Journal of Advances in Mathematics and Computer Science

xx(x): ....., 2026; Article no.JAMCS.xxxxx

ISSN: 2456-9968

(Past name: British Journal of Mathematics & Computer Science, Past ISSN: 2231-0851)

Hybrid Secondary Decomposition and CNN BiLSTM Model for Nonlinear Monthly Runoff Prediction

Research Article

Received: XX January 2026
Accepted: XX January 2026
Online Ready: XX January 2026

Abstract

Addressing the strong nonlinear and non-stationary characteristics of monthly runoff series, which pose challenges to single decomposition or single-model prediction approaches, this paper proposes a hybrid prediction model (ISCBL) combining ICEEMDAN-SSA secondary decomposition with CNN-BiLSTM deep learning. The secondary decomposition is guided by sample entropy to identify and further decompose the most complex intrinsic mode function (IMF), thereby enhancing feature stationarity. The model is validated using monthly runoff data (1971–2023) from the Lanzhou hydrological station in the upper Yellow River, with data split into training (1972–2013) and testing (2013–2023) sets. The results show that the ISCBL model achieves a Nash-Sutcliffe efficiency coefficient (NSE) of 0.9733 and a mean absolute percentage error (MAPE) of 8.35% on the test set, significantly outperforming baseline and primary decomposition models. This study verifies the effectiveness of the sample entropy-guided secondary decomposition strategy in extracting multi-scale features and the advantage of CNN-BiLSTM in capturing spatiotemporal dependencies. The proposed framework provides a robust and accurate method for complex hydrological sequence prediction, offering valuable technical support for water resource management in the Yellow River Basin.

Keywords: Monthly runoff prediction; Secondary decomposition; ICEEMDAN; Singular Spectrum Analysis; CNN-BiLSTM

2020 Mathematics Subject Classification: 68T07; 86A05; 62M10

1 Introduction

Runoff prediction is fundamental work for water resources planning and management. Accurate runoff prediction can provide a scientific basis for decision-makers, effectively improve water resource utilization efficiency, reduce flood disaster losses (Saha and Chandra Pal, 2024), and is of great significance for flood control and disaster reduction, reservoir operation, and optimal allocation of water resources (Ma et al., 2024).

Traditional single-runoff prediction models (WEI & Yuan, 2026) still have limitations when dealing with complex runoff sequences (Du et al., 2023). The runoff formation process is influenced by various factors such as climate, topography, and human activities, exhibiting strong nonlinear and non-stationary characteristics (Nanda and Sen, 2021).

In recent years, the combination of signal decomposition and deep learning (Fang et al., 2026) has provided new ideas for runoff prediction (Xu et al., 2025). Colominas M. A. (Colominas et al., 2014) proposed an Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (ICEEMDAN), which significantly reduces the mode mixing index and better solves the mode mixing problem through dynamic noise injection and residual calculation. Singular Spectrum Analysis (SSA), as a non-parametric spectral estimation method, is particularly suitable for extracting periodic and trend components in time series (Gu et al., 2024). Sibtain Muhammad (Sibtain et al., 2022) et al. proposed a multivariate ultra-short-term wind speed forecasting model (VIL) based on the VMD-ICEEMDAN multistage decomposition method and a Long Short-Term Memory (LSTM) network. Chen Shu et al. used EEMD-VMD two-stage decomposition combined with Support Vector Machines (SVM) and a simple average method (SAM) for annual runoff forecasting, significantly improving prediction accuracy (Chen et al., 2021). Li Menghang (Li et al., 2024) et al. employed a hybrid approach combining Empirical Mode Decomposition and Wavelet Denoising (EMD-WD) with Convolutional Neural Networks (CNN) and Long Short-Term Memory networks (LSTM) to construct two novel models (EMD-WD-CNN and EMD-WD-LSTM) for predicting reference crop evapotranspiration (ET0). Secondary decomposition methods (Yang et al., 2026), by combining different decomposition algorithms (Liang et al., 2026), show significant advantages in runoff prediction.

The upper Yellow River, as the core runoff generation area of the Yellow River basin, has a hydrological process that plays a decisive role in the water resources allocation of the entire basin (Li et al., 2025). This paper takes the Lanzhou hydrological station in the upper Yellow River as the research object, proposes a hybrid prediction model (ISCBL) based on ICEEMDAN-SSA secondary decomposition and CNN-BiLSTM deep learning, and validates it using monthly runoff data from 1971 to 2023 from the Lanzhou station. It aims to improve the accuracy and reliability of runoff prediction,

provide new ideas for monthly runoff prediction, and offer key technical support for the optimal operation of cascade reservoirs in the upper Yellow River.

