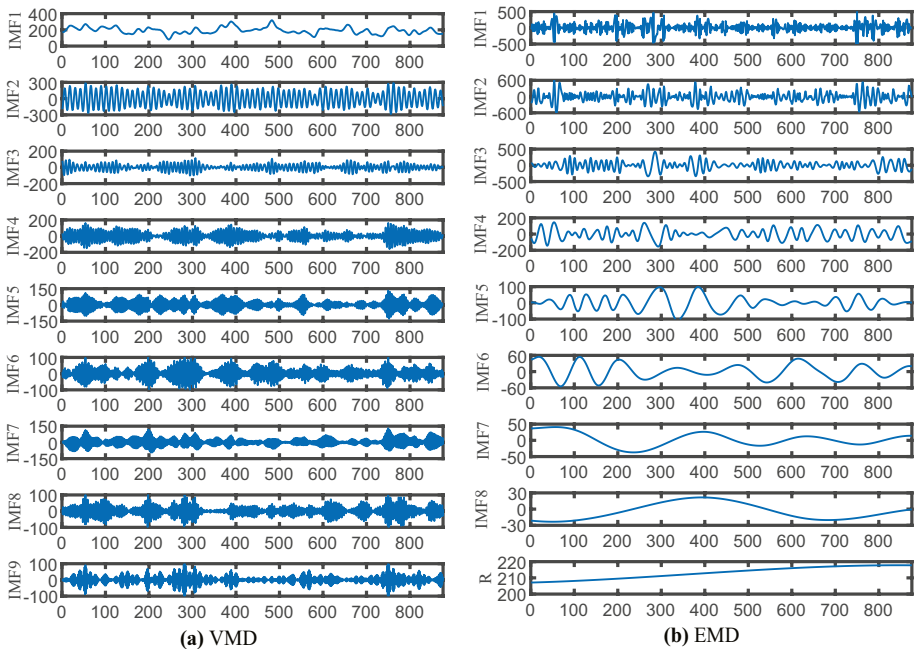

(a) VMD **(b)** EMD

Fig. 3 Decomposed results of monthly runoff data in Xinfengjiang Reservoirs

could be considered. The PACF values for the original and decomposed subsequences of the Xinfengjiang Reservoir data are shown in Fig. 4. On the basis of Fig. 4, the input variables for each sequence of Xinfengjiang Reservoir data could be determined. From Table 1, it can be easily seen that the numbers of input variables for the original and decomposed data are similar but not always the same, indicating the complex and variable features of the data from the two reservoirs.

3.4 Model Development

To confirm the feasibility of the proposed method, five models were employed for comparison, namely, backpropagation neural networks (BPNN), support vector machine (SVM), LSTM, VMD-LSTM and EMD-LSTM models. The details of the models are stated as follows.

1. BPNN, SVM and LSTM models.

   The original monthly runoff data were used to calibrate the parameters of the BPNN, SVM and LSTM models. In this study, the input variables for the three models were set based on PACF values of the original series. For the BPNN model, three layers were employed, the output nodes were set as 1, and the hidden nodes were set by a trial-and-error procedure. For the SVM model, the radial basis function was chosen as the kernel function, and the genetic algorithm was used to optimize the parameters. For the LSTM model, the number of hidden layers was 2, the output nodes were set as 1 and the
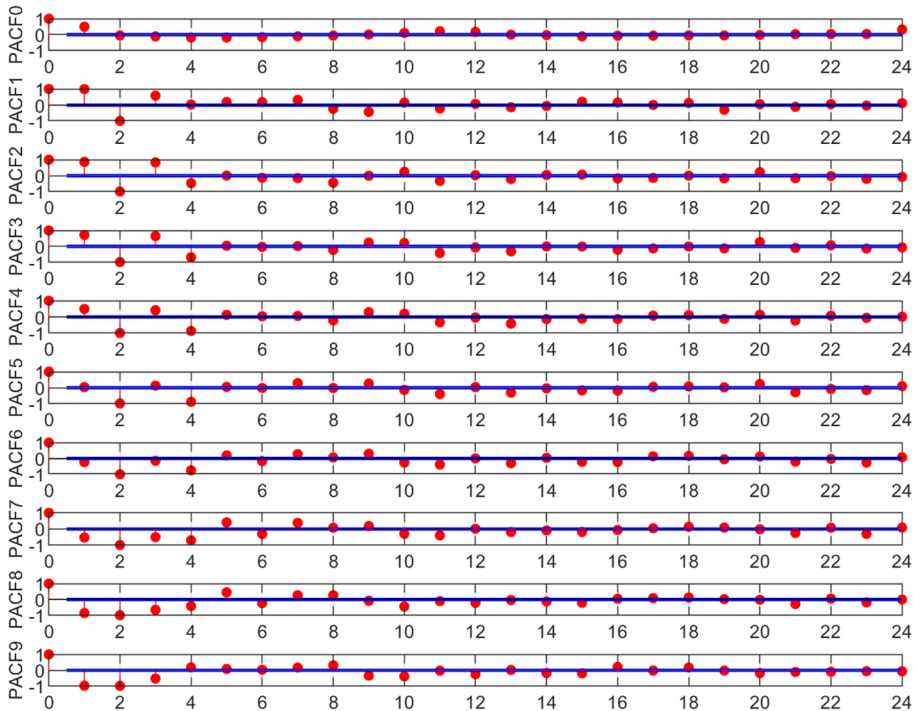

