


Hydrological time series prediction by extreme learning machine and sparrow search algorithm

Bao-fei Feng*, Yin-shan Xu, Tao Zhang  and Xiao Zhang

Bureau of Hydrology, Changjiang Water Resources Commission, Wuhan 430010, China

*Corresponding author. E-mail: fengbf@cjh.com.cn

 TZ, 0000-0002-1842-2717

ABSTRACT

In general, accurate hydrological time series prediction information is of great significance for the rational planning and management of water resource systems. Extreme learning machine (ELM) is an effective tool proposed for the single-layer feedforward neural network in regression and classification problems. However, the standard ELM model falls into a local minimum with a high probability in hydrological prediction problems since the randomly assigned parameters (like input-hidden weights and hidden biases) often remain unchanged in the learning process. For effectively improving the prediction accuracy, this paper develops a hybrid hydrological forecasting model where the emerging sparrow search algorithm (SSA) is firstly used to determine the satisfying parameter combinations of the ELM model, and then the Moore–Penrose generalized inverse method is chosen to analytically obtain the weight matrix between the hidden layer and output layer. The proposed method is used to forecast the long-term daily runoff series collected from three real-world hydrological stations in China. Based on several performance evaluation indexes, the results show that the proposed method outperforms several ELM variants optimized by other evolutionary algorithms in both training and testing phases. Hence, an effective evolutionary machine-learning tool is developed for accurate hydrological time series forecasting.

Key words: extreme learning machine, hydrologic forecasting, sparrow search algorithm

HIGHLIGHTS

- A hybrid method using extreme learning machine and sparrow search algorithm is developed for hydrological forecasting.
- The proposed method finds better results than several traditional methods, providing an alternative tool for hydrological forecasting.

1. INTRODUCTION

Hydrological time series forecasting has been an important research theme in hydrology, water resources and other related engineering construction fields (Du & Weng 2021; Kaluarachchi 2021; Zhang & Ariaratnam 2021). Accurate hydrological prediction information plays an important role in many producing activities (Chen *et al.* 2019; Pérez Lespier *et al.* 2019; Feng *et al.* 2021a; Niu *et al.* 2021a, 2021b), like hydropower production, flood control, peak operation, and water supply. Due to the comprehensive impact of various factors (like meteorological process, underlying surface, and human activities), the natural hydrological process often shows the characteristics of strong nonlinearity, high mutability, and uncertainty (Peng *et al.* 2019; Won & Kim 2020; Sun *et al.* 2021). Then, many researchers and engineers all over the world have paid great attention to deepening the research and development of hydrological forecasting methods (Tamura *et al.* 1990; Jiang *et al.* 2018; Ghasemlounia & Sagheblian 2021). In the past few years, numerous hydrological forecasting approaches have been successfully developed and these existing approaches can be roughly classified into two types: process-based approaches and data-based approaches (Leon *et al.* 2020; Tian 2020; Viccione *et al.* 2020; Birbal *et al.* 2021). In general, the process-based approaches usually involve advanced physical formulas and abundant knowledge to describe the complicated meteorological and hydrological processes. Due to the limitation of the cognitive levels and model descriptions, unsatisfactory forecasting results often become unavoidable, which limits their applications in practical engineering (Chen *et al.* 2020; Dalkılıç & Hashimi 2020; Wu *et al.* 2020). With the

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY-NC-ND 4.0), which permits copying and redistribution for non-commercial purposes with no derivatives, provided the original work is properly cited (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

rapid development of information technology, many data-based approaches have been introduced to improve this problem, like artificial neural network, adaptive neuro-fuzzy inference system, and support vector machine (Chen & Chau 2016; Moghayedj & Windapo 2019; Adib *et al.* 2021; Ji *et al.* 2021). By gaining knowledge from history, data-based approaches gradually produce satisfactory forecasting results in hydrology even though the detailed physical process of the hydrological time series is not well understood (Feng & Niu 2021; Feng *et al.* 2021b; Niu *et al.* 2021c). Owing to their easy implementation and high forecasting ability, data-based approaches are becoming more and more popular in practical engineering (Guo *et al.* 2013; Yuan *et al.* 2018; He *et al.* 2019).

Single-layer feedforward neural network (SLFN) has proved to be an effective tool for regression and classification problems (Chua & Wong 2011). Generally, traditional gradient-based training tools are widely used to determine the computation parameters of the SLFN model. However, many applications show that the gradient-based methods are easily trapped into local minima because the network structure features are not well considered. To alleviate the defects of the gradient-based methods, extreme learning machine (ELM) has been successfully developed in recent years (Peng *et al.* 2017; Zhou *et al.* 2018; Luo *et al.* 2019). In ELM, the input-hidden weights and hidden biases are randomly determined to analytically obtain the hidden-output weights by the Moore–Penrose generalized inverse method. Compared with the gradient-based methods, ELM shows the advantages of faster training speed, stronger generalization ability and fewer computation parameters. Nevertheless, it is difficult for the ELM model to achieve optimal results because the network parameters are determined in a random manner. In other words, the standard ELM method tends to fall into a local minimum because the values of input-hidden weights and hidden biases are not well chosen. In order to further improve the generalization ability of the ELM method, meta-heuristic evolutionary algorithms are used to optimize model structure and hyperparameters (Niu *et al.* 2021d, 2021e), like particle swarm optimization, gravitational search algorithm, cooperation search algorithm and sine cosine algorithm (Mei *et al.* 2018; Niu *et al.* 2021f).

