



OPEN

Monthly runoff prediction based on a coupled VMD-SSA-BiLSTM model

Xianqi Zhang^{1,2,3}, Xin Wang^{1✉}, Haiyang Li¹, Shifeng Sun¹ & Fang Liu¹

The accurate prediction of monthly runoff in the lower reaches of the Yellow River is crucial for the rational utilization of regional water resources, optimal allocation, and flood prevention. This study proposes a VMD-SSA-BiLSTM coupled model for monthly runoff volume prediction, which combines the advantages of Variational Modal Decomposition (VMD) for signal decomposition and preprocessing, Sparrow Search Algorithm (SSA) for BiLSTM model parameter optimization, and Bi-directional Long and Short-Term Memory Neural Network (BiLSTM) for exploiting the bi-directional linkage and advanced characteristics of the runoff process. The proposed model was applied to predict monthly runoff at GaoCun hydrological station in the lower Yellow River. The results demonstrate that the VMD-SSA-BiLSTM model outperforms both the BiLSTM model and the VMD-BiLSTM model in terms of prediction accuracy during both the training and validation periods. The Root-mean-square deviation of VMD-SSA-BiLSTM model is 30.6601, which is 242.5124 and 39.9835 lower compared to the BiLSTM model and the VMD-BiLSTM model respectively; the mean absolute percentage error is 5.6832%, which is 35.5937% and 6.3856% lower compared to the other two models, respectively; the mean absolute error was 19.8992, which decreased by 136.7288 and 25.7274 respectively; the square of the correlation coefficient (R^2) is 0.93775, which increases by 0.53059 and 0.14739 respectively; the Nash–Sutcliffe efficiency coefficient was 0.9886, which increased by 0.4994 and 0.1122 respectively. In conclusion, the proposed VMD-SSA-BiLSTM model, utilizing the sparrow search algorithm and bidirectional long and short-term memory neural network, enhances the smoothness of the monthly runoff series and improves the accuracy of point predictions. This model holds promise for the effective prediction of monthly runoff in the lower Yellow River.

Runoff simulation and prediction play a vital role in water resource management, regulation, and rational planning. They are crucial for the efficient utilization of water resources, flood control, and disaster reduction. With the rapid advancement of artificial intelligence technology, numerous deep learning algorithms have emerged, and comprehensive forecasting models based on intelligent methods and numerical weather prediction have been proposed. These models involve various optimization algorithms such as Chaos Optimization Algorithm¹, bald eagle search optimization algorithm², Particle Swarm Optimization (PSO)³, and artificial neural network models^{4,5}, which have deepened their intersectionality with hydrology.

The pursuit of a runoff prediction model with high accuracy and applicability has been a topic of constant concern in hydrological forecasting. Both domestic and foreign researchers have conducted extensive studies to improve the accuracy of prediction models, resulting in fruitful results. For instance, Feng et al.⁶ combined quantum behavioural particle swarm optimization algorithms with variational modal decomposition and SVM to build a monthly runoff prediction model. Lei et al.⁷ proposed a prediction model that combines empirical mode decomposition and Metropolis Hastings sampling Bayesian model for hydrological prediction. Roy⁸ used the LSTM model to predict ET0 in multiple watersheds with daily and multi-step forward predictions. The LSTM model demonstrated strong adaptability to various prediction indicators. Xu Dongmei et al.⁹ utilized LSSVM and CEEMDAN to predict monthly runoff at Changshui hydrological station. Fan Hongxiang and others developed a meteorological runoff model for the Poyang Lake basin based on the Long short-term memory neural network, effectively simulating the runoff process, capturing extreme runoff values, and reflecting short-term fluctuations. The Long Short Term Memory (LSTM) neural network model has been widely employed in runoff prediction

¹Water Conservancy College, North China University of Water Resources and Electric Power, Zhengzhou 450046, China. ²Collaborative Innovation Center of Water Resources Efficient Utilization and Protection Engineering, Zhengzhou 450046, China. ³Technology Research Center of Water Conservancy and Marine Traffic Engineering, Henan Province, Zhengzhou 450046, China. ✉email: 928139226@qq.com

due to its nonlinear prediction capability, faster convergence speed, and long-term memory effect. However, when the LSTM model learns time series, it faces challenges such as a poor early feature memory effect, leading to the loss of features in the initial learned part, and difficulties in fully capturing the entire time series features¹⁰.

As research progresses, there is an increasing demand for higher accuracy in runoff prediction. Seo¹¹ and He¹² developed a combined runoff prediction model based on the VMD algorithm, demonstrating improved prediction accuracy after data decomposition and reconstruction. Ba et al.¹³ used the sparrow search algorithm (SSA) to optimize artificial neural networks (ANN model), showcasing the advantages of simple implementation, high search accuracy, fast convergence speed, stability, and robustness. Chen and Liang¹⁴ combined the empirical modal decomposition method (EMD), attention mechanism, and BiLSTM neural network to predict the daily longitudinal flow at the Qingxi River site in Xuanhan County, obtaining accurate results that met prediction requirements. Regarding the study of bidirectional long short-term memory neural networks, Li and Jiang¹⁵ used STL for sequence decomposition, temporal convolution, a bidirectional long short-term memory network (TCN-BiLSTM) for feature learning the decomposed series, and interdependent moment feature emphasis using DMAAttention to predict the concentration of air pollutants, achieving accurate prediction accuracy. Sathi et al.¹⁶ used a CNN-BiLSTM model to predict the attention to the induced electric field of a transcranial magnetic stimulation coil with similarly excellent results. Current research on monthly runoff prediction models, both domestically and internationally, mainly focuses on coupling optimization algorithms with neural networks to improve prediction accuracy by constructing coupled models. Overall, as the complexity of hydrological phenomena deepens, the development of conceptual and physical hydrological models encounters certain bottlenecks¹⁷, and the accuracy of predictions needs further improvement. Therefore, this article proposes a VMD-SSA-BiLSTM coupled model suitable for predicting monthly runoff in the lower Yellow River. The model combines the variational mode decomposition (VMD) used for signal decomposition and preprocessing, the sparse search algorithm (SSA) used for parameter optimization of the BiLSTM model, and the bidirectional Long short-term memory neural network (BiLSTM) used to utilize the bidirectional links and advanced characteristics of the runoff process. The proposed VMD SSA-BiLSTM model combines the sparrow search algorithm and the bidirectional Long short-term memory neural network, which enhances the smoothness of monthly runoff series and improves the accuracy of point prediction.