Fig. 4 PACF values of each series from the Xinfengjiang Reservoir

Table 1 The selected input values of each series for the Xinfengjiang and Guangzhao Reservoirs

Reservoir	No	Series	Input variables	Numbers of input
Xinfengjiang	1	Original	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8}$	8
	2	IMF_1	$x_{t-1}, x_{t-2}, x_{t-3}$	3
	3	IMF_2	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}$	4
	4	IMF_3	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}$	4
	5	IMF_4	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}$	4
	6	IMF_5	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}$	4
	7	IMF_6	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}$	7
	8	IMF_7	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}$	7
	9	IMF_8	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8}$	8
	10	IMF_9	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}$	4
Guangzhao	1	Original	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}$	7
	2	IMF_1	$x_{t-1}, x_{t-2}, x_{t-3}$	3
	3	IMF_2	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}$	4
	4	IMF_3	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}$	4
	5	IMF_4	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}$	5
	6	IMF_5	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8}$	8
	7	IMF_6	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}$	5
	8	IMF_7	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}$	5
	9	IMF_8	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}$	4

hidden units for each hidden layer were set by a trial-and-error procedure. In addition, the hyperparameters, i.e., epoch and learning rate, were also set by a trial-and-error procedure.

2. VMD-LSTM and EMD-LSTM models.

For the VMD-LSTM and EMD-LSTM models, there were there three main steps to be implemented. First, the original monthly runoff data were decomposed into several subsequences using VMD or EMD. Second, the standard LSTM model was employed to simulate each subsequence, and the input variables for each subsequence are listed in Table 1. Finally, the results for each subsequence were aggregated as the final results.

3.5 Forecast Results

3.5.1 Results for the Xinfengjiang Reservoir

According to the aforementioned methods, the original monthly runoff series and extracted subsequences were simulated. The detailed evaluation indices of different models over the calibration and validation periods for the Xinfengjiang Reservoir are presented in Table 2. It can be intuitively found that compared with the BPNN, SVM, LSTM, EMD-LSTM and VMD-LSTM models, VMD-GWO-LSTM could yield the best results in terms of all four evaluation indices in both the calibration and validation periods. For instance, compared with the standalone BPNN model, the proposed hybrid VMD-GWO-LSTM model could provide better forecast accuracy with decreases of 77.95% and 75.57% in terms of RMSE

Table 2 Comparison of evaluation indexes of different models for the Xinfengjiang Reservoir

Models	Calibration				Validation			
	RMSE	MAPE	R	CE	RMSE	MAPE	R	CE
BPNN	176.912	116.940	0.558	0.245	175.616	124.456	0.540	0.193
SVM	162.401	92.020	0.604	0.364	159.082	95.410	0.587	0.338
LSTM	143.215	73.108	0.713	0.505	154.706	76.900	0.640	0.373
EMD-LSTM	76.179	52.945	0.931	0.861	130.014	81.677	0.802	0.557
VMD-LSTM	39.524	30.710	0.984	0.962	43.224	32.676	0.977	0.951
VMD-GWO-LSTM	36.950	29.348	0.985	0.967	38.720	30.395	0.981	0.961

and MAPE and increases of 81.67% and 397.93% in terms of R and CE during the validation period, respectively. As seen in Table 3, the hybrid models, such as EMD-LSTM and VMD-LSTM, consisting of LSTM and decomposed methods outperformed the standalone LSTM model in terms of all four evaluation indices during the calibration and validation periods. For example, compared with the LSTM method, the VMD-LSTM model performed better with decreases of 72.06% and 57.51% in terms of RMSE and MAPE and increases of 52.66% and 154.96% in terms of R and CE during the validation period. In addition, Table 2 also reveals that the proposed hybrid model VMD-GWO-LSTM performed slightly better than the VMD-LSTM model in terms of the four measures both in the calibration and validation periods.

To detect the performance of tracing dynamic changes in the monthly runoff, a comparison of forecasted versus observed runoff data using BPNN, SVM, LSTM, EMD-LSTM, VMD-LSTM and VMD-GWO-LSTM for the Xinfengjiang Reservoir is depicted in Fig. 5. On the whole, all models could simulate monthly runoff to some extent except for significant differences in peak flow prediction, indicating that different models have different abilities to simulate peak runoff. To comprehend the performance of the models, the scatter diagrams for the Xinfengjiang Reservoir show fewer scatters with the VMD-GWO-LSTM than the other five models and are consistent with the results in Table 2.

In addition, to assess the performance of the proposed hybrid mode in peak flow forecasting, peak flow estimates of different models over the validation period for the Xinfengjiang Reservoir can be processed by statistical analysis. As shown in Table 3, the absolute average of the relative error of the BPNN, SVM, LSTM, EMD-LSTM, VMD-LSTM and VMD-GWO-LSTM for forecasting the 21 peak flows were 38.9%, 46.0%, 43.7%, 23.2%, 10.8% and 9.4%, respectively. It can be easily concluded that in the aspect of peak flow forecasting, the VMD-GWO-LSTM model can yield much better forecast accuracy than BPNN, SVM, LSTM and EMD-LSTM models and outperform slightly better forecasts than the VMD-LSTM model.