2 Research Methods

2.1 ICEEMDAN

ICEEMDAN, an improved version of the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise, features an innovative adaptive noise injection mechanism. This mechanism dynamically adjusts noise amplitude, significantly mitigating mode mixing. In this study, the ICEEMDAN parameters were set as: noise standard deviation Nstd = 0.4, number of realizations NR = 10, and maximum iterations MaxIter = 50, following common practices for hydrological time series. The algorithm's main steps are outlined below:

Step 1: Add specific noise to the original signal:

$$x_k(t) = x(t) + \varepsilon_0 E_1(\omega_k(t)), \quad (k = 1, 2, \cdots, n), \tag{2.1}$$

where $x_k(t)$ represents the k-th signal, ε_0 represents the noise standard deviation of the signal during the first decomposition, $\omega_k(t)$ is the k-th added Gaussian white noise, and $E_1(*)$ represents the first intrinsic mode function (IMF) obtained from the ICEEMDAN decomposition.

Step 2: Calculate the local mean of all new signals to obtain the first residual and the first modal component:

$$r_1(t) = \frac{1}{n}\sum_{i=1}^{n} M(x_i(t)), \qquad \text{IMF}_1 = x - r_1(t), \tag{2.2}$$

where $M(x_i(t))$ represents the local mean function.

Step 3: Introduce Gaussian white noise into the initial residual component to obtain the second local mean and the second modal component:

$$M(r_1(t) + \epsilon_1 E(\omega(t))), \quad \text{IMF}_2 = r_1(t) - \frac{1}{n}\sum_{i=1}^{n} M[r_1(t) + \epsilon_2 E(\omega(t))]. \tag{2.3}$$

Step 4: By analogy, obtain the k-th residual:

$$r_k(t) = \frac{1}{n}\sum_{i=1}^{n} M[r_{k-1}(t) + \epsilon_{k-1} E_k(\omega(t))]. \tag{2.4}$$

Then obtain the k-th modal component:

$$IMF_k = r_{k-1}(t) - r_k(t). \tag{2.5}$$

Step 5: Repeat step 4 until the residual is a monotonic function. Finally, the original signal can be expressed as:

$$x(t) = r(t) + \sum_{i=1}^{n} IMF_i. \tag{2.6}$$

2.2 Singular Spectrum Analysis (SSA)

The SSA decomposition process primarily consists of three stages: embedding, singular value decomposition, and reconstruction:

Step 1: Embedding: Select a suitable window length L and transform the time series $x_1, x_2, \cdots, x_n$ into a trajectory matrix. In this study, the window length L was set to 7 months to effectively capture the seasonal patterns in the monthly runoff series:

$$X = \begin{bmatrix} X_1 & X_2 & \cdots & X_{N-L+1} \\ X_2 & X_3 & \cdots & X_{N-L+2} \\ \vdots & \vdots & \ddots & \vdots \\ X_L & X_{L+1} & \cdots & X_N \end{bmatrix}. \tag{2.7}$$

Step 2: Singular Value Decomposition (SVD): Perform SVD decomposition on matrix X: $X = U\Sigma V^T$, where: U is the left singular vector matrix; Σ is the diagonal matrix of singular values; V^T is the transpose of the right singular vector matrix. Then calculate the covariance matrix: $C = XX^T$, solve for the eigenvalues of the covariance matrix and their corresponding eigenvectors: $\lambda_1 > \lambda_2 > \cdots > \lambda_L \geq 0$ and $U_1, U_2, \cdots, U_L$, and finally obtain:

$$X = \sum_{m=1}^{L} \sqrt{\lambda_m} U_m V_m^T, \quad m = 1, 2, \cdots, L. \tag{2.8}$$

Step 3: Reconstruction: First, calculate the projection of the lagged sequence:

$$a_i^m = X_i U_m = \sum_{j=1}^{L} x_{i+j} U_{m,j}, \quad 0 \leq i \leq N - L, \tag{2.9}$$

where X_i is the i-th column of the trajectory matrix, a_i^m is the time principal component, i.e., the weight of X_i in the original time series period. Finally, obtain the expression

for the reconstructed series:

$$x_i = \sum_{k=1}^{L} x_i^k, \quad i = 1, 2, \cdots, N. \tag{2.10}$$

2.3 CNN

The Convolutional Neural Network (CNN) module uses a CNN to train sample data to obtain the output (Xie et al., 2023). Its neural network structure is shown in Fig. 1. It is mainly used to process time series data, using 16 convolution kernels with a size of 3 to perform sliding calculations along the time dimension, extracting local spatiotemporal features of the runoff series. Each convolutional layer is followed by Batch Normalization (BatchNorm) and a ReLU activation function. The former accelerates training convergence by standardizing intermediate outputs, while the latter introduces nonlinearity to enhance feature expression capability. Subsequently, the time dimension of the feature map is compressed by 50% through a max-pooling layer, retaining the most significant feature responses. This hierarchical structure enables the CNN to automatically learn abstract representations from raw runoff data $\rightarrow$ local features $\rightarrow$ key temporal patterns, providing high-information-density input features for the subsequent BiLSTM layer.