Theory and method

Sparrow Search Algorithm (SSA). The Sparrow Search Algorithm (SSA) is an intelligent optimization algorithm that draws inspiration from the foraging and anti-predation behavior of sparrows¹⁸. The optimization process of SSA can be described as follows:

The formula for updating the discoverer's position $X_{i,j}^{t+1}$ is:

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^{t+1} \exp[-i/\alpha i_{ter_{max}}] & R_2 < S_T \\ X_{i,j}^t + QL & R_2 \geq S_T \end{cases} \quad (1)$$

where $i_{ter_{max}}$ is the maximum number of iterations, $x_{i,j}$ represents the position of the i -th sparrow in the j -th dimension, α is a random number in the range (0, 1), R_2 is the early warning value in the range [0, 1], S_T is the safe value in the range [0.5, 1], Q is a random number following a normal distribution, and L is a $1 \times d$ matrix with all elements set to 1.

Additionally, the joiner location $X_{i,j}^{t+1}$ is updated using a specific process:

$$X_{i,j}^{t+1} = \begin{cases} Q \exp[(X_{worst} - X_{i,j}^t)/i^2] & i > n/2 \\ X_p^{t+1} + |X_{i,j}^t - X_p^{t+1}| A^+ L & Others \end{cases} \quad (2)$$

In equation $A^+ = A^T(AA^T)^{-1}$, X_p is the optimal position of the current discoverer; X_{worst} is the current global worst position, and A is a $1 \times d$ matrix with randomly assigned elements of either 1 or -1.

Furthermore, it is assumed that a certain percentage (10–20%) of the sparrow flock are aware of the danger¹⁹. These vigilant sparrows swiftly move towards the safety zone, and the mathematical expression for the location of the vigilantes $X_{i,j}^{t+1}$ is:

$$X_{i,j}^{t+1} = \begin{cases} X_{best}^t + \beta |X_{i,j}^t - X_{best}^t| & f_i > f_g \\ X_{i,j}^t + K \left(\frac{|X_{i,j}^t - X_{worst}^t|}{(f_i - f_w) + \varepsilon} \right) & f_i = f_g \end{cases} \quad (3)$$

In the formula, X_{best} is the current global optimal position, β is a random number used as a step control parameter following a normal distribution with a mean of 0 and a variance of 1, K is a random number within the range [-1, 1], f_i is the fitness value of the individual sparrow, f_g is the current global optimal fitness value, f_w is the current global worst adaptation value, and ε is a constant²⁰.

The algorithmic steps of the Sparrow Search Algorithm (SSA)²¹ are as follows:

Step 1 Initialize the population, specify the proportion of predators and joiners, and set the number of iterations.

Step 2 Calculate the fitness values for each sparrow and sort them in descending order.

Step 3 Update the positions of the discoverers.

- Step 4* Update the positions of the joiners.
Step 5 Update the positions of the vigilant sparrows (those aware of danger).
Step 6 Calculate the fitness values and update the positions of the sparrows.
Step 7 If the desired requirements are met, output the result; otherwise, repeat steps 2 to 6.

The flow chart illustrating the Sparrow Search Algorithm is depicted in Fig. 1.

In this study, the Sparrow Search Algorithm (SSA) is employed to optimize three key parameters of the BiL-STM model: the number of hidden units, the maximum training period, and the initial learning rate.

Variational mode decomposition (VMD). Variational Modal Decomposition (VMD) is a non-recursive decomposition model that employs an adaptive variational method to determine frequency bands and estimate corresponding modes, effectively balancing the errors between them²². VMD aims to decompose a given real input signal $f(t)$ into discrete sub-signals or modes, denoted as μ_k , where each mode u_k is primarily concentrated around a center frequency w_k ²³. The specific steps for VMD to decompose $f(t)$ into k subsequences²⁴ are as follows:

- Step 1* Calculate the analytic signal and construct the spectrum for each mode μ_k using the Hilbert transform.
Step 2 Shift the frequency spectrum of the modes to the baseband by utilizing their respective estimated center frequencies.
Step 3 Estimate the bandwidth by demodulating the Gaussian smoothness of the signal, which is represented by the L_2 norm of the gradient.

This results in a constrained variational problem, which can be expressed as follows:

$$\left\{ \begin{array}{l} \min_{\{u_k\}, \{w_k\}} \left\{ \sum_k \|\partial(t) \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-jw_k t} \|_2^2 \right\} \\ \text{s.t. } \sum_k u_k = f \end{array} \right\} \quad (4)$$

In the formula, u_k is the k -th modal component, w_k is the frequency center associated with the k -th mode, and $\delta(t)$ is the unit pulse function. To handle the constrained variational problem, quadratic penalty terms and Lagrange multiplier λ are introduced to transform it into an unconstrained problem. The alternating direction multiplier method (ADMM) is then employed for solving the transformed problem²⁵. In this study, the VMD

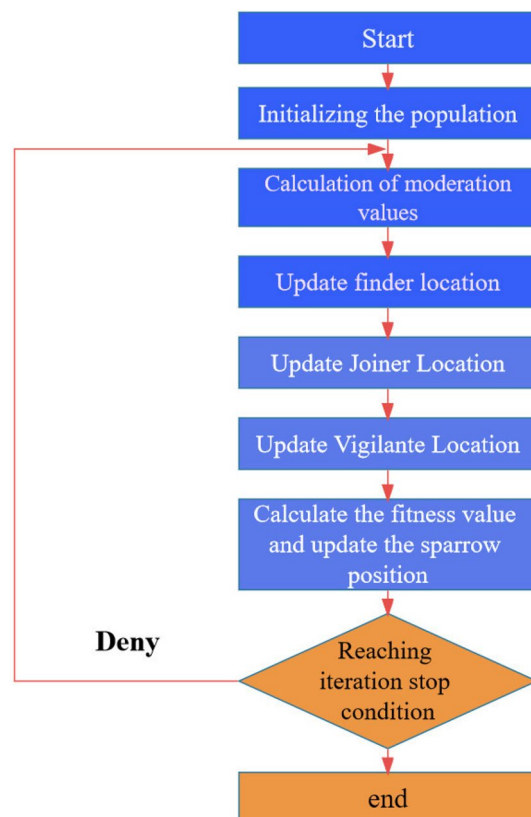


Figure 1. Sparrow search algorithm optimization flowchart.

technique is applied to decompose the monthly runoff time series into a set of relatively smooth sub-series data. This decomposition helps enhance the prediction accuracy of the model.

Bidirectional long short-term memory network (BiLSTM). Bidirectional Long Short-Term Memory (BiLSTM) is an advanced variation of traditional bidirectional recurrent neural networks that replaces regular RNN units with LSTM units. It consists of both forward and backward LSTM components²⁶. By incorporating information from both past and future contexts of a sequence, BiLSTM effectively captures a comprehensive range of features. The hidden layer of BiLSTM consists of two parts: the forward LSTM cell state and the backward LSTM cell state. Historical sequences are fed into the hidden layer through the input layer, enabling forward and backward computations. The model learns from the past and future sequence features to produce the final output result²⁷. The network structure of BiLSTM, as illustrated in Fig. 2, where $X(t)$ represents the input of the network and $Y(t)$ represents the output of the network²¹.