3.5.2 Results for the Guangzhao Reservoir

The statistics of different models over the calibration and validation periods for the Guangzhao Reservoir are shown in Table 4. It can be easily seen that the hybrid methods, namely, EMD-LSTM, VMD-LSTM and VMD-GWO-LSTM, display better performance than the standalone BPNN, SVM and LSTM methods. Furthermore, Table 4 also reveals that the forecast accuracy of the LSTM model can be enhanced under the

Table 3 Peak flow estimates of different models for the Xinfengjiang Reservoir during the validation period

Peak NO	Observed peaks (m³/s)	BPNN (m³/s)	SVM (m³/s)	LSTM (m³/s)	EMD-LSTM (m³/s)	VMD-LSTM (m³/s)	VMD-GWO-LSTM(m³/s)	Relative error(%)					
								BPNN	SVM	LSTM	EMD-LSTM	VMD-LSTM	VMD-GWO-LSTM
1	494.0	481.1	269.0	510.0	402.1	493.5	492.3	-2.6	-45.6	3.2	-18.6	-0.1	-0.4
2	368.0	256.2	232.6	247.6	235.6	421.4	401.5	-30.4	-36.8	-32.7	-36.0	14.5	9.1
3	803.0	458.8	408.5	296.9	653.2	751.2	739.4	-42.9	-49.1	-63.0	-18.7	-6.5	-7.9
4	645.0	433.9	329.2	524.6	561.6	572.8	589.8	-32.7	-49.0	-18.7	-12.9	-11.2	-8.6
5	362.0	385.8	311.7	145.4	212.9	388.7	385.6	6.6	-13.9	-59.8	-41.2	7.4	6.5
6	498.0	137.4	168.3	265.3	184.4	427.3	459.2	-72.4	-66.2	-46.7	-63.0	-14.2	-7.8
7	618.0	526.7	324.4	479.4	508.7	508.8	526.4	-14.8	-47.5	-22.4	-17.7	-17.7	-14.8
8	353.0	381.8	319.5	233.4	335.3	332.6	344.2	8.2	-9.4	-33.9	-5.0	-5.8	-2.5
9	336.0	389.9	239.5	495.3	243.1	303.6	332.8	16.0	-28.7	47.4	-27.7	-9.7	-0.9
10	203.0	319.8	280.4	493.7	197.4	255.8	275.6	57.6	38.1	143.2	-2.8	26.0	35.8
11	1496.0	439.3	351.0	412.9	1649.7	1301.1	1321.0	-70.6	-76.5	-72.4	10.3	-13.0	-11.7
12	783.8	493.5	422.4	441.7	592.1	738.2	766.3	-37.0	-46.1	-43.6	-24.5	-5.8	-2.2
13	687.5	311.7	251.4	436.1	610.4	680.8	698.2	-54.7	-63.4	-36.6	-11.2	-1.0	1.6
14	1066.0	299.6	260.5	470.4	619.8	940.2	956.8	-71.9	-75.6	-55.9	-41.9	-11.8	-10.2
15	228.2	278.9	252.5	468.5	159.9	320.0	325.4	22.2	10.6	105.3	-29.9	40.2	42.6
16	867.5	512.7	382.7	469.6	897.9	748.6	769.1	-40.9	-55.9	-45.9	3.5	-13.7	-11.3
17	369.6	176.1	186.8	408.9	291.8	394.0	409.8	-52.3	-49.4	10.6	-21.0	6.6	10.9
18	442.3	500.5	352.2	424.0	313.4	414.8	433.3	13.2	-20.4	-4.1	-29.1	-6.2	-2.0
19	860.9	374.5	302.5	515.3	829.3	785.5	799.0	-56.5	-64.9	-40.1	-3.7	-8.8	-7.2
20	616.2	292.6	233.3	450.7	570.9	589.3	599.8	-52.5	-62.1	-26.9	-7.4	-4.4	-2.7
21	493.6	196.4	214.6	465.0	195.6	482.8	493.6	-60.2	-56.5	-5.8	-60.4	-2.2	0.0
Average (absolute)								38.9	46.0	43.7	23.2	10.8	9.4
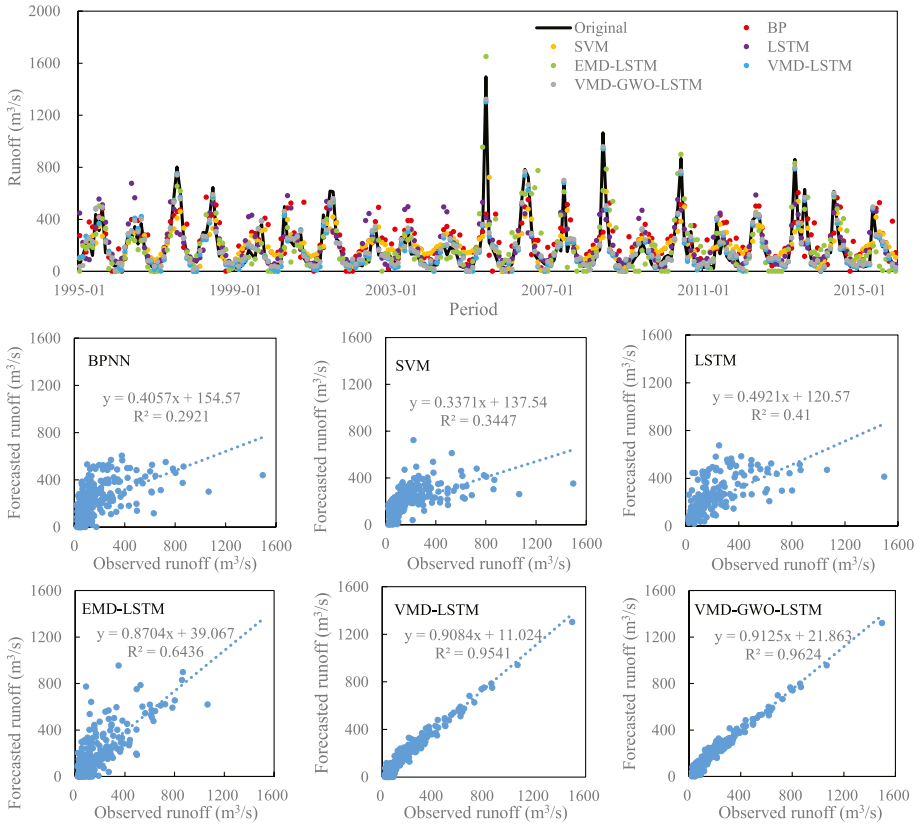
Fig. 5 Comparison of the forecast results for the Xinfengjiang Reservoir during the validation period