As a novel swarm intelligence algorithm, the sparrow search algorithm (SSA), inspired by the sparrow's wisdom, foraging and anti-predation behaviors, has been developed to solve the global optimization problems (Tuerxun *et al.* 2021; Yang *et al.* 2021; Yuan *et al.* 2021). In SSA, several modules are carefully designed to balance exploration and exploitation, like producers for searching for food, scroungers for monitoring the producers and watchers for avoiding danger (Truchet *et al.* 2016; Wang *et al.* 2021; Zhang & Ding 2021). The SSA method is employed to deal with a group of numerical functions and engineering problems. The results show that the SSA method outperforms several mature evolutionary algorithms with respect to search rate, solution precision, and local minimum avoidance. Due to its satisfactory search capability, the SSA method is gradually becoming popular in many research fields. However, there are few reports about using the SSA method to improve ELM performance in hydrological forecasting. To fill this research gap, this research develops a hybrid forecasting method where the SSA method is used to search for satisfying network parameters of the ELM model. Then, the hybrid method is used to forecast the long-term runoff time series in different working conditions. The comparative results demonstrate that the developed method outperforms several traditional methods in both training and testing phases. Therefore, it can be concluded that the SSA method is a useful optimizer for finding a better neural network structure for accurate time series simulation, while an effective evolutionary extreme learning machine tool is provided here for hydrological forecasting.

The rest of this study is summarized as below: Section 2 gives information on the hybrid method; Section 3 compares the engineering practicability of the proposed method in hydrological forecasting; and the conclusions are given at the end.

2. METHODS

2.1. Sparrow search algorithm (SSA)

Sparrow search algorithm (SSA) is an emerging evolutionary algorithm based on the sparrow's foraging and anti-predation behaviors (Zhou & Wang 2021). Compared with several traditional evolutionary algorithms, SSA has stronger global search ability and faster convergence speed in global optimization problems. In SSA, the population is divided into two different groups: one is the producer group possessing larger search steps to find food, and another is the scrounger group that follows the producers to find food. During the search process, the scroungers have larger probabilities to find food via the following behaviors, while the roles of both producers and scroungers are adjusted dynamically to find more high-quality food sources. Then, the mathematical model of the SSA method is given as below:

Step 1: Parameter definition, of the number of sparrows (N), the number of producers (PN) and scroungers ($N-PN$), the maximum iteration (g_{\max}). The location of the i th sparrow can be defined as $\mathbf{x}_i = (x_{i,1}, x_{i,2}, \dots, x_{i,D})$ while $f(\mathbf{x}_i)$ represents

the fitness value of the i th sparrow. Then, the initial swarm \mathbf{x} can be expressed as below:

$$\mathbf{x} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,d} \\ \vdots & x_{ij} & \vdots \\ x_{N,1} & \cdots & x_{N,D} \end{bmatrix} \quad (1)$$

where x_{ij} is the j th element's value of the i th sparrow. D is the number of decision variables.

Step 2: The producers' positions are updated by:

$$\mathbf{x}_i^{g+1} = \begin{cases} \mathbf{x}_i^g \cdot \exp\left[\frac{-i}{\alpha \cdot g_{\max}}\right] & \text{if } (R < ST) \\ \mathbf{x}_i^g + Q \cdot \mathbf{L} & \text{else } i \in [1, PN], \end{cases} \quad (2)$$

where g is the iteration index; α is a random number uniformly distributed in the range of $[0, 1]$; Q is a random number obeying the standard normal distribution; and \mathbf{L} is a $1 \times D$ matrix whose elements are set as 1. $R \in [0, 1]$ and $ST \in [0.5, 1.0]$ denote the alarm value and safety threshold. If $R < ST$, the producers will execute the extensive search mode without the predators' influence; if $R \geq ST$, the predators have been found by some sparrows and all the sparrows should fly to the safe areas.

Step 3: The scroungers' positions are updated by:

$$\mathbf{x}_i^{g+1} = \begin{cases} Q \cdot \exp\left[\frac{\mathbf{G}_{\text{worst}} - \mathbf{x}_i^g}{i^2}\right] & \text{if } (i > N/2) \\ \mathbf{S}_{\text{best}} + |\mathbf{x}_i^g - \mathbf{S}_{\text{best}}| \cdot \mathbf{A}^+ \cdot \mathbf{L} & \text{else } i \in [PN + 1, N], \end{cases} \quad (3)$$

where \mathbf{S}_{best} is the producer's best-known location; $\mathbf{G}_{\text{worst}}$ is the global worst-known location found by the swarm; \mathbf{A} is a $1 \times D$ matrix whose elements are randomly chosen from the set $\{1, -1\}$; and $\mathbf{A}^+ = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}$. If $i > N/2$, the i th scrounger should search in other areas to find energy; otherwise, the i th scrounger is foraging in the area around \mathbf{S}_{best} .