SSA optimization BiLSTM. To obtain optimal parameters for model prediction, the Sparrow Search Algorithm (SSA) was employed to optimize the number of hidden units, the maximum training period, and the initial learning rate of the Bidirectional Long Short-Term Memory (BiLSTM) network. The construction of the SSA-BiLSTM model involved the following steps:

- Step 1* The dataset was divided into training and testing sets, which were then normalized.
- Step 2* The training set was used as the input vector for training the BiLSTM network.
- Step 3* The Sparrow Search Algorithm was applied to optimize the number of hidden units, the maximum training cycle, and the initial learning rate of the BiLSTM model, resulting in the identification of optimal parameters. These parameters were selected to establish the model.
- Step 4* The optimized BiLSTM model was trained again, and the results were compared.
- Step 5* Termination conditions were evaluated. If met, the loop was exited and the prediction results were outputted. Otherwise, the process was repeated.
- Step 6* The test set was fed into the optimized BiLSTM model, and the prediction results were obtained.

The corresponding flow chart depicting these steps is illustrated in Fig. 3.

VMD-SSA-BiLSTM model prediction process. The VMD-SSA-BiLSTM model is a combination of Variational Modal Decomposition (VMD), Sparrow Search Algorithm (SSA), and Bidirectional Long Short-Term Memory (BiLSTM) networks. The prediction steps for monthly runoff using this model are as follows:

- Step 1* Select the monthly runoff data from the previous n months as input to the model.
- Step 2* Decompose the original runoff sequence using VMD to obtain k components.
- Step 3* Define the search range for the sparrow population size N , maximum iteration number M , and parameter range. Then, establish the coupled model of SSA and BiLSTM by selecting Mean Squared Error (MSE) as the objective function in the optimization algorithm.
- Step 4* Input each component separately into the SSA-BiLSTM prediction models to obtain k individual prediction models.
- Step 5* Sum up the predicted values from the k prediction models to obtain the final predicted values for monthly runoff.

The flow chart depicting the construction of the VMD-SSA-BiLSTM model is presented in Fig. 4.

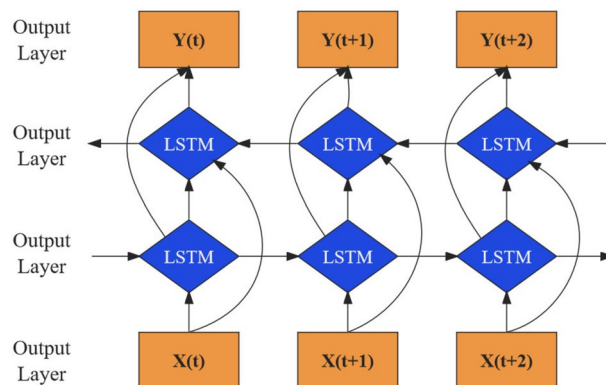


Figure 2. Construction of bidirectional long short-term memory neural network.

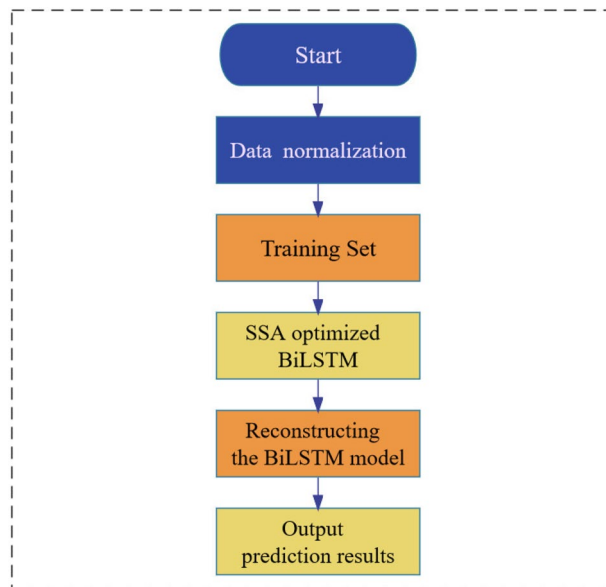


Figure 3. Flow chart of SSA optimized BiLSTM model.

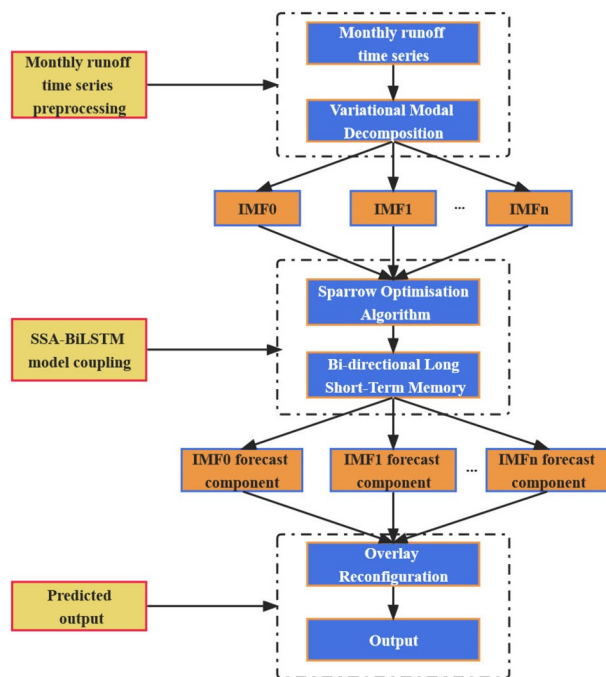


Figure 4. VMD-SSA-BiLSTM model flowchart.

Example applications

Data sources. The Yellow River basin, spanning between 96°–119° E and 32°–42° N, covers an area of approximately 79.5×10^4 km². It is divided into three sections: the upper, middle, and lower reaches. Over the past few decades, the measured runoff in the basin has exhibited a declining trend, accompanied by significant inter-annual variations and differences in intra-annual distribution. These challenges have resulted in severe water supply and demand issues throughout the basin. The problem of sedimentation exacerbates the conflict between water supply and demand, particularly in the downstream section. Hence, accurate runoff prediction in the lower Yellow River basin, including the Gaocun hydrological station, is crucial. Gaocun hydrological station, established in 1934 in Dongming County, Heze City, is located in the lower reaches of the Yellow River basin. It serves as the gateway station for the Yellow River's entry into Shandong Province. The station plays a vital role in managing and developing the Yellow River, as well as in flood control, water resource management, and pro-

viding hydrological information for Shandong's social and economic development. In this article, the monthly meridional flow data of Gaocun Hydrological Station in the lower reaches of the Yellow River from 1950 to 2021 (852 months in total) were selected, and the data information was obtained from the measured data of Gaocun Hydrological Station in the lower reaches of the Yellow River, and was checked for triteness. The location of the study area is shown in Fig. 5, and this figure is created using ArcMap 10.2, URL:www.arcgis.com.