condition of optimized hyperparameters. For instance, compared to the SVM model, the VMD-LSTM model can provide better forecast accuracy with decreases of 59.06% and 65.66% in terms of RMSE and MAPE and increases of 31.73% and 80.51% in terms of R and CE during the validation period, respectively. Compared to the VMD-LSTM model, the VMD-GWO-LSTM model can provide better forecast accuracy with

Table 4 Comparison of evaluation indexes of different models for the Guangzhao Reservoir

Models	Calibration				Validation			
	RMSE	MAPE	R	CE	RMSE	MAPE	R	CE
BPNN	162.371	73.088	0.762	0.570	148.575	67.225	0.689	0.373
SVM	157.653	69.643	0.771	0.595	131.621	66.394	0.728	0.508
LSTM	134.204	32.806	0.845	0.706	128.575	50.008	0.759	0.530
EMD-LSTM	105.202	55.066	0.941	0.884	87.303	36.484	0.897	0.783
VMD-LSTM	39.146	19.660	0.989	0.975	53.880	22.798	0.959	0.917
VMD-GWO-LSTM	37.987	17.572	0.990	0.976	34.415	17.921	0.984	0.966

decreases of 36.13% and 21.39% in terms of RMSE and MAPE and increases of 2.61% and 5.34% in terms of R and CE during the validation period, respectively. Hence, this reconfirms that the proposed hybrid model is superior to the other models utilized in this study.

The forecast results of different models for the Guangzhao Reservoir during the validation phase are shown in Fig. 6. It is clear from the hydrographs that the BPNN model had the worst performance in tracing dynamic changes in the monthly runoff, and the remaining models had satisfactory forecast results. It can be intuitively found that the VMD-GWO-LSTM model could offer the least forecast results among the six models and had the best performance with a trendline very near the observed data line.

Table 5 lists the statistics of the peak flow estimates of different models for the Guangzhao Reservoir during the validation period. From Table 5, the absolute average of the relative error of the BPNN, SVM, LSTM, EMD-LSTM, VMD-LSTM and VMD-GWO-LSTM models for forecasting the 18 peak flows were 31.4%, 33.2%, 31.7%, 15.8%, 7.6% and 6.2%, respectively. Thus, in terms of peak flow forecasting, the VMD-GWO-LSTM model can perform much better than the BPNN, SVM, LSTM and EMD-LSTM models, and slightly better than the VMD-LSTM model. As a consequence, the VMD-GWO-LSTM model is an efficient method for monthly runoff forecasting due to its superior performance over comparable models during the validation period.
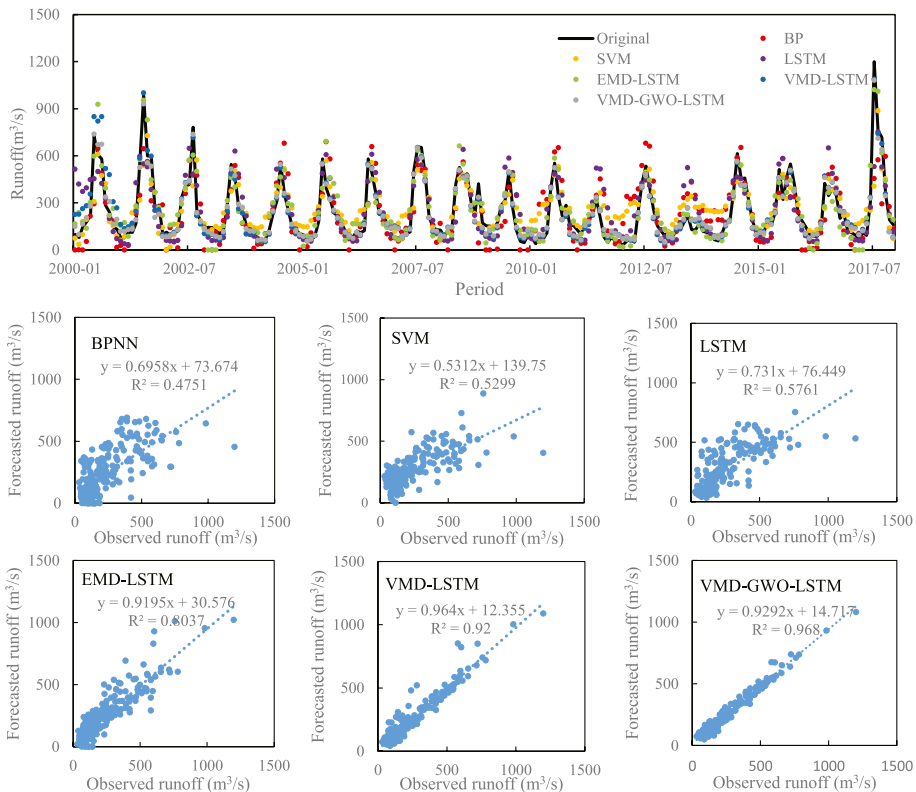