Step 4: To avoid possible danger, about 10%–20% of sparrows in the swarm are randomly selected as the scouts and their positions are updated by:

$$\mathbf{x}_i^{g+1} = \begin{cases} \mathbf{G}_{\text{best}} + \beta \cdot |\mathbf{x}_i^g - \mathbf{G}_{\text{best}}| & \text{if } (f(\mathbf{x}_i^g) > f(\mathbf{G}_{\text{worst}})) \\ \mathbf{x}_i^g + \frac{K \cdot |\mathbf{x}_i^g - \mathbf{G}_{\text{worst}}|}{f(\mathbf{x}_i^g) + f(\mathbf{G}_{\text{worst}}) + \theta} & \text{if } (f(\mathbf{x}_i^g) = f(\mathbf{G}_{\text{worst}})) \end{cases} \quad (4)$$

where β is a random number obeying the standard normal distribution; \mathbf{G}_{best} is the global best-known location found by far; $K \in [-1, 1]$ is a random number representing the search step size; and θ is a small constant used to avoid the denominator being zero. If $f(\mathbf{x}_i^g) > f(\mathbf{G}_{\text{worst}})$, the i th sparrow at the edge of the swarm will easily find predators; otherwise, the i th sparrow at the center of the swarm should be close to other sparrows for antipredation.

Step 5: Update the best as well as worst fitness values of the swarm to get all the sparrows' new positions.

Step 6: If the termination condition is not met, return to Step 2 for the next cycle; otherwise, the global best-known position found by the sparrow population will be treated as the final solution for the target problem.

2.2. Extreme learning machine (ELM)

Extreme learning machine (ELM), shown in Figure 1, is an effective training tool developed to resolve a single-layer feedforward neural network (SLFN). After randomly determining the values of the input-hidden weights and hidden biases, ELM makes use of the classical Moore–Penrose generalized inverse to analytically calculate the hidden-output weights (Huang *et al.* 2006a; Cambria *et al.* 2013; Huang *et al.* 2014). Different from the conventional learning methods for SLFN, ELM makes an obvious reduction in the size of the computational parameters. In this way, ELM possesses the merits of faster execution speed, fewer learning parameters and stronger generalization ability in comparison with the traditional gradient-based methods (Yadav *et al.* 2016).

For a training dataset with N samples, the outputs of the ELM model with L hidden neurons can be given as below:

$$\mathbf{o}_i = \sum_{l=1}^L \beta_l \cdot g(\mathbf{w}_l \cdot \mathbf{a}_i + b_l), i = 1, 2, \dots, N \quad (5)$$

where \mathbf{a}_i and \mathbf{o}_i are the inputs and outputs associated with the i th training sample; \mathbf{w}_l is the weight vector connecting the l th hidden neuron with all the neurons in the input layer; β_l is the weight vector connecting the l th hidden neuron with all the neurons in the output layer; b_l is the bias of the l th hidden neuron; and $g(\cdot)$ denotes the activation function of the l th hidden neuron.

In the ELM theory, it is believed that the neural network model is capable of ideally approximating all the considered training data without any deviation. Then, the above equation can be modified as below:

$$\mathbf{H}\beta = \mathbf{O} \quad (6)$$

$$\mathbf{H} = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{a}_1 + b_1) & \cdots & g(\mathbf{w}_L \cdot \mathbf{a}_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(\mathbf{w}_1 \cdot \mathbf{a}_N + b_1) & \cdots & g(\mathbf{w}_L \cdot \mathbf{a}_N + b_L) \end{bmatrix}_{N \times L} \quad (7)$$

$$\beta = \begin{bmatrix} \beta_{1,1} & \cdots & \beta_{1,m} \\ \vdots & \ddots & \vdots \\ \beta_{L,1} & \cdots & \beta_{L,m} \end{bmatrix}_{L \times m} \quad (8)$$

$$\mathbf{O} = \begin{bmatrix} o_{1,1} & \cdots & o_{1,m} \\ \vdots & \ddots & \vdots \\ o_{N,1} & \cdots & o_{N,m} \end{bmatrix}_{N \times m} \quad (9)$$

where \mathbf{H} is the output matrix associated with the hidden layer; β is the weight matrix connecting the hidden layer with the output layer; and \mathbf{O} is the output matrix associated with all training data.