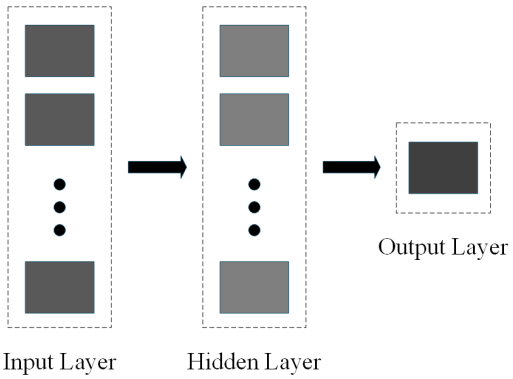


Figure 1: *CNN neural network structure diagram*

2.4 BiLSTM

LSTM (Long Short-Term Memory) is a special type of recurrent neural network that can effectively capture and transmit long-term dependencies in sequence data (Xue et al., 2025). The core gating mechanism structure of the LSTM neural network is

shown in Fig. 2, controlling information flow through the forget gate, input gate, and output gate, thus effectively capturing long-term dependencies.

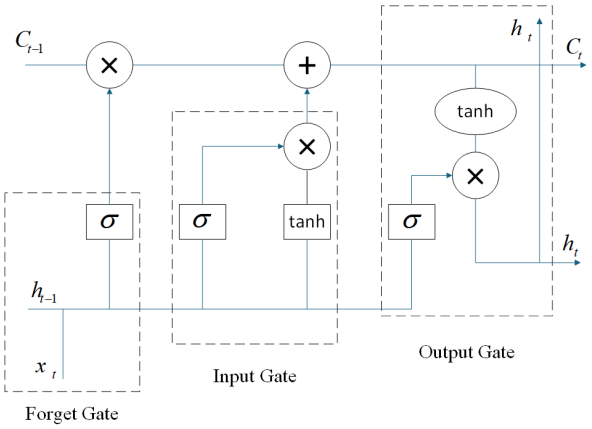


Figure 2: *LSTM neural network schematic diagram*

BiLSTM (Bidirectional LSTM) introduces a bidirectional structure on the basis of LSTM, considering both past and future information simultaneously (Liu et al., 2024), making it suitable for tasks requiring contextual information. The BiLSTM algorithm structure is shown in Fig. 3. The calculation method of BiLSTM:

$$\vec{h}_t = \text{LSTM}(x_t, \vec{h}_{t-1}), \quad \overleftarrow{h}_t = \text{LSTM}(x_t, \overleftarrow{h}_{t-1}), \quad y_t = \sigma(W_y[\vec{h}_t, \overleftarrow{h}_t] + b_y), \quad (2.11)$$

where $\vec{h}_t$ is the state of the forward LSTM hidden layer at time t; $\overleftarrow{h}_t$ is the state of the backward LSTM hidden layer at time t; x_t is the input data at time t; y_t is the output data at time t.

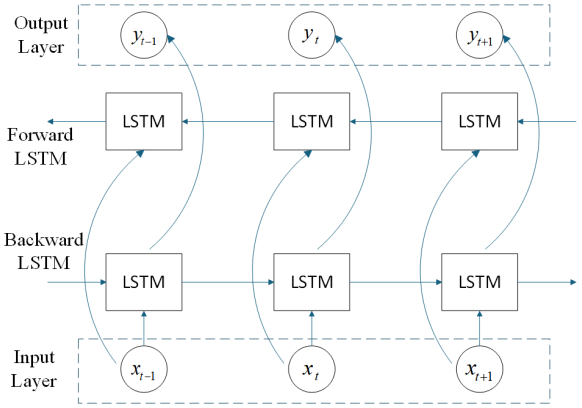


Figure 3: *BiLSTM algorithm structure diagram*

2.5 CNN-BiLSTM

The CNN-BiLSTM model constructed in this paper adopts a cascaded hybrid architecture. First, local spatiotemporal features are extracted through the CNN layer (Hu et al., 2024). After batch normalization and ReLU activation, dimensionality is reduced through a max-pooling layer with a stride of 2. Then, a bidirectional LSTM layer is connected to capture bidirectional temporal dependencies, with 10% Dropout added to prevent overfitting, the predicted value is ultimately output via a fully connected layer. The model minimizes the mean square error through the Adam optimizer and adopts a dynamic learning rate decay strategy. To ensure reproducibility, a fixed random seed (42) was set for model initialization, data shuffling during training, and all stochastic operations. The parameter settings of the model are shown in Table 1:

Table 1: *Combined model parameter table*

Module	Parameter Name	Parameter Value
Input Layer	Time Steps	12
	Feature Dimension	1
CNN Module	Number of Kernels	16
	Kernel Size	3
BiLSTM Module	Hidden Units	25
	Dropout Rate	0.1
	Optimizer	Adam
Training Params	Learning Rate	0.01
	Batch Size	32
	Training Epochs	150

The basic structure diagram of the model is shown in Fig. 4. Its prediction process is divided into four stages:

(1) Data preprocessing: Convert the original monthly runoff series into a supervised learning format with 12-month time steps and perform MinMax normalization.