Fig. 6 Comparison of the forecast results for the Guangzhao Reservoir during the validation period

Table 5 Peak flow estimates of different models for the Guangzhao Reservoir during the validation period

Peak NO	Observed peaks (m³/s)	BPNN (m³/s)	SVM (m³/s)	LSTM (m³/s)	EMD-LSTM (m³/s)	VMD-LSTM (m³/s)	VMD-GWO-LSTM (m³/s)	Relative error(%)					
								BP	SVM	LSTM	EMD-LSTM	VMD-LSTM	VMD-GWO-LSTM
1	723.5	293.8	308.1	454.3	597.1	848.7	738.2	-59.4	-57.4	-37.2	-17.5	17.3	2.0
2	983.3	644.1	539.3	548.6	953.6	1001.5	931.1	-34.5	-45.2	-44.2	-3.0	1.9	-5.3
3	782.6	484.0	407.5	477.1	603.5	717.1	736.4	-38.1	-47.9	-39.0	-22.9	-8.4	-5.9
4	544.8	382.6	373.4	490.3	460.7	483.8	515.7	-29.8	-31.5	-10.0	-15.4	-11.2	-5.3
5	546.4	552.6	399.6	523.2	465.2	525.7	544.5	1.1	-26.9	-4.3	-14.9	-3.8	-0.4
6	580.2	357.1	339.5	487.2	393.0	551.0	546.7	-38.5	-41.5	-16.0	-32.3	-5.0	-5.8
7	580.5	295.9	282.2	437.2	291.8	547.9	541.0	-49.0	-51.4	-24.7	-49.7	-5.6	-6.8
8	656.2	542.0	508.2	576.5	635.7	653.0	649.6	-17.4	-22.6	-12.2	-3.1	-0.5	-1.0
9	526.0	492.7	454.3	497.9	661.7	510.0	508.3	-6.3	-13.6	-5.4	25.8	-3.1	-3.4
10	500.2	399.0	352.0	585.4	498.6	489.4	495.6	-20.2	-29.6	17.0	-0.3	-2.1	-0.9
11	552.7	624.3	474.3	483.4	584.9	523.9	523.3	13.0	-14.2	-12.5	5.8	-5.2	-5.3
12	320.4	347.6	327.5	521.3	311.4	332.4	336.4	8.5	2.2	62.7	-2.8	3.7	5.0
13	535.0	679.8	495.1	445.0	471.0	484.7	483.7	27.1	-7.5	-16.8	-12.0	-9.4	-9.6
14	219.9	357.1	321.9	524.8	224.6	267.4	266.9	62.4	46.4	138.6	2.1	21.6	21.4
15	614.1	588.4	434.5	464.8	563.2	556.8	564.1	-4.2	-29.2	-24.3	-8.3	-9.3	-8.1
16	548.0	338.2	397.4	333.6	444.7	493.3	508.5	-38.3	-27.5	-39.1	-18.8	-10.0	-7.2
17	423.9	191.3	263.8	471.5	275.8	465.2	461.9	-54.9	-37.8	11.2	-34.9	9.7	9.0
18	1199.6	454.5	405.7	531.1	1021.1	1087.9	1081.4	-62.1	-66.2	-55.7	-14.9	-9.3	-9.9
Average (absolute)								31.4	33.2	31.7	15.8	7.6	6.2

3.6 Discussion

The statistics of the forecast results yielded by the models clearly indicate that the proposed model can offer the best performance among these models. In reality, once built, the proposed model can be used for one-step monthly runoff forecasting on the condition that the decomposition of the observed runoff data is executed and the forecast results of each subseries are aggregated. Generally, multistep monthly runoff forecasting can also be carried out iteratively. That is, the results of the current one-step monthly runoff forecasting are decomposed and selected as inputs to forecast the next one-step monthly runoff.