From the mathematical viewpoint, \mathbf{H} and \mathbf{O} can be regarded as an independent matrix and dependent matrix of N training samples, and then β becomes the coefficient matrix to be determined. By this time, Equation (6) will be transformed to a standard linear system and then the weight matrix β can be analytically deduced by determining the optimal solution of the above

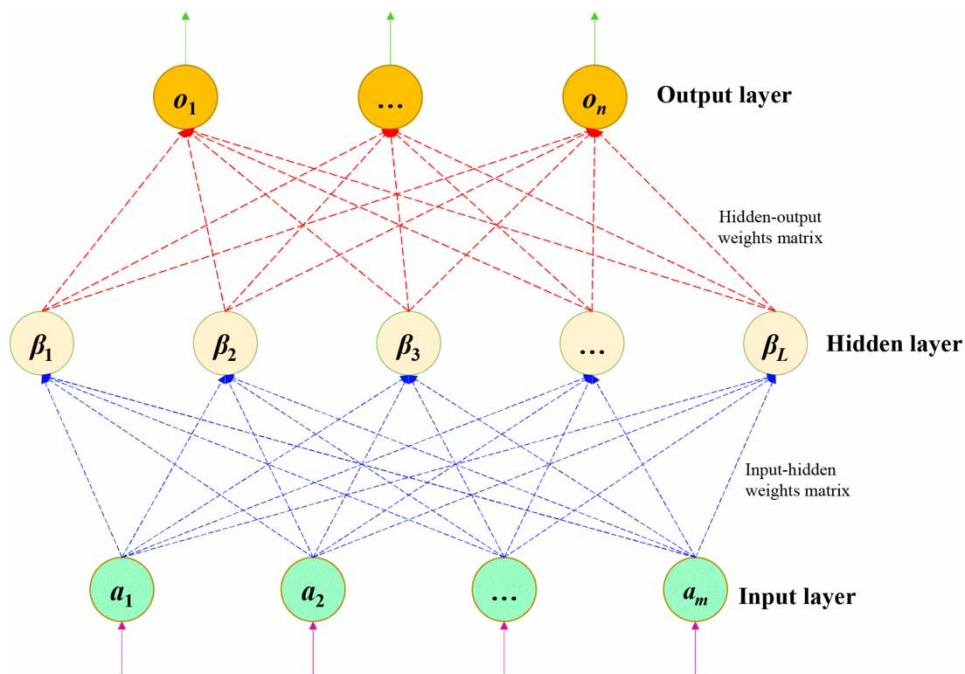


Figure 1 | Sketch map of the extreme learning machine model.

linear system, which can be expressed as below:

$$\tilde{\beta} = H^{\dagger} O \quad (10)$$

where H^{\dagger} denotes the Moore–Penrose generalized inverse matrix of H .

2.3. Hydrological time series forecasting model based on ELM and SSA

As mentioned above, the ELM model can achieve good performance and make an obvious reduction in the execution time by stochastically determining the model parameters, rather than by iterative adjustment. ELM has the merits of faster training speed and stronger learning ability (Huang *et al.* 2006b; Cao *et al.* 2012; Huang *et al.* 2015). However, the random determination of parameters associated with the hidden layer may fail to produce suboptimal solutions, which then affects the forecasting capability of the model in practice. In order to address this defect, this paper develops an effective ELM-SSA model where the hidden biases and input-hidden weights of the ELM model are iteratively optimized by the SSA method, rather than the random assignment at the initial phase and no adjustment in the late learning phase. Figure 2 shows the flowchart of the proposed method. For the sake of simplicity, the proposed method with n input variables, L hidden neurons and one output variable is used to forecast the hydrological time series. In other words, the developed method has n input nodes, L hidden nodes and 1 output node. Then, the execution steps of the proposed method are given as:

Step 1: Data normalization. All the considered data should be normalized into the range of [0,1] before being divided into the training and testing datasets, which can be expressed as below:

$$Q'_i = a \cdot \frac{Q_i - \min_{1 \leq i \leq n} \{Q_i\}}{\max_{1 \leq i \leq n} \{Q_i\} - \min_{1 \leq i \leq n} \{Q_i\}} + b \quad (11)$$

where Q'_i and Q_i are the i th normalized and original data; n is the amount of data; and a and b are the adjusting coefficients.

Step 2: Parameter setting, such as the number of sparrows and iterations in SSA, the number of hidden neurons as well as the activation function of hidden neurons in ELM. Here, the classical sigmoid function is chosen as the activation module for data mapping in the hidden layer.

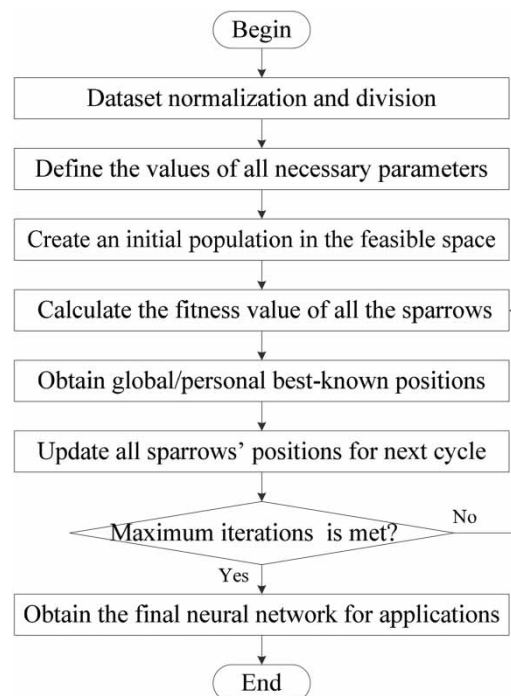


Figure 2 | Flowchart of the proposed method.