(2) Feature extraction: The CNN layer automatically learns local fluctuation patterns of the runoff series (such as seasonal mutations). The output feature map has its dimensions compressed to 6×16 after pooling.

(3) Temporal modeling: The BiLSTM layer analyzes the long-term dependencies of the feature map from forward/backward dimensions, outputting a 25-dimensional hidden state.

(4) Result reconstruction: The fully connected layer maps the LSTM output to

the predicted value, and the final runoff prediction result is obtained after inverse normalization.

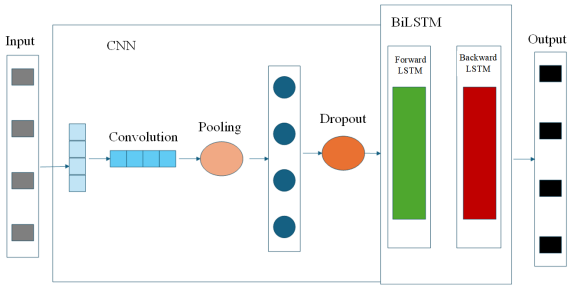


Figure 4: *CNN-BiLSTM model structure diagram*

2.6 ISCBL

First, the original runoff series is initially decomposed by the ICEEMDAN model to obtain a series of relatively smooth IMFs and a residual. Then, the sample entropy of each component is calculated to quantify its complexity and irregularity (Liu and Zhao, 2024). Based on this, the IMF with the highest sample entropy value is selected for secondary decomposition using the SSA model to extract the main periodic components to enhance its stationarity and predictability. Then, each processed component is input into the CNN-BiLSTM prediction model for training and prediction. Finally, the final runoff prediction value is reconstructed by integrating the prediction results of each component. The overall flow of the ISCBL model is shown in Fig. 5.

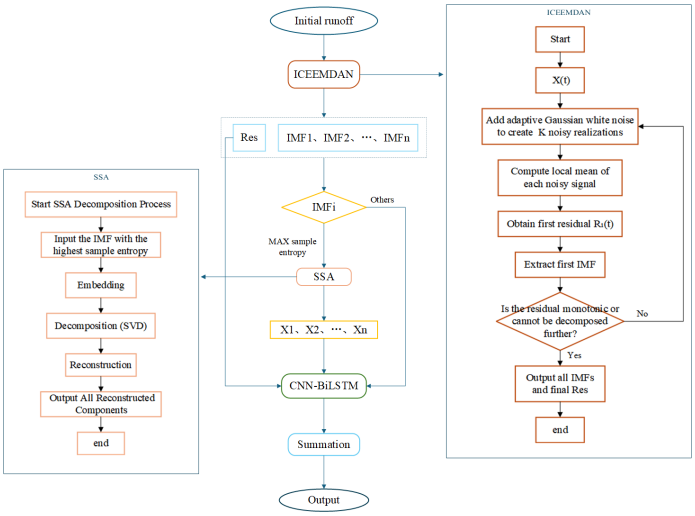


Figure 5: *ISCBL model construction flowchart*

2.7 Evaluation Metrics

In order to evaluate the veracity of the model's predictions, this study has chosen several widely recognized evaluation metrics, namely the mean absolute error (MAE), root mean square error (RMSE), Mean Absolute Percentage Error (MAPE), Nash-Sutcliffe efficiency coefficient (NSE), and Coefficient of Determination (R^2). These metrics are utilized to thoroughly assess the model's performance. The corresponding formulas for calculating these evaluation indicators are presented as follows:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \qquad \text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{2.12}$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|\hat{y}_i - y_i|}{y_i} \times 100\%, \qquad \text{NSE} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \hat{y}_{\text{avg}})^2} \tag{2.13}$$

$$R^2 = \left[\frac{\sum_{i=1}^{n}(y_i - y_{\text{avg}})(\hat{y}_i - \hat{y}_{\text{avg}})}{\sqrt{\sum_{i=1}^{n}(y_i - y_{\text{avg}})^2 \sum_{i=1}^{n}(\hat{y}_i - \hat{y}_{\text{avg}})^2}}\right]^2, \tag{2.14}$$

where y_i is the actual value of the i-th sample. $\hat{y}_i$ is the predicted value of the i-th sample. y_{avg} is the average of all measured values. $\hat{y}_{avg}$ is the average of all predicted values.

3 Case Analysis

3.1 Study Area and Data

The total area of the Yellow River basin is about 795,000 square kilometers, making it the second largest basin in China. This study uses monthly runoff data from the Lanzhou hydrological station in the upper reaches of the Yellow River basin. The Lanzhou station is located on the main stream of the upper Yellow River, with a

length of 152 kilometers and a controlled drainage area of about 222,600 square kilometers. Its multi-year average runoff volume is 31.44 billion cubic meters. As an important controlling hydrological station on the upper Yellow River, it is a key node for water resources dispatch, flood control, and disaster reduction in the Yellow River basin.