According to the forecast results provided by BPNN, SVM and LSTM, it can be directly found that there are significant differences in terms of the four evaluation indices, demonstrating the importance of model selection and model parameter calibration. For the BPNN model, the gradient-based training algorithms have some drawbacks, such as overfitting and local optima. The ordinary SVM employing the structural risk-minimization principle can obtain good generalization performance. Nonetheless, the performance of SVM usually relies on the optimization algorithm to optimize the parameters, and many studies can be found in the literature (Feng et al. 2020). As a deep learning algorithm, the LSTM can overcome the vanishing/exploding gradient problem faced by traditional RNN and can exhibit good generalization performance in hydrological time series prediction (Kratzert et al. 2018). Influenced by many factors such as human activities and climate change, runoff usually contains multifrequency components (Niu et al. 2019). Hence, it is difficult to use a standalone prediction model to completely simulate runoff precisely because only one resolution component is used and the underlying multiscale phenomena cannot be unraveled. According to the literature (Lv et al. 2020; Zuo et al. 2020), adopting decomposition methods can effectively forecast the accuracy of the LSTM model. As decomposition methods, EMD and VMD are utilized to identify the multifrequency components to decrease the modeling difficulty. Therefore, the EMD-LSTM and VMD-LSTM models performed better than the standalone LSTM. Although many successful applications of the LSTM have not involved how to optimize the hyperparameters, it is still worth considering hyperparameter optimization to enhance model performance, and swarm intelligent algorithms (i.e., GWO) can be selected as possible solutions. As revealed by Yuan et al. (2018), the hyperparameter optimization of LSTM models can enhance model performance. Consequently, the proposed VMD-GWO-LSTM model outperformed the VMD-LSTM model. Hence, the framework of the "decomposition-optimization-model" in using the LSTM, such as VMD-PSO-LSTM, was verified effectively for hydrological forecasting (Wang et al. 2021).

The probable causes of the VMD-GWO-LSTM model being superior to the comparable models can be generally attributed to the contribution of VMD and hyperparameter optimization based on the GWO in the LSTM. VMD can decompose the monthly runoff time series into several subsequences and can reveal the underlying multiscale phenomena implied in the monthly runoff time series. Each subsequence was simulated by the LSTM with hyperparameter optimization conducted by the GWO, which can identify the dynamic changes and decrease the modeling difficulty. Meanwhile, automatic optimization of hyperparameters of the LSTM conquers the drawbacks of presetting parameters, easily causing lower forecast accuracy.

Although the feasibility of the VMD-GWO-LSTM model was verified with monthly runoff data derived from two reservoirs, further research should be conducted in the

future. Although the GWO has stronger robustness and searching ability than the PSO in solving optimization problems, the comparison of VMD-GWO-LSTM and VMD-PSO-LSTM models was not made in this study and can be carried out in the future. It is necessary to involve new and excellent decomposition algorithms to enhance the quality of subsequences. Of course, more machine learning techniques should be investigated and verified to improve the single model forecast accuracy. Furthermore, the standard swarm optimization algorithms, for example the GWO used in this study, should be modified to improve the quality of parametric optimization for the models.

4 Conclusion

In this study, a hybrid model, VMD-GWO-LSTM, is proposed for forecasting monthly runoff. This innovation was implemented in three steps. First the original monthly runoff data were decomposed into several subsequences. Second, each subsequence was simulated by a standalone LSTM model, of which the hyperparameters, including learning rate, epochs and hidden layer neurons, were optimized by GWO. Finally, all outputs of the standalone LSTM for each subsequence were aggregated as the final forecast results. Monthly runoff data derived from two reservoirs (Xinfengjiang and Guangzhao Reservoirs) located in China were employed to investigate the proposed hybrid model. To evaluate the model performance, four commonly used statistical evaluation indices were utilized, and five models, namely, BPNN, SVM, LSTM, EMD-LSTM and VMD-LSTM, were used for comparison. The results indicated that the proposed model outperformed the five models in terms of all four evaluation indices. The proposed method is easy to understand and implement. Hence, it is feasible and promising for improving the forecasting accuracy of monthly runoff prediction. Furthermore, it also provides a useful tool for solving other hydrological time series forecasting, such as water level forecasting.

Declarations

References

Abuzied SM, Mansour BMH (2019) Geospatial hazard modeling for the delineation of flash flood-prone zones in Wadi Dahab basin. Egypt J Hydroinform 21(1):180–206. https://doi.org/10.2166/hydro.2018.043

Abuzied SM, Pradhan B (2020) Hydro-geomorphic assessment of erosion intensity and sediment yield initiated debris-flow hazards at Wadi Dahab Watershed. Egypt Georisk 15(3):221–246. https://doi.org/10.1080/17499518.2020.1753781

Abuzied S, Yuan M, Ibrahim S, Kaiser M, Saleem T (2016) Geospatial risk assessment of flash floods in Nuweiba area. Egypt J Arid Environ 133:54–72. https://doi.org/10.1016/j.jaridenv.2016.06.004

Altan A, Karasu S, Zio E (2021) A new hybrid model for wind speed forecasting combining long short-term memory neural network, decomposition methods and grey wolf optimizer. Appl Soft Comput 100:106996. https://doi.org/10.1016/j.asoc.2020.106996

ASCE-Task-Committee (2000) Artificial neural networks in hydrology-II: Hydrological applications. J Hydrol Eng 5(2):124–137. https://doi.org/10.1061/(ASCE)1084-0699(2000)5:2(124)

Bournas A, Baltas E (2021) Increasing the efficiency of the Sacramento model on event basis in a mountainous river basin. Environ Process 8(2):943–958. https://doi.org/10.1007/s40710-021-00504-4

Budamala V, Mahindrakar AB (2020) Integration of adaptive emulators and sensitivity analysis for enhancement of complex hydrological models. Environ Process 7(4):1235–1253. https://doi.org/10.1007/s40710-020-00468-x

Colominas MA, Schlotthauer G, Torres ME (2014) Improved complete ensemble EMD: A suitable tool for biomedical signal processing. Biomed Signal Process Control 14:19–29. https://doi.org/10.1016/j.bspc.2014.06.009