The monthly runoff data measured at the Gaocun hydrological station from 1950 to 2021 were utilized in this study. The dataset was divided into a training set, comprising the initial 70% of the data, and a test set, comprising the remaining 30%. Figure 6 presents the graphical representation of the monthly runoff series in the study area.

VMD decomposition. To predict the 13th month based on the inter-annual variation pattern, the first 12 months of each component were selected as inputs for the model. The selection of the number of modes (k) in the VMD decomposition greatly impacts the decomposition effect. A larger number of decompositions can result in frequency mixing, while a smaller number may lead to the loss of original signal information. Different

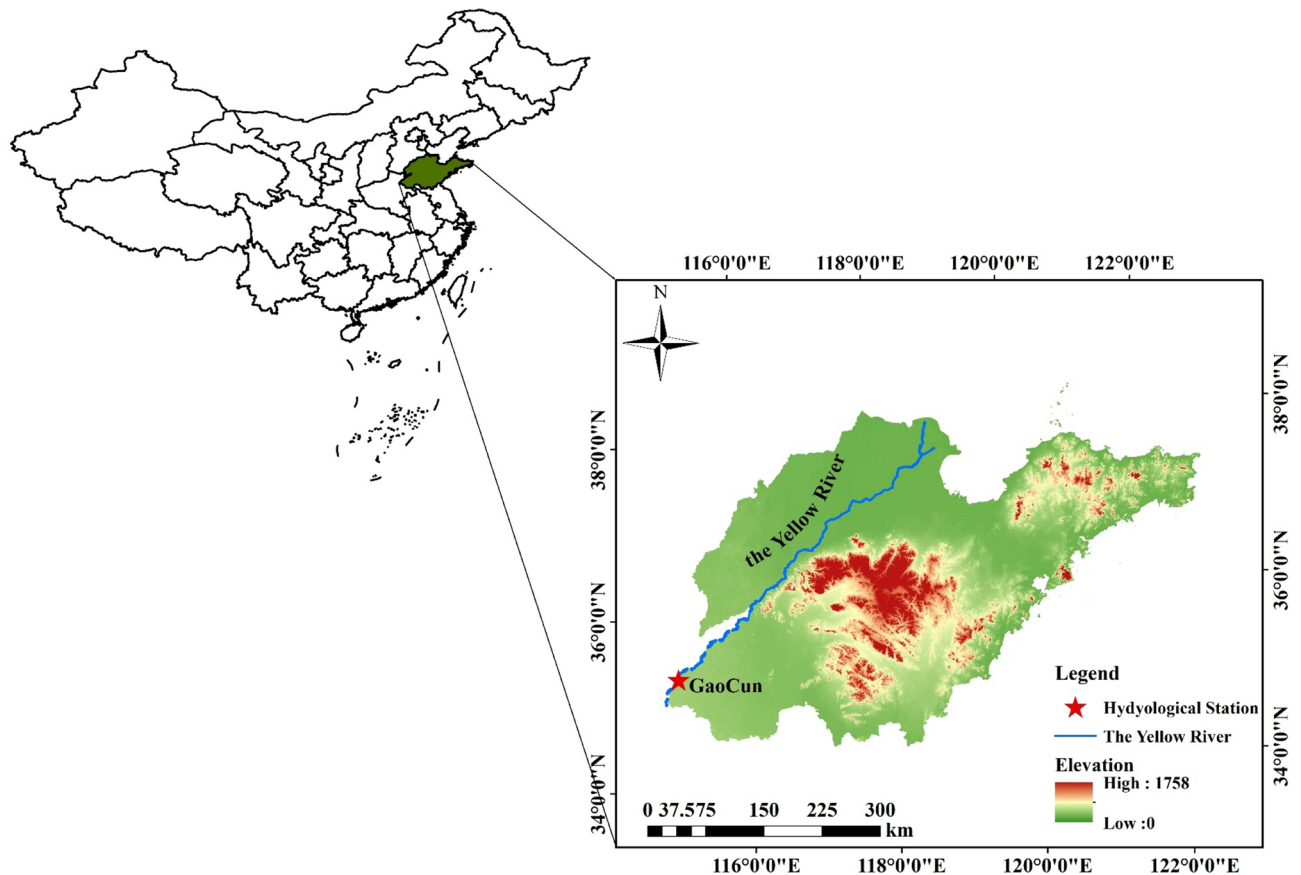


Figure 5. Location map of the study area.

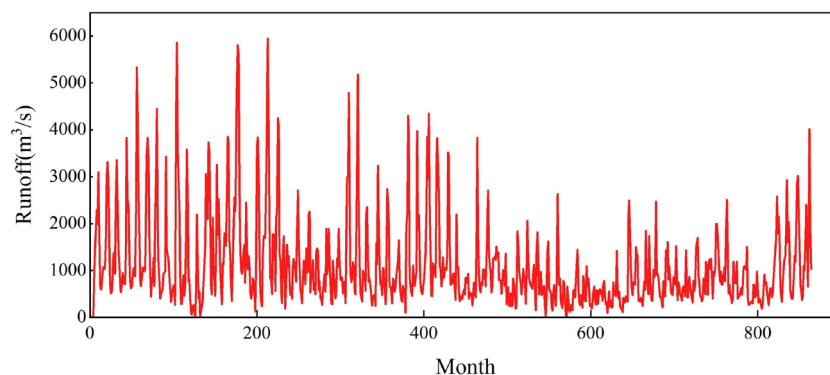


Figure 6. Monthly runoff series graphs.

k values were tested, and it was observed that when k was set to 5, the corresponding center frequencies were more dispersed. Hence, $k=5$ was chosen. Figure 7 illustrates the decomposed subsequences with 5 different frequencies. In contrast to Fig. 6, the original runoff sequence does not exhibit a clear pattern of amplitude change. However, in Fig. 7, as the number of modal components increases, the amplitude change of the sequence becomes periodic, resulting in a more stable sequence. This preprocessing step enables the SSA-BiLSTM model to better capture the feature information in the data and make accurate predictions.