The hydrological process in this area is influenced by both glacier-snow melt from the Tibetan Plateau and monsoon precipitation, presenting complex non-stationary characteristics (Xu et al., 2025), providing a typical sample for hydrological model research. This paper uses monthly runoff data from January 1971 to October 2023 from the Lanzhou hydrological station (see Fig. 6). Data from January 1972 to June 2013 is used for model training, and data from July 2013 to October 2023 is used for model testing.

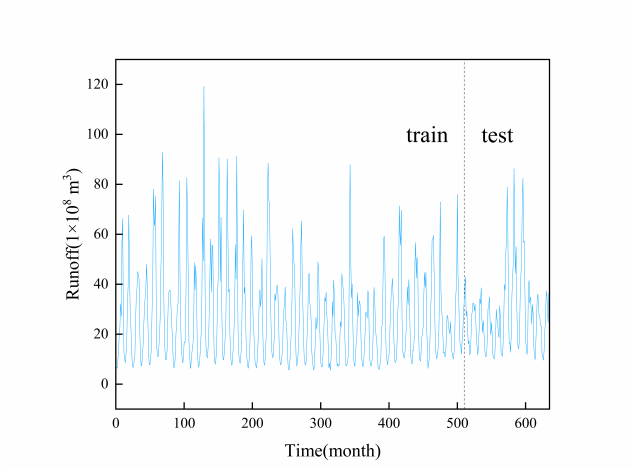


Figure 6: *Measured monthly runoff time series at Lanzhou station*

3.2 ICEEMDAN Decomposition Results

The original runoff series is decomposed using the ICEEMDAN method, obtaining multiple IMFs and a final residual (see Fig. 7). Each component exhibits characteristics of different time scales. Low-frequency components reflect long-term trends, while high-frequency components reflect short-term fluctuations. These decomposed components carry distinct hydrological signatures. The high-frequency IMFs (e.g., IMF1-2) correspond to short-term hydrological processes such as intense rainfall events and rapid snowmelt pulses, which are common in the upper Yellow River basin due to its monsoon-influenced and alpine climate. In contrast, the low-frequency IMFs (e.g., IMF5-6) and the residual component capture slower, long-term variations related to seasonal baseflow dynamics, gradual groundwater discharge,

and the overarching impacts of climate change or human activities. This multi-scale separation effectively disentangles the complex runoff signals, providing more stationary and interpretable inputs for the subsequent prediction model.

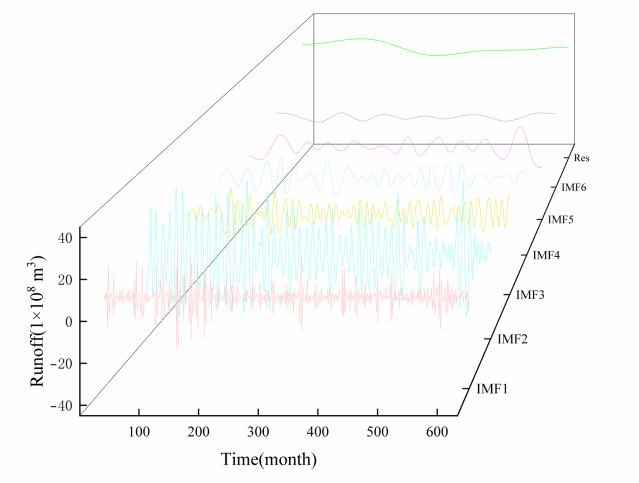


Figure 7: *ICEEMDAN decomposition results*

3.3 SSA Secondary Decomposition

Before secondary decomposition, utilize the sample entropy to assess the complexity of the decomposed components. The sample entropy of each component obtained from the ICEEMDAN decomposition is calculated. The calculation results are shown in Table 2. Screen out the component with the highest complexity, and decompose it twice using the SSA technique to extract the main periodic components (see Fig. 8), effectively reducing noise interference.