Step 3: Population initialization. Set the counter $g = 1$ and then randomly create all the element values of each sparrow of the initial swarm in the feasible state space. Each sparrow represents a possible ELM model containing all the parameters associated with the hidden layer.

Step 4: Problem evaluation. Calculate the hidden-output weights to obtain the fitness value of all the sparrows. For the i th sparrow at the g th cycle \mathbf{x}_i^g , its fitness value $f(\mathbf{x}_i^g)$ is obtained by the following formulas:

$$\hat{\boldsymbol{\beta}}_{i,g} = \mathbf{H}_{i,g}^\dagger \mathbf{O} \quad (12)$$

$$\mathbf{H}_{i,g} = \begin{bmatrix} g(\mathbf{w}_{1,i,g} \cdot \mathbf{a}_1 + b_{l,i,g}) & \cdots & g(\mathbf{w}_{L,i,g} \cdot \mathbf{a}_1 + b_{L,i,g}) \\ \vdots & \ddots & \vdots \\ g(\mathbf{w}_{1,i,g} \cdot \mathbf{a}_N + b_{l,i,g}) & \cdots & g(\mathbf{w}_{L,i,g} \cdot \mathbf{a}_N + b_{L,i,g}) \end{bmatrix}_{N \times L} \quad (13)$$

$$f(\mathbf{x}_i^g) = \sqrt{\frac{1}{N} \sum_{s=1}^N \left\| o_s - \sum_{l=1}^L \hat{\boldsymbol{\beta}}_{l,i,g} g(\mathbf{w}_{l,i,g} \cdot \mathbf{a}_s + b_{l,i,g}) \right\|_2^2} \quad (14)$$

where N is the number of training samples; o_s is the outputs of the s th training sample; $\mathbf{H}_{i,g}$ is the hidden-layer output matrix; and $\mathbf{H}_{i,g}^\dagger$ represents the Moore–Penrose generalized inverse of $\mathbf{H}_{i,g}$.

Step 5: Population updating. Obtain the global best-known or worst-known sparrows of the swarm, as well as the best-known locations of all the producers. Next, compute the necessary computation parameters to update all the sparrows' positions.

Step 6: Set $g = g + 1$. If the maximum iteration is met, go to Step 4 for the next cycle; otherwise, the global best sparrow is treated as the ideal parameters of the hidden layer and then the Moore–Penrose generalized inverse method is used to obtain the hidden-output weights. The optimized ELM model by far is obtained for applications.

3. CASE STUDIES

3.1. Engineering background

To test the feasibility of the proposed method, the Yangtze River of China is selected as the case study. The flood events at Yangtze River usually occur in the wet season between May and October, while the other months belong to the dry season. During the flood season, the spatial–temporal distribution of rainfall is largely affected by monsoon activities and subtropical anticyclones. Meanwhile, the many huge reservoirs represented by the Three Gorges project, the world's largest hydropower plant, are put into operation in succession, resulting in great changes in runoff features. As a result, it becomes more and more difficult to accurately forecast the runoff time series under the changing environment.

The daily runoff time series collected from three hydrological stations located on the Yangtze River are used for comparative study. The three hydrological stations (A ~ C) play an important role in monitoring the water changing tendency and guaranteeing the safe operation of Yangtze River. Specially, hydrological station B is located on the south tributary, while hydrological stations A and C are located on the mainstream of Yangtze River. The water at stations A and B flows to station C. Figure 3 illustrates the studied daily runoff time series of the three hydrological stations. The forecasting model should have strong adaptive capacities to respond to the obvious differences of the runoff series at the three hydrological stations. For the runoff time series of each hydrological station, the collected data is divided into a different sub-dataset, where the first 70% of the data is used for training and validating the model's parameter, while the other data is used for testing.

3.2. Performance evaluation indicators

To fully show the forecasting ability of the proposed method, several evaluation indexes are used to analyze the prediction level. Root mean squared error (RMSE) is chosen as the first evaluation index to measure the model's performance in high flow, which is defined as below:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \tilde{y}_i)^2} \quad (15)$$

where y_i and \tilde{y}_i are the i th target and predicted data; and n is the amount of data for comparison.