Chu HB, Wei JH, Jiang Y (2021) Middle- and long-term streamflow forecasting and uncertainty analysis using lasso-DBN-bootstrap model. Water Resour Manag 35(8):2617–2632. https://doi.org/10.1007/s11269-021-02854-y

Dragomiretskiy K, Zosso D (2014) Variational mode decomposition. IEEE Trans on Signal Process 62(3):531–544. https://doi.org/10.1109/tsp.2013.2288675

El Harraki W, Ouazar D, Bouziane A, El Harraki I, Hasnaoui D (2021) Streamflow prediction upstream of a dam using SWAT and assessment of the impact of land use spatial resolution on model performance. Environ Process 8(3):1165–1186. https://doi.org/10.1007/s40710-021-00532-0

ElSaid A, El Jamiy F, Higgins J, Wild B, Desell T (2018) Optimizing long short-term memory recurrent neural networks using ant colony optimization to predict turbine engine vibration. Appl Soft Comput 73:969–991. https://doi.org/10.1016/j.asoc.2018.09.013

Fang W, Huang S, Ren K, Huang Q, Huang G, Cheng G, Li K (2019) Examining the applicability of different sampling techniques in the development of decomposition-based streamflow forecasting models. J Hydrol 568:534–550. https://doi.org/10.1016/j.jhydrol.2018.11.020

Feng ZK, Niu WJ (2021) Hybrid artificial neural network and cooperation search algorithm for nonlinear river flow time series forecasting in humid and semi-humid regions. Knowl Based Syst 211:106580. https://doi.org/10.1016/j.knosys.2020.106580

Feng ZK, Niu WJ, Tang ZY, Jiang ZQ, Xu Y, Liu Y, Zhang HR (2020) Monthly runoff time series prediction by variational mode decomposition and support vector machine based on quantum-behaved particle swarm optimization. J Hydrol 583:124627. https://doi.org/10.1016/j.jhydrol.2020.124627

Feng ZK, Niu WJ, Tang ZY, Xu Y, Zhang HR (2021) Evolutionary artificial intelligence model via cooperation search algorithm and extreme learning machine for multiple scales nonstationary hydrological time series prediction. J Hydrol 595:126022. https://doi.org/10.1016/j.jhydrol.2021.126062

He X, Luo J, Zuo G, Xie J (2019) Daily runoff forecasting using a hybrid model based on variational mode decomposition and deep neural networks. Water Resour Manag 33(4):1571–1590. https://doi.org/10.1007/s11269-019-2183-x

Hochreiter S, Schmidhuber J (1997) Long Short-Term Memory. Neural Comput 9(8):1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Kratzert F, Klotz D, Brenner C, Schulz K, Herrnegger M (2018) Rainfall-runoff modelling using long short-term memory (LSTM) networks. Hydrol Earth Syst Sci 22(11):6005–6022. https://doi.org/10.5194/hess-22-6005-2018

Li FG, Ma GW, Chen SJ, Huang WB (2021) An ensemble modeling approach to forecast daily reservoir inflow using bidirectional long- and short-term memory (Bi-LSTM), variational mode decomposition (VMD), and energy entropy method. Water Resour Manag 35(9):2941–2963. https://doi.org/10.1007/s11269-021-02879-3

Li Y, Wu H, Liu H (2018) Multi-step wind speed forecasting using EWT decomposition, LSTM principal computing, RELM subordinate computing and IEWT reconstruction. Energy Convers Manag 167:203–219. https://doi.org/10.1016/j.enconman.2018.04.082

Liao SL, Li G, Sun QY, Li ZF (2016) Real-time correction of antecedent precipitation for the Xinanjiang model using the genetic algorithm. J Hydroinform 18(5):803–815. https://doi.org/10.2166/hydro.2016.168

Liao SL, Liu ZW, Liu BX, Cheng CT, Jin XF, Zhao ZP (2020) Multistep-ahead daily inflow forecasting using the ERA-Interim reanalysis data set based on gradient-boosting regression trees. Hydrol Earth Syst Sci 24(5):2343–2363. https://doi.org/10.5194/hess-24-2343-2020

Liu H, Mi XW, Li YF (2018) Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, LSTM network and ELM. Energy Convers Manag 159:54–64. https://doi.org/10.1016/j.enconman.2018.01.010

Lv N, Liang XX, Chen C, Zhou Y, Li J, Wei H, Wang H (2020) A long Short-Term memory cyclic model with mutual information for hydrology forecasting: A Case study in the xixian basin. Adv Water Resour 141:103622. https://doi.org/10.1016/j.advwatres.2020.103622

Mirjalili S, Mirjalili SM, Lewis A (2014) Grey Wolf Optimizer Adv Eng Softw 69:46–61. https://doi.org/10.1016/j.advengsoft.2013.12.007

Nakisa B, Rastgoo MN, Rakotonirainy A, Maire F, Chandran V (2018) Long short term memory hyperparameter optimization for a neural network based emotion recognition framework. IEEE Access 6:49325–49338. https://doi.org/10.1109/access.2018.2868361

Ni L, Wang D, Singh VP, Wu J, Wang Y, Tao Y, Zhang J (2020) Streamflow and rainfall forecasting by two long short-term memory-based models. J Hydrol 583:124296. https://doi.org/10.1016/j.jhydrol.2019.124296