Table 2: *Sample entropy calculation results table*

Component Name	Sample Entropy Value
IMF1	1.2393
IMF2	0.72042
IMF3	0.62261
IMF4	0.54037
IMF5	0.306
IMF6	0.15483

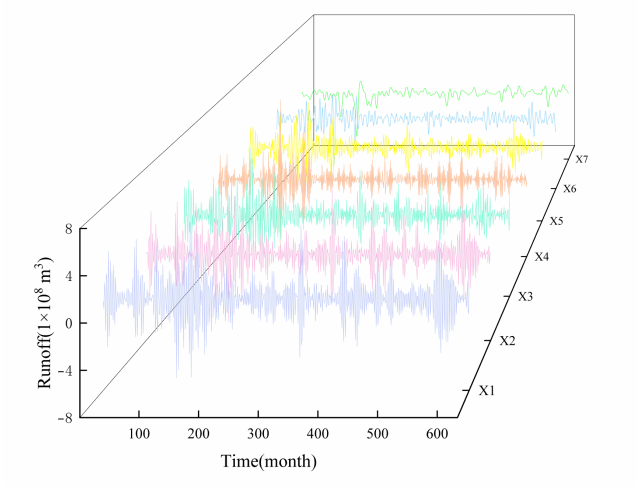


Figure 8: *SSA decomposition results*

3.4 Determination of Lag Time

ACF and PACF indicators guide the selection of input variables. According to the ACF and PACF curves of the Lanzhou station shown in Fig. 9, the ACF peaks at lag 12 (months), while most PACF values fall within the confidence interval. Therefore, the input variable is set to 12, meaning that subsequent prediction data will be modeled based on the previous 12 data points.

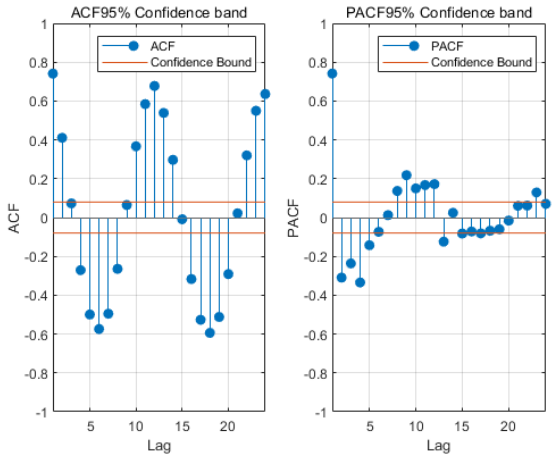
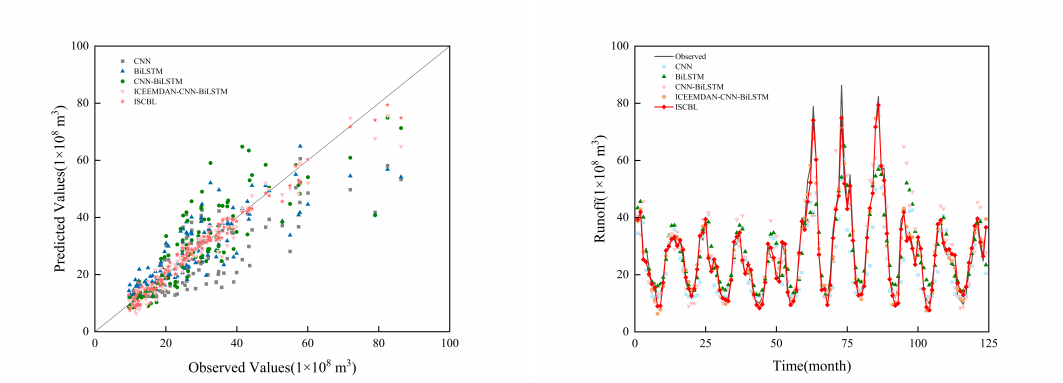


Figure 9: *ACF and PACF curves of Lanzhou station*

3.5 Results and Discussion

To verify the effectiveness of the combined model, the CNN, BiLSTM, and CNN-BiLSTM models without ICEEMDAN decomposition processing, the ICEEMDAN-CNN-BiLSTM model with ICEEMDAN decomposition processing, and the final ISCBL model are set up for monthly runoff prediction. The prediction results of each model are compared. The prediction effects of the CNN, BiLSTM, CNN-BiLSTM, primary decomposition model, and the final ISCBL model are shown in Fig. 10a. The fitting situation of each model with the measured values is shown in Fig. 10b. Comparative analysis shows that the predicted values of the ISCBL combined model fit best with the actual values, demonstrating the best simulation effect.



(a) *Prediction effects of different models*

(b) *Fitting situation of predictions from different models*

Figure 10: *Model comparison results*

The prediction performance comparison of each model is shown in Table 3:

Table 3: *Comparison of prediction performance of different models*

model	MAE	RMSE	MAPE(%)	NSE	R^2
CNN	7.2911	9.7137	28.4681	0.6022	0.6192
BiLSTM	6.0782	8.5830	22.5859	0.6894	0.6969
CNN-BiLSTM	5.6101	8.1864	18.9827	0.7175	0.7408
ICEEMDAN-CNN-BiLSTM	2.4935	3.6929	10.0394	0.9425	0.9452
ISCBL	1.8797	2.5172	8.3486	0.9733	0.9750

As model complexity increases and decomposition methods are optimized, prediction accuracy improves significantly. Single models (CNN, BiLSTM) have obvious limitations: CNN performs inadequately in capturing long-term temporal dependencies,

with an NSE of only 0.6022, while BiLSTM, although improving temporal modeling, still has a MAPE of 22.59%, indicating insufficient extraction of local features.