The second index is set as the mean absolute percentage error (MAPE) for measuring the proportional error of the developed model. MAPE is often sensitive to the prediction error of large-magnitude data but insensitive to that of small-magnitude

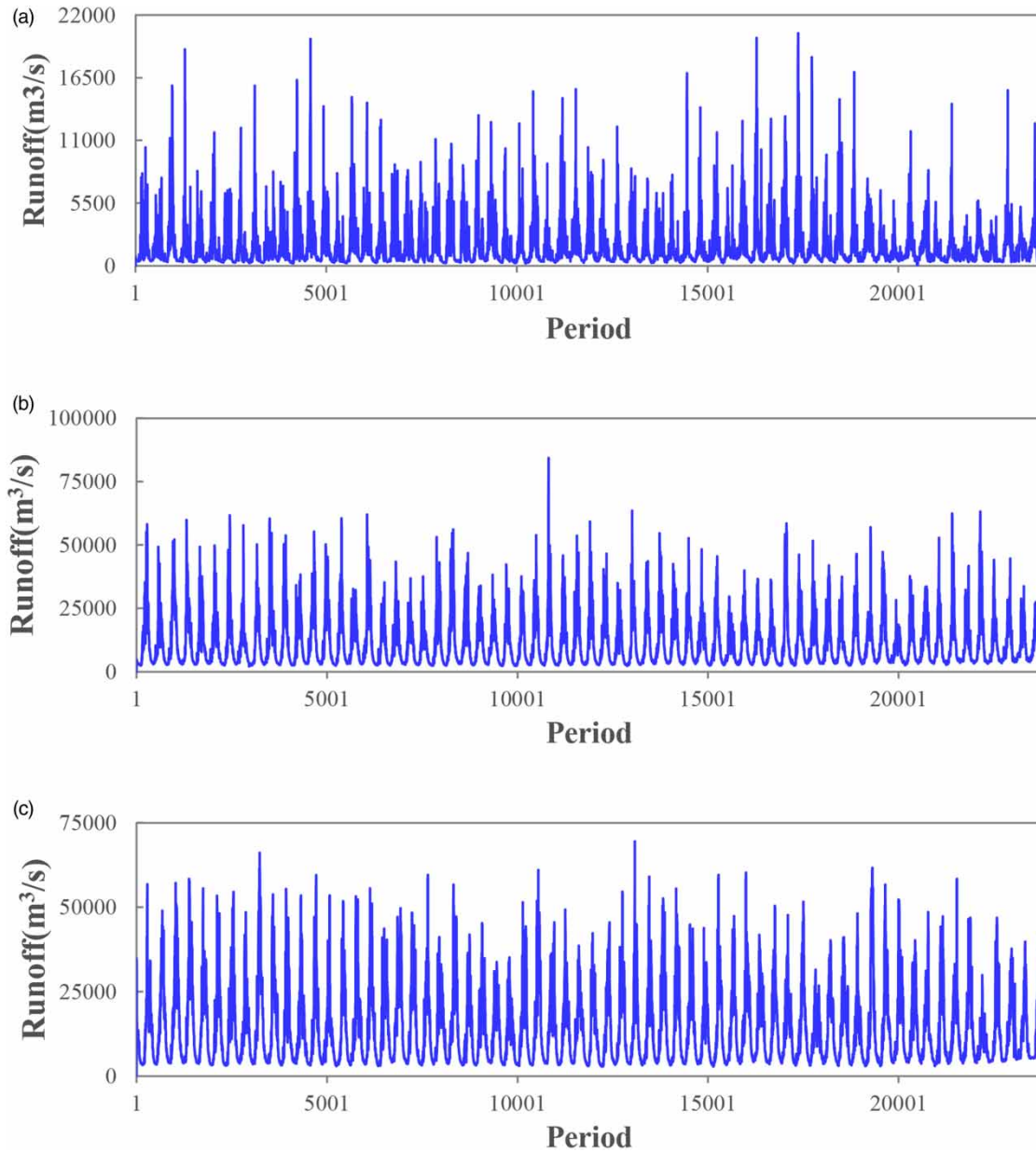


Figure 3 | Daily runoff series of the three hydrological stations: (a) station A, (b) station B, (c) station C.

data. The MAPE definition is given as:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\tilde{y}_i - y_i}{y_i} \right| \times 100\% \quad (16)$$

The third index is chosen as the coefficient of correlation (R) for reflecting the linear relationship between the target and predicted data, which is defined as below:

$$R = \frac{\sum_{i=1}^n [(y_i - y_{avg})(\tilde{y}_i - \tilde{y}_{avg})]}{\sqrt{\sum_{i=1}^n (y_i - y_{avg})^2 (\tilde{y}_i - \tilde{y}_{avg})^2}} \quad (17)$$

where y_{avg} and \tilde{y}_{avg} denote the mean value of all the target and predicted data.