Niu WJ, Feng ZK, Liu S, Chen YB, Xu YS, Zhang J (2021) Multiple hydropower reservoirs operation by hyperbolic grey wolf optimizer based on elitism selection and adaptive mutation. Water Resour Manag 35(2):573–591. https://doi.org/10.1007/s11269-020-02737-8

Niu WJ, Feng ZK, Zeng M, Feng BF, Min YW, Cheng CT, Zhou JZ (2019) Forecasting reservoir monthly runoff via ensemble empirical mode decomposition and extreme learning machine optimized by an improved gravitational search algorithm. Appl Soft Comput 82:105589. https://doi.org/10.1016/j.asoc.2019.105589

Noorbeh P, Roozbahani A, Moghaddam HK (2020) Annual and monthly dam inflow prediction using bayesian networks. Water Resour Manag 34:2933–2951. https://doi.org/10.1007/s11269-020-02591-8

Nourani V, Alami MT, Aminfar MH (2009) A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation. Eng Appl Arti Intell 22(3):466–472. https://doi.org/10.1016/j.engappai.2008.09.003

Nuaekaew K, Artrit P, Pholdee N, Bureerat S (2017) Optimal reactive power dispatch problem using a two-archive multi-objective grey wolf optimizer. Expert Syst Appl 87:79–89. https://doi.org/10.1016/j.eswa.2017.06.009

Riahi-Madvar H, Dehghani M, Memarzadeh R, Gharabaghi B (2021) Short to long-term forecasting of river flows by heuristic optimization algorithms hybridized with ANFIS. Water Resour Manag 35(4):1149–1166. https://doi.org/10.1007/s11269-020-02591-8

Roushangar K, Ghasempour R, Nourani V (2021) The potential of integrated hybrid pre-post-processing techniques for short- to long-term drought forecasting. J Hydroinform 23(1):117–135. https://doi.org/10.2166/hydro.2020.088

Şen Z (2021) Reservoirs for water supply under climate change impact—A Review. Water Resour Manag 35(11):3827–3843. https://doi.org/10.1007/s11269-021-02925-0

Seo Y, Kim S, Kisi O, Singh VP (2015) Daily water level forecasting using wavelet decomposition and artificial intelligence techniques. J Hydrol 520:224–243. https://doi.org/10.1016/j.jhydrol.2014.11.050

Shahid F, Zameer A, Mehmood A, Raja MAZ (2020) A novel wavenets long short term memory paradigm for wind power prediction. Appl Energ 269:115098. https://doi.org/10.1016/j.apenergy.2020.115098

Shu XS, Ding W, Peng Y, Wang ZR, Wu J, Li M (2021) Monthly streamflow forecasting using convolutional neural network. Water Resour Manag 35(15): 5089–5104. https://doi.org/10.1007/s11269-021-02961-w

Sibtain M, Li XS, Bashir H, Azam MI (2021) A hybrid model for runoff prediction using variational mode decomposition and artificial neural network. Water Resour+ 48(5): 701–712. https://doi.org/10.1134/s0097807821050171

Tikhamarine Y, Souag-Gamane D, Ahmed AN, Kisi O, El-Shafie A (2020) Improving artificial intelligence models accuracy for monthly streamflow forecasting using grey Wolf optimization (GWO) algorithm. J Hydrol 582:124435. https://doi.org/10.1016/j.jhydrol.2019.124435

Wan F, Wang FQ, Yuan WL (2017) The reservoir runoff forecast with artificial neural network based on ant colony optimization. Appl Ecol Env Res 15(4): 497–510. https://doi.org/10.15666/aeer/1504_497510

Wang XJ, Wang YP, Yuan PX, Wang L, Cheng DL (2021) An adaptive daily runoff forecast model using VMD-LSTM-PSO hybrid approach. Hydrol Sci J 66(9):1488–1502. https://doi.org/10.1080/02626667.2021.1937631

Wang YX, Liu FY, Jiang ZS, He SL, Mo QY (2017) Complex variational mode decomposition for signal processing applications. Mech Syst Signal Process 86:75–85. https://doi.org/10.1016/j.ymssp.2016.09.032

Wen X, Feng Q, Deo RC, Wu M, Yin Z, Yang L, Singh VP (2019) Two-phase extreme learning machines integrated with the complete ensemble empirical mode decomposition with adaptive noise algorithm for multi-scale runoff prediction problems. J Hydrol 570:167–184. https://doi.org/10.1016/j.jhydrol.2018.12.060

Yu J, Wang S, Xi L (2008) Evolving artificial neural networks using an improved PSO and DPSO. Neurocomputing 71(4):1054–1060. https://doi.org/10.1016/j.neucom.2007.10.013

Yuan XH, Chen C, Lei XH, Yuan YB, Adnan RM (2018) Monthly runoff forecasting based on LSTM-ALO model. Stoch Environ Res Risk Assess 32(8):2199–2212. https://doi.org/10.1007/s00477-018-1560-y

Zhang M, Jiang ZN, Feng K (2017) Research on variational mode decomposition in rolling bearings fault diagnosis of the multistage centrifugal pump. Mech Syst Signal Process 93:460–493. https://doi.org/10.1016/j.ymssp.2017.02.013

Zuo GG, Luo JG, Wang N, Lian YN, He XX (2020) Decomposition ensemble model based on variational mode decomposition and long short-term memory for streamflow forecasting. J Hydrol 585:124776. https://doi.org/10.1016/j.jhydrol.2020.124776