The simple combination of the two (CNN-BiLSTM) partially compensates for the defects, increasing the NSE to 0.7175, but its RMSE is as high as 8.1864, failing to solve the non-stationarity problem of the original data.

The primary decomposition model (ICEEMDAN-CNN-BiLSTM) significantly improves performance through signal decomposition, achieving an NSE of 0.9425 and reducing MAE by 55%, proving that signal decomposition can effectively handle the non-stationarity of complex time series data. However, with MAPE=10.04%, high-frequency component noise still affects accuracy.

The final hybrid model combining ICEEMDAN-SSA secondary decomposition and CNN-BiLSTM shows comprehensive advantages. The NSE reaches 0.9733, which is 61.7% and 41.2% higher than the single CNN and BiLSTM models, respectively. The MAE is further reduced by 25% compared to the primary decomposition model. Significant improvements are seen in all other indicators, verifying the superiority of the ”decomposition enhancement + deep learning fusion” strategy.

Fig. 11 shows the violin plots for the Lanzhou station, where Model 1 to Model 5 correspond to the CNN, BiLSTM, CNN-BiLSTM, ICEEMDAN-CNN-BiLSTM, and ISCBL models, respectively. The violin plot of the measured values reflects the statistical distribution of the real runoff. The prediction result distributions of the undecomposed models (CNN, BiLSTM, CNN-BiLSTM) show significant differences from the measured values, indicating that single or simple hybrid models have limited ability to capture complex nonlinear patterns.

The distribution of ICEEMDAN-CNN-BiLSTM is closer to the measured values, indicating that decomposition processing effectively extracts different frequency components in the sequence, improving prediction accuracy. The ISCBL model shows the best distributional agreement with the measured values, indicating that secondary decomposition further optimizes the model's fitting and generalization ability.

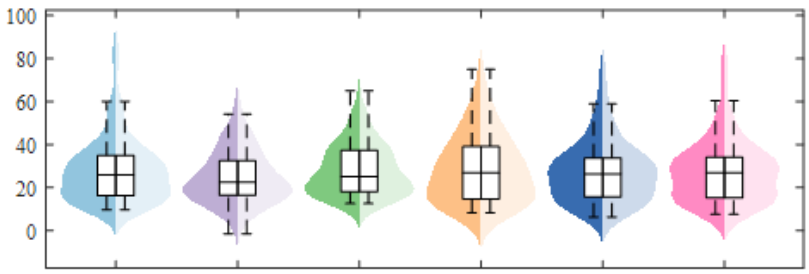


Figure 11: *Violin plots of different models*

4 Conclusions

Targeting the nonlinear and non-stationary characteristics of the runoff series at the Lanzhou Hydrological Station in the upper Yellow River, this paper proposes a novel hybrid prediction model based on ICEEMDAN-SSA secondary decomposition and CNN-BiLSTM for the nonlinear and non-stationary characteristics of the runoff series at the Lanzhou hydrological station in the upper Yellow River. Validation using monthly runoff observation data from 1971 to 2023 shows that the model exhibits significant advantages under complex hydrological conditions. Secondary decomposition can effectively extract multi-scale features of the runoff series, significantly improving prediction accuracy. The CNN-BiLSTM hybrid model can capture spatiotemporal features simultaneously, demonstrating superior performance in runoff prediction. The ISCBL model achieves an NSE of 0.9733 and a MAPE of 8.35% on the test set, providing an effective tool for runoff prediction in the upper Yellow River. The promising validation at the Lanzhou station notwithstanding, the framework's generalizability across basins with diverse climatic and geographical settings necessitates further multi-station testing. Future work will address these aspects through expanded case studies, comprehensive sensitivity analyses, and the integration of additional data sources to enhance the model's applicability and reliability under varying hydrological regimes.

Disclaimer (Artificial Intelligence)

Authors hereby declare that generative AI technologies (ChatGPT 4o, OpenAI) have been used during manuscript editing for polishing academic expression, optimizing logical coherence.

Details of the AI usage are given below:

1. AI tool: ChatGPT (Version 4o), OpenAI (https://openai.com/).
2. Usage purpose: Academic expression polishing and grammatical revision of the manuscript.
3. Post-use step: The authors reviewed all AI-edited content and take full responsibility for the publication content.

Acknowledgement

This work was supported in part by National Natural Science Foundation of China (Grant No. 12272135), the Key Scientific Research Projects of Henan Higher Education Institutions (Grant No. 24B110007); the Graduate Education Reform and Quality Improvement Project of Henan Province (Grant No. YJS2025KC03, YJS2026AL002).

Competing Interests

Authors have declared that no competing interests exist.

References

Saha, A., & Chandra Pal, S. (2024). Application of machine learning and emerging remote sensing techniques in hydrology: A state-of-the-art review and current research trends. *Journal of Hydrology*, 632, 130907.