Parameter settings. The BiLSTM model utilizes two hidden layers with 20 neurons in the first hidden layer ($H1$), 20 neurons in the second hidden layer ($H2$), 100 training iterations (E), and a learning rate (η) of 0.005. In the VMD-SSA-BiLSTM model, the sparrow population size (N) is set to 3, and the maximum number of iterations (M) is 5. Among the sparrow population, 20% are discoverers, while the rest are enrollees. The warning value is set to 0.8, indicating the presence or absence of predators. A value less than 0.8 signifies the absence of predators, while a value greater than or equal to 0.8 indicates the presence of predators, necessitating foraging in safer areas. The sparrow search algorithm explores the parameter ranges [1,100] for $H1$ and $H2$, [1,100] for E , and [0.001,0.01] for η in the search for optimal BiLSTM parameters.

Forecast results. The monthly runoff data from 1950 to 2021 at Gaocun Hydrological Station in the lower reaches of the Yellow River were used for prediction. The dataset consisted of 864 months. The first 70% of the dataset (months 1–610) was used for training, while the remaining 30% (months 611–864) was used for testing. Extensive experiments were conducted, and the obtained prediction results are presented in Fig. 8.

Based on the observation of Fig. 8, the VMD-SSA-BiLSTM coupled model shows a generally accurate prediction of monthly runoff. Although there are a few instances where the predicted values deviate significantly from the actual values, overall, the predicted values align closely with the actual values. This indicates that the model exhibits good prediction accuracy, supporting the conclusion that the model is reasonable.

Model comparison and analysis. Four models, namely LSTM, BiLSTM, VMD-BiLSTM, and VMD-SSA-BiLSTM, were employed to predict the monthly runoff of Gaocun Hydrological Station. A statistical analysis was conducted on the error indicators, including Root-mean-square deviation ($RMSE$), mean absolute percentage error ($MAPE$), mean absolute error (MAE), correlation coefficient square (R^2), and Nash coefficient (NSE) of the four models. The following are the calculation formulas for the five error indicators,

$$\text{Formula 1 : } RMSE = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^N (y_i - p_i)^2}$$

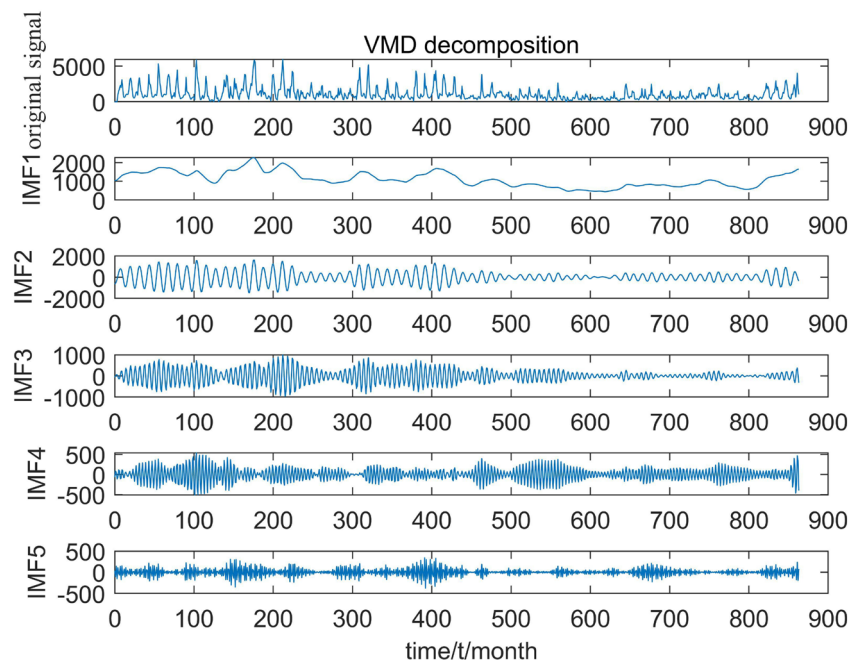


Figure 7. VMD decomposition diagram.

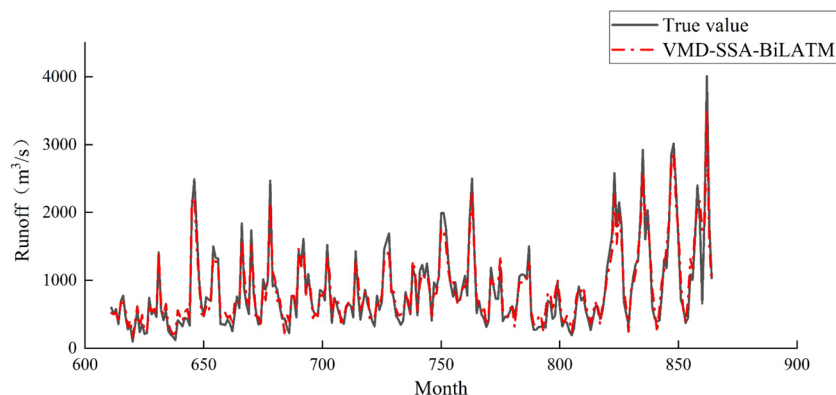


Figure 8. Runoff prediction curve for august.

$$\text{Formula 2 : } MAPE = \frac{100}{N} \cdot \sum_{i=1}^N \frac{y_i - p_i}{y_i}, y \neq 0$$

$$\text{Formula 3 : } MAE = \frac{1}{N} \sum_{i=1}^N |y_i - p_i|$$

$$\text{Formula 4 : } R^2 = \frac{SSR}{SST} = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2}$$

$$\text{Formula 5 : } ESE = 1 - \frac{\sum_{t=1}^T (Q_o^t - Q_m^t)^2}{\sum_{t=1}^T (Q_o^t - \bar{Q}_o)^2}$$

The results of the analysis are presented in Table 1, as depicted below.

Table 1 presents the results of the analysis, revealing that the BiLS™ model outperforms the LS™ model in predicting monthly runoff. During the validation period, the RMSE index decreases by 48.3789, the MAPE index decreases by 8.1182%, the MAE index decreases by 45.6839, the R2 index increases by 0.2665, and the Nash index increases by 0.0928. Hence, BiLS™ is selected as the basis for model construction in this study.

Furthermore, the VMD-BiLS™ model, formed by incorporating VMD decomposition and reconstruction into the BiLS™ model, yields more accurate prediction results. Additionally, the VMD-SSA-BiLS™ model, optimized using the sparrow search algorithm (SSA), demonstrates significant advantages over the first two models during the retraining and validation periods. Specifically, the RMSE index decreases by 242.5124 and 45.9835, the MAPE index decreases by 35.5937% and 6.3856%, the MAE index decreases by 136.7288 and 25.7274, the R2 index increases by 0.53059 and 0.14739, and the NSE index increases by 0.4994 and 0.1122, respectively.

Based on the comparative analysis of the error metrics, it is evident that the VMD-SSA-BiLS™ model exhibits the highest prediction accuracy.