The last index is set as the Nash–Sutcliffe efficiency (CE) for assessing the predictive capability of the developed model, which is defined as below:

$$CE = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - y_{\text{avg}})^2} \quad (18)$$

3.3. Comparison with the ELM method at station A

Table 1 lists multiple-step-ahead forecasting results of the ELM method and the hybrid method at both training and testing phases for station A. It can be clearly seen that at the same forecasting period, the proposed method obtains the best results among all the forecasting periods; with the increasing number of the forecasting period, the performances of all the forecasting methods become gradually worse while the results of the proposed method are still better than those of the ELM method. For instance, in the training phase, the proposed method makes about 15.70% and 16.62% improvements in the MAPE value for the one-step and two-step forecasting periods; as the forecasting period increases from 1 to 6, the RMSE value of the proposed method is increased from 577.774 to 1,105.495 at the testing phase, better than that of the ELM model from 656.59 to 1,146.196. Hence, it can be concluded that the organic combination of the SSA and ELM methods can effectively improve the forecasting results.

Figure 4 shows the multiple-step-ahead forecasting results of various methods for daily runoff of station A at the testing period. The two methods can effectively track the variation tendency of runoff series on the whole-time window, demonstrating the feasibility of the extreme learning machine method in streamflow prediction. On the other hand, the standard ELM method is not as good as the proposed method at the same forecasting period because its correlation coefficient values are obviously smaller than those of the hybrid method. In addition, with the increasing forecasting period, the correlation coefficient values of the proposed method are slowly reduced while the decreasing amplitudes of the ELM method in the correlation coefficient value are relatively large. Thus, the proposed method can make effective improvements on the robustness of the ELM method.

The relative forecasting errors of the two methods for daily streamflow at station A during the testing phase are drawn in Figure 5. As the forecasting period is equal to 1, the relative forecasting errors of the two methods are rather close; in the same forecasting period, the relative forecasting error of the ELM method varies in a relatively large zone in comparison with that of the proposed method; in addition, with the increasing forecasting period, the relative forecasting errors of the two methods gradually become larger. The reason lies in that for a long-term forecasting task, more uncertain factors are involved in the complex hydrological process and thereby it becomes much more difficult for the forecasting method to capture the dynamic change of runoff. Hence, the proposed method using SSA to optimize the parameters of the ELM model can produce satisfactory forecasting results.

Table 1 | Multi-step-ahead forecasting results of various forecasting models with different inputs at station A

Forecasting period	Model	Training				Testing			
		RMSE	MAPE	R	CE	RMSE	MAPE	R	CE
$\tau = 1$	ELM	828.001	21.113	0.890	0.792	656.590	19.337	0.887	0.786
	Proposed	677.504	17.798	0.928	0.860	577.774	17.027	0.913	0.834
$\tau = 2$	ELM	1,127.157	37.799	0.784	0.614	877.421	31.331	0.786	0.617
	Proposed	1,046.716	31.518	0.817	0.667	835.184	27.282	0.809	0.653
$\tau = 3$	ELM	1,295.671	52.177	0.700	0.490	997.539	42.134	0.712	0.505
	Proposed	1,232.407	40.925	0.734	0.538	952.819	34.205	0.742	0.549
$\tau = 4$	ELM	1,373.460	56.665	0.653	0.426	1,050.117	45.324	0.673	0.452
	Proposed	1,334.638	47.605	0.677	0.458	1,020.208	39.312	0.696	0.482
$\tau = 5$	ELM	1,432.996	62.165	0.613	0.376	1,095.431	49.764	0.636	0.403
	Proposed	1,391.922	51.821	0.641	0.411	1,063.061	42.662	0.664	0.438
$\tau = 6$	ELM	1,506.925	71.634	0.557	0.310	1,146.196	57.730	0.592	0.347
	Proposed	1,462.935	57.607	0.591	0.349	1,105.495	48.027	0.629	0.392