Ma, X., Li, Z., Ren, Z., Shen, Z., Xu, G., & Xie, M. (2024). Predicting future impacts of climate and land use change on streamflow in the middle reaches of China's Yellow River. *Journal of Environmental Management*, 370, 123000.

Du, S., Jiang, S., Ren, L., Yuan, S., Yang, X., Liu, Y., Gong, X., & Xu, C.-Y. (2023). Control of climate and physiography on runoff response behavior through use of catchment classification and machine learning. *Science of The Total Environment*, 899, 166422.

Nanda, A., & Sen, S. (2021). A complex network theory based approach to better understand the infiltration-excess runoff generation thresholds. *Journal of Hydrology*, 603, 127038.

Xu, W., Chen, J., & Corzo, G. (2025). Combining data augmentation and hybrid modeling approaches for deep learning-based monthly streamflow forecasting. *Journal of Hydrology*, 659, 133318.

Colominas, M. A., Schlotthauer, G., & Torres, M. E. (2014). Improved complete ensemble EMD: A suitable tool for biomedical signal processing. *Biomedical Signal Processing and Control*, 14, 19–29.

Gu, J., Hung, K., Ling, B. W.-K., Chow, D. H.-K., Zhou, Y., Fu, Y., & Pun, S. H. (2024). Generalized singular spectrum analysis for the decomposition and analysis of non-stationary signals. *Journal of the Franklin Institute*, 361(6), 106696.

Sibtain, M., Bashir, H., Nawaz, M., Hameed, S., Imran Azam, M., Li, X., Abbas, T., & Saleem, S. (2022). A multivariate ultra-short-term wind speed forecasting model by employing multistage signal decomposition approaches and a deep learning network. *Energy Conversion and Management*, 263, 115703.

Chen, S., Ren, M., & Sun, W. (2021). Combining two-stage decomposition based machine learning methods for annual runoff forecasting. *Journal of Hydrology*, 603, 126945.

Li, M., Zhou, Q., Han, X., & Lv, P. (2024). Prediction of reference crop evapotranspiration based on improved convolutional neural network (CNN) and long short-term memory network (LSTM) models in Northeast China. *Journal of Hydrology*, 645, 132223.

Li, Z., Liu, F., Li, H., Liu, M., Li, Z., Feng, Q., Xu, B., & Liu, X. (2025). Driving forces of the spatiotemporal supply pattern of runoff in the source region of Yellow River. *Journal of Hydrology: Regional Studies*, 60, 102515.

Xie, Y., Sun, W., Ren, M., Chen, S., Huang, Z., & Pan, X. (2023). Stacking ensemble learning models for daily runoff prediction using 1D and 2D CNNs. *Expert Systems with Applications*, 217, 119469.

Xue, H., Guo, C., Dong, G., Zhang, C., Lian, Y., & Yuan, Q. (2025). Prediction of runoff in the upper reaches of the Hei River based on the LSTM model guided by physical mechanisms. *Journal of Hydrology: Regional Studies*, 58, 102218.

Liu, J., Zhang, X., Wu, X., Yang, Y., & Zheng, Y. (2024). Impacts of LULC changes on runoff from rivers through a coupled SWAT and BiLSTM model: A case study in Zhanghe River Basin, China. *Ecological Informatics*, 84, 102866.

Hu, F., Yang, Q., Yang, J., Luo, Z., Shao, J., & Wang, G. (2024). Incorporating multiple grid-based data in CNN-LSTM hybrid model for daily runoff prediction in the source region of the Yellow River Basin. *Journal of Hydrology: Regional Studies*, 51, 101652.

Liu, X., & Zhao, H. (2024). Analyzing watershed system state through runoff complexity and driver interactions using multiscale entropy and deep learning. *Ecological Indicators*, 168, 112779.

Fang, X., Wang, Y., Lv, S., Liu, L., Fan, S., & Liu, L. (2026). Lightweight Transformer-based nonlinear equalization for robust IQ recovery in WDM PDM CO-OFDM systems. *Optics Communications*, 607, 132837-132837.

Liang, J., Yue, J., Xin, Y., Pan, S., Tian, J., & Sun, J. (2026). Short-Term photovoltaic power forecasting based on K-means++ clustering, secondary decomposition and TCN-BiLSTM-Attention model. *Electric Power Systems Research*, 255, 112749-112749.

Yang, Z., Dong, Q., Zhang, X., Zhu, H., & Cheng, Z. (2026). Development of a Two-Stage LSTM for Multi-Step Runoff Forecasting Using a XAJ Model and EEMD. *Water Resources Management*, 40(2), 81-81.

WEI, W., & Yuan, Y. (2026). VMD-DCA-BiGRU wind turbine bearing fault diagnosis method with attention mechanism integration. *Measurement*, 266, 120466-120466.