	Models	LS™	BiLS™	VMD-BiLS™	VMD-SSA-BiLS™
Training period	RMSE	395.6334	311.0775	88.6966	37.7533
	MAPE/%	56.6173	48.3651	14.7437	5.6832
	MAE	213.64	173.51	46.7185	19.8992
	R2	0.12118	0.39017	0.77656	0.93775
	NSE	0.3523	0.4423	0.8364	0.9717
Validation period	RMSE	321.5514	273.1725	76.6436	30.6601
	MAPE/%	48.4025	40.2843	11.0762	4.6906
	MAE	197.6471	151.9632	40.9618	15.2344
	R2	0.14662	0.41312	0.79632	0.94371
	NSE	0.3964	0.4892	0.8764	0.9886

Table 1. Parameter rate and test performance statistics for the four models.

Discussion

From Fig. 9, it can be seen that the predicted results of monthly runoff using the LSTM model and BiLSTM model alone differ significantly from the measured data, resulting in poor prediction performance. However, after comparing the two, it can be concluded that the BiLSTM model has slightly better prediction performance than LSTM; However, the predicted values obtained by using VMD-BiLSTM and VMD-SSA-BiLSTM models are relatively consistent with the measured values, but it is difficult to distinguish which model has the better prediction effect based on Fig. 9 alone; Furthermore, according to the scatter plot in Fig. 10, it can be seen that the prediction accuracy of the VMD-SSA-BiLSTM model is slightly better than that of VMD-BiLSTM. Figure 11 shows the Box plot drawn from the model prediction data and the measured data. It can be seen from the graphic comparison that the LSTM model used alone has the worst prediction effect, while the VMD-SSA-BiLSTM model has the best prediction effect. In addition, it can be seen from the Taylor diagram in Fig. 12 that the prediction accuracy of the VMD-SSA-BiLSTM model is the highest. In summary, both the single LSTM model and the BiLSTM model are not effective in predicting monthly runoff. This may be due to the fact that a single time series direct prediction cannot take into account the physical conditions of the basin, considering that the upstream of the Gaocun hydrological station will be influenced by other large and small hydropower stations on runoff, and the mechanism of runoff formation is relatively complex, making the single prediction model less accurate and unable to meet the actual prediction needs. The VMD decomposition can reduce the noise of the original runoff series and extract the complex and

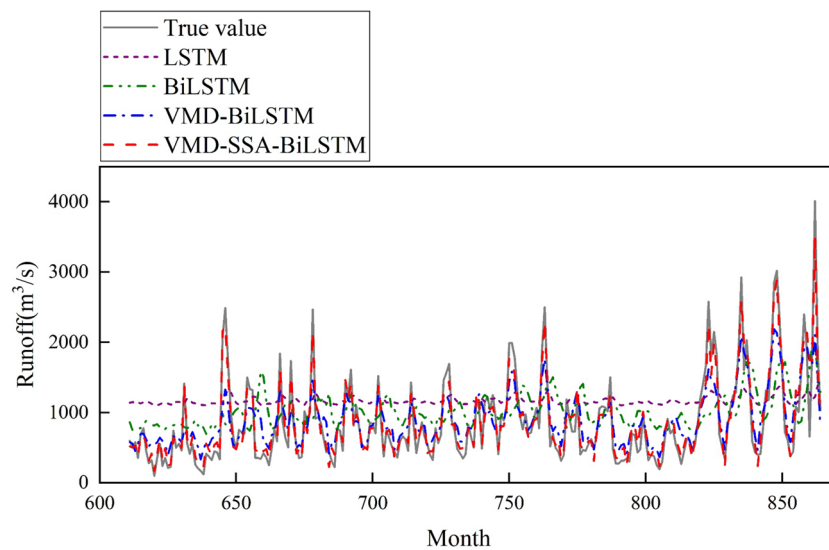


Figure 9. Comparison of the prediction process of the four models during the validation period.

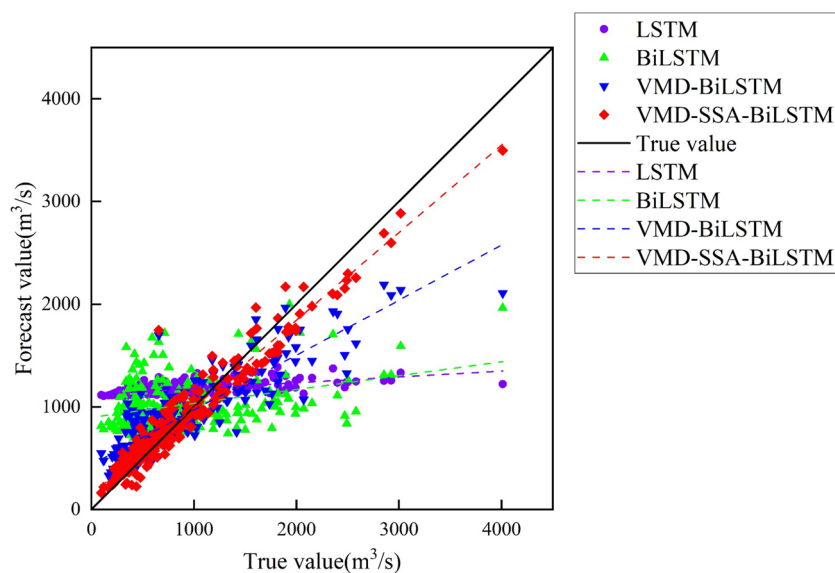


Figure 10. Scatter plot of predicted and measured monthly runoff for the validation period of the four models.