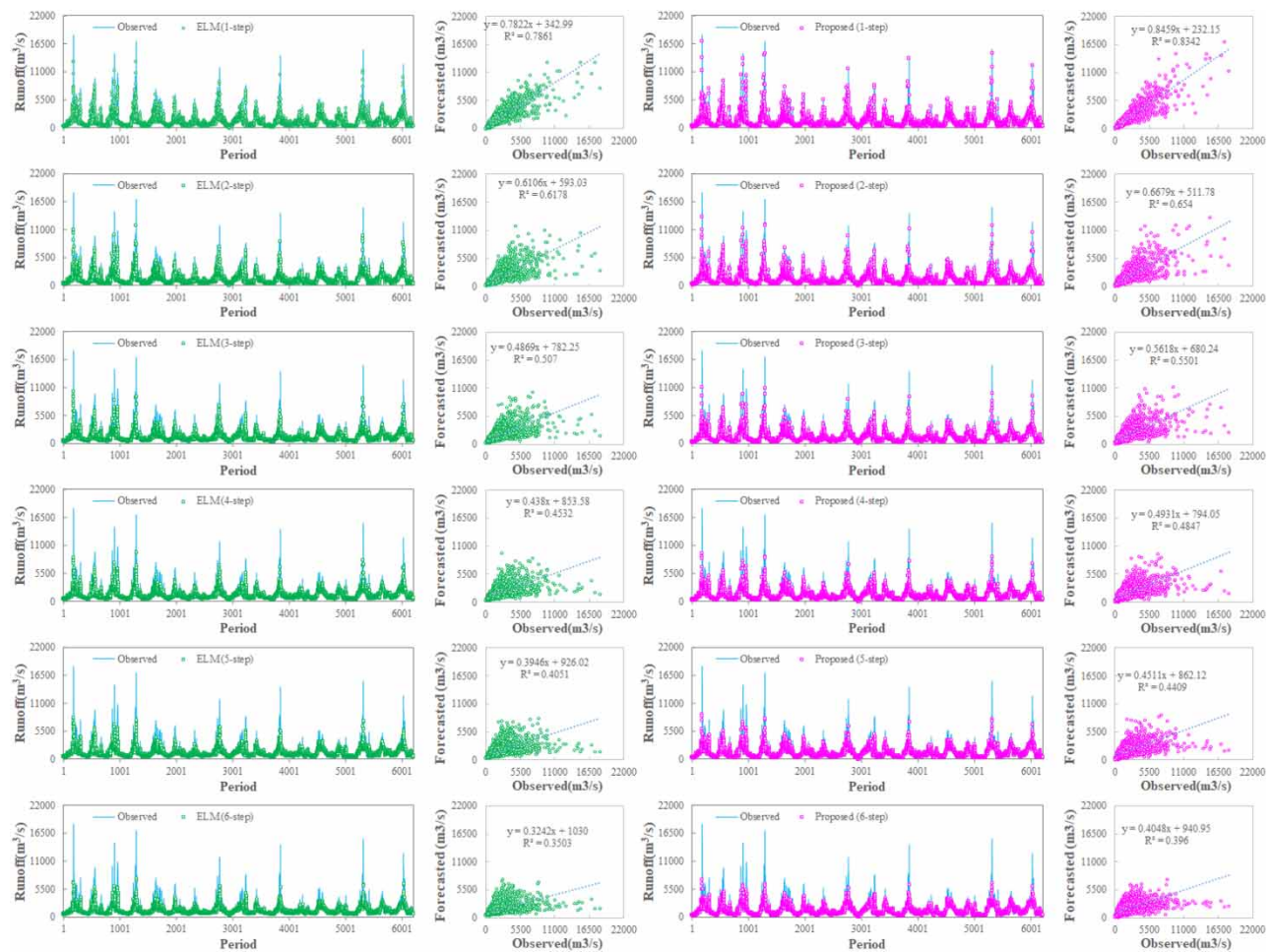


Figure 4 | Multi-step-ahead forecasting results of various forecasting models for station A at the testing phase.

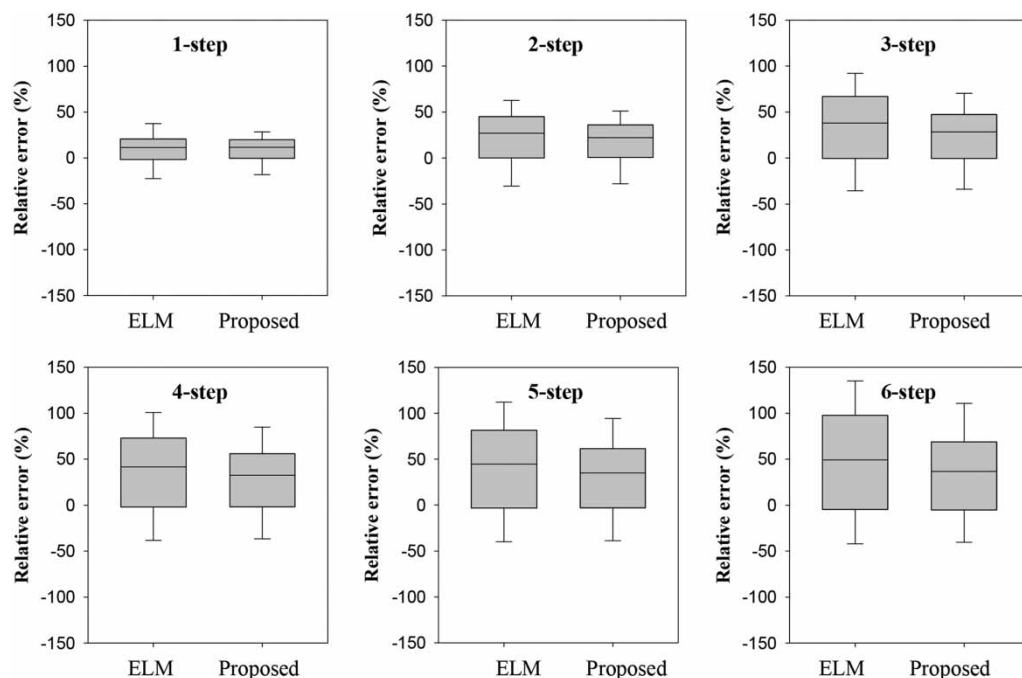