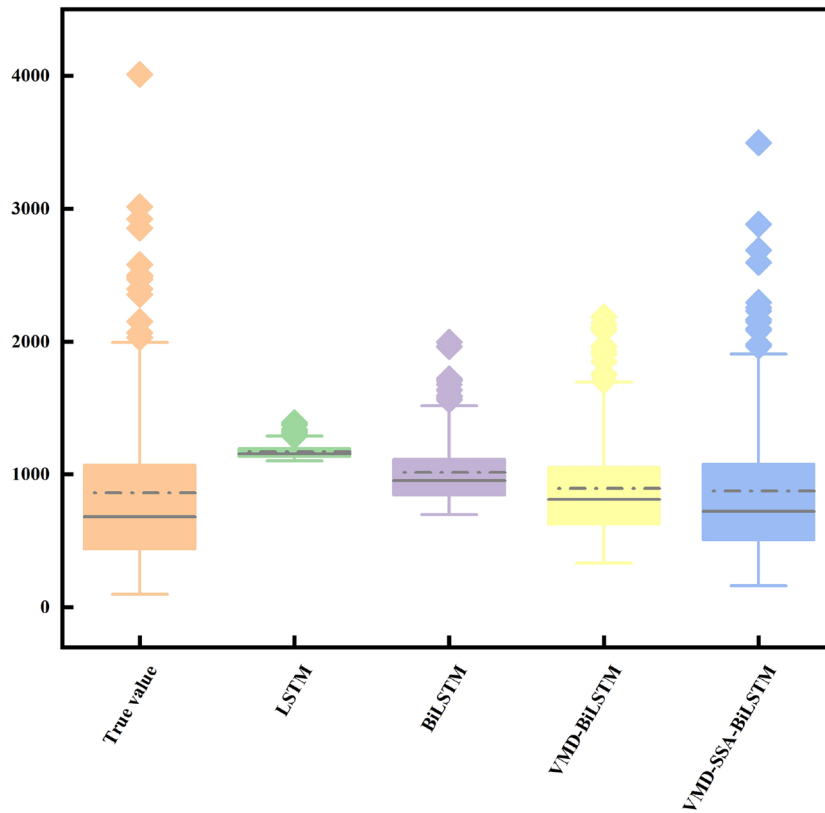


Figure 11. Box plot of predicted and measured monthly runoff for the validation period of the four models.

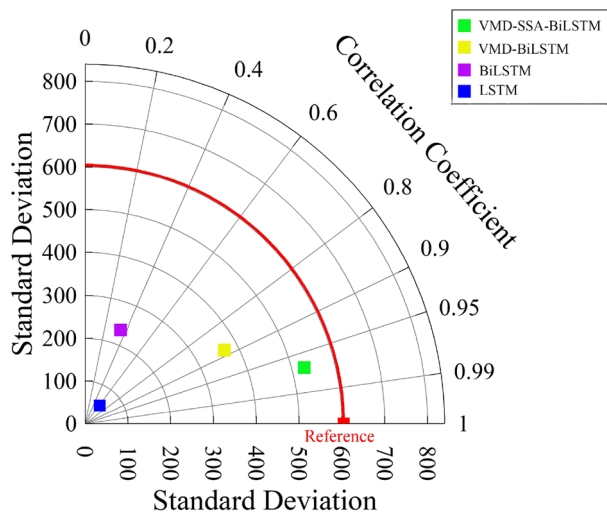


Figure 12. Taylor diagram of predicted and measured monthly runoff for the validation period of the four models.

effective information implied in the runoff data, which to a certain extent can reflect the intrinsic mechanism of watershed runoff formation. The SSA further optimizes the BiLS™ model to obtain the optimal parameters for the number of hidden units, the maximum training period and the initial learning rate, which improves the efficiency of the model parameter selection. The coupled VMD-SSA-BiLS™ model can achieve high prediction accuracy in both the training and validation periods, which shows that the VMD-SSA-BiLS™ model proposed in this paper is feasible for the prediction of monthly runoff. In comparison with other research articles on runoff prediction, this paper firstly addresses the relatively complexity of runoff formation mechanisms in the study area by performing data noise reduction prior to parameter optimisation.

Conclusions

- (1) This study proposes a monthly runoff prediction model based on the coupling of VMD-SSA-BiLS™, which combines variational mode decomposition (VMD), sparrow search algorithm (SSA), and bidirectional Long short-term memory neural network (BiLS™). Compared with previous studies on runoff prediction, this paper constructs a coupled VMD-SSA-BiLS™ model, which combines the advantages of the three and improves the prediction accuracy; at the same time, the BiLS™ neural network, which has been less applied in runoff prediction previously, is used as one of the research methods, and good results have been achieved. The model effectively addresses the nonstationarity of monthly runoff series and enhances the accuracy of point prediction. Additionally, the nonparametric Kernel density estimation provides prediction intervals for monthly runoff without assuming error distribution in advance. This compensates for the limitations of point prediction models in describing prediction results and offers a novel approach to monthly runoff prediction.
- (2) The model was applied to predict monthly runoff at the Takamura hydrological station. The obtained results include a root mean square error (*RMSE*) of 30.6601, a mean absolute percentage error (*MAPE*) of 4.6906%, a mean absolute error (*MAE*) of 15.2344, a mean correlation coefficient squared (*R*²) of 0.94371, and a Nash coefficient (*NSE*) of 0.9886.
- (3) The VMD-SSA-BiLS™ coupled model fully utilizes the benefits of VMD decomposition for data noise reduction and the SSA optimization algorithm for optimizing BiLS™ neural network parameters. Compared to the BiLS™ model and the VMD-BiLS™ model, the VMD-SSA-BiLS™ model achieves notable improvements. Specifically, the *RMSE* is reduced by 242.5124 and 45.9835, the *MAPE* is reduced by 35.5937% and 6.3856%, the *MAE* is reduced by 136.7288 and 25.7274, the *R*² is increased by 0.53059 and 0.14739, and the *NSE* is increased by 0.4994 and 0.1122, respectively. In this paper, only a single-step prediction of the monthly runoff series during the flood season is presented. The sparrow optimisation algorithm has better optimisation performance than the traditional population optimisation algorithm, but still suffers from too fast convergence and the tendency to fall into local optima. The advantage of the model proposed in this article lies in the idea of "decomposition-prediction-reconstruction", which improves the final prediction accuracy with each step of model construction. Then, the BiLS™ model used for prediction, which has the advantages of bidirectional link and advanced characteristics, is also very suitable for runoff prediction. In addition, regarding the limitations of the model, the sparrow optimization algorithm has the problem of too fast convergence for parameter optimization, which can easily result in only local optima. This model has good performance in predicting single time series runoff data, but further improvement is needed when applied to predicting multiple series data.
- (4) In this paper, compared with other research articles, the more novel BiLS™ model was chosen for the prediction method, and in terms of data processing, the VMD decomposition method was used for noise reduction of the data, together with the SSA optimisation algorithm coupled with the model, to obtain a high prediction accuracy.
- (5) In this paper, only a single-step prediction of the monthly runoff series during the flood season is presented. In subsequent studies, the proposed model can be used to make further multi-step predictions on the basis of the single-step prediction. The sparrow optimisation algorithm has better optimisation performance than the traditional population optimisation algorithm, but still suffers from too fast convergence and the tendency to fall into local optima. In the future, suitable algorithms should also be selected for forecasting for the selected study area, further achieving extended research from algorithm improvement and model construction. For future scope of the model, the model can be used for further multi-step prediction based on single step prediction. Also it can be tried to use this model in other predictions such as annual runoff prediction, seasonal runoff prediction.

Data availability

Data and materials are available from the corresponding author upon request.

Received: 15 June 2023; Accepted: 27 July 2023

Published online: 12 August 2023

References

1. Kaffas, K. *et al.* Forecasting soil erosion and sediment yields during flash floods: The disastrous case of Mandra, Greece, 2017. *Earth Surf. Process. Landf. J. British Geomorphol. Res. Gr.* **47**(7), 1744–1760. <https://doi.org/10.1002/esp.5344> (2022).
2. Wang, W. *et al.* An improved bald eagle search algorithm with Cauchy mutation and adaptive weight factor for engineering optimization. *Comput. Model. Eng. Sci.* **136**(2), 1603–1642 (2023).
3. Han, H. & Morrison, R. R. Improved runoff forecasting performance through error predictions using a deep-learning approach. *J. Hydrol.* <https://doi.org/10.1016/j.jhydrol.2022.127653> (2022).
4. Aoulmi, Y., Marouf, N., Rasouli, K. & Panahi, M. Runoff predictions in a semiarid watershed by convolutional neural networks improved with metaheuristic algorithms and forced with reanalysis and climate data. *J. Hydrol. Eng.* **28**(7), 04023018 (2023).
5. Bi, K. *et al.* Accurate medium-range global weather forecasting with 3D neural networks. *Nature* <https://doi.org/10.1038/s41586-023-06185-3> (2023).
6. Feng, Z.-K. *et al.* 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> (2020).
7. Lei, G., Yin, J., Wang, W., Wang, H. & Liu, C. Hydrological frequency analysis in changing environments based on empirical mode decomposition and Metropolis–Hastings sampling Bayesian models. *J. Hydrol. Eng.* **28**(9), 04023027 (2023).

8. Roy, D. K. *et al.* Daily prediction and multi-step forward forecasting of reference evapotranspiration using LSTM and Bi-LSTM models. *Agronomy* **12**(3), 594 (2022).
9. Xu, D. M., Zhuang, W. T. & Wang, W. C. Monthly runoff prediction based on CEEMDAN-WD-PSO-LSSVM model. *China Rural Water Hydropower*, (8), 54–58. <https://doi.org/10.3969/j.issn.1007-2284.2021.08.010> (2021).
10. Rahimzad, M. *et al.* Performance comparison of an LSTM-based deep learning model versus conventional machine learning algorithms for streamflow forecasting. *Water Resour. Manag.* **35**(12), 4167–4187. <https://doi.org/10.1007/s11269-021-02937-w> (2021).
11. Seo, Y., Kim, S. & Singh, V. P. Machine learning models coupled with variational mode decomposition: A new approach for modeling daily rainfall-runoff. *Atmosphere* **9**(7), 251 (2018).
12. He, X., Luo, J., Zuo, G. & Xie, J. Daily runoff forecasting using a hybrid model based on variational mode decomposition and deep neural networks. *Water Resour. Manag.* **33**, 1571–1590 (2019).
13. Ba, H. *et al.* Improving ANN model performance in runoff forecasting by adding soil moisture input and using data preprocessing techniques. *Hydrol. Res. Int. J.* **49**(3/4), 744–760. <https://doi.org/10.2166/nh.2017.048> (2018).
14. Liang, C., Xiaoying, B. & Xinzhi, Z. Research on runoff prediction model based on EMD-ATT-BILSTM. *Mod. Comput.* **28**(1), 18–24. <https://doi.org/10.3969/j.issn.1007-1423.2022.01.003> (2022).
15. Li, W. & Jiang, X. Prediction of air pollutant concentrations based on TCN-BiLSTM-DMAttention with STL decomposition. *Sci. Rep.* **13**(1), 4665 (2023).
16. Sathi, K. A., Hosain, M. K., Hossain, M. A. & Kouzani, A. Z. Attention-assisted hybrid 1D CNN-BiLSTM model for predicting electric field induced by transcranial magnetic stimulation coil. *Sci. Rep.* **13**(1), 2494 (2023).
17. Zhang, S., Chen, J. & Gu, L. Overall uncertainty of climate change impacts on watershed hydrology in China. *Int. J. Climatol.* **42**(1), 507–520. <https://doi.org/10.1002/joc.7257> (2022).
18. Le, X. H. *et al.* Comparison of deep learning techniques for river streamflow forecasting. *IEEE Access* **9**, 71805–71820 (2021).
19. New circuits research findings from Huaiyin institute of technology reported (study on load monitoring and demand side management strategy based on Elman neural network optimized by sparrow search algorithm). *Network Daily News* (Feb.25), 54 (2022).
20. Li, B. *et al.* Tri-stage optimal scheduling for an islanded microgrid based on a quantum adaptive sparrow search algorithm. *Energy Convers. Manag.* **261**(Jun.), 115639.1-115639.21. <https://doi.org/10.1016/j.enconman.2022.115639> (2022).
21. Zhou, J., Xu, Z. & Wang, S. A novel hybrid learning paradigm with feature extraction for carbon price prediction based on Bi-directional long short-term memory network optimized by an improved sparrow search algorithm. *Environ. Sci. Pollut. Res.* **29**(43), 65585–65598. <https://doi.org/10.1007/s11356-022-20450-4> (2022).
22. Recent research from Beijing University of chemical technology highlight findings in automation science (gear fault diagnosis based on variational modal decomposition and wide plus narrow visual field neural networks). *Network Daily News*. (Jan.19), 71–72 (2022).
23. Lu, J., Fu, Y. & Yue, J. Natural gas pipeline leak diagnosis based on improved variational modal decomposition and locally linear embedding feature extraction method. *Trans. Inst. Chem. Eng. Process Saf. Environ. Prot. Part B* **164**, 857–867. <https://doi.org/10.1016/j.psep.2022.05.043> (2022).
24. Dalian university of technology researchers target support vector machines (Bolt loosening detection of rocket connection structure based on variational modal decomposition and support vector machines). (Jul.13), 75–76 (2022).
25. Qu, J. *et al.* An intermittent fault diagnosis method of analog circuits based on variational modal decomposition and adaptive dynamic density peak clustering. *Soft Comput. Fusion Found. Methodol. Appl.* **26**(17), 8603–8615. <https://doi.org/10.1007/s00500-022-07226-1> (2022).
26. Recent research from Pohang University of science and technology (POSTECH) highlight findings in data retention (bi-directional long short-term memory neural network modeling of data retention characterization in 3-D triple-level cell NAND ...). *Network Daily News*. (Aug.22), 40–41 (2022).
27. Roy, S. S. *et al.* Accurate detection of bearing faults using difference visibility graph and bi-directional long short-term memory network classifier. *IEEE Transact. Ind. Appl.* **58**(4), 4542–4551. <https://doi.org/10.1109/TIA.2022.3167658> (2022).

Author contributions

All authors contributed to the study conception and design. writing and editing: X.Z.; chart editing: X.W.; preliminary data collection: S.S., H.L. and F.L. All authors read and approved the final manuscript.

Funding

This work was supported by the Key Scientific Research Project of Colleges and Universities in Henan Province (CN) [Grant Numbers 17A570004].

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to X.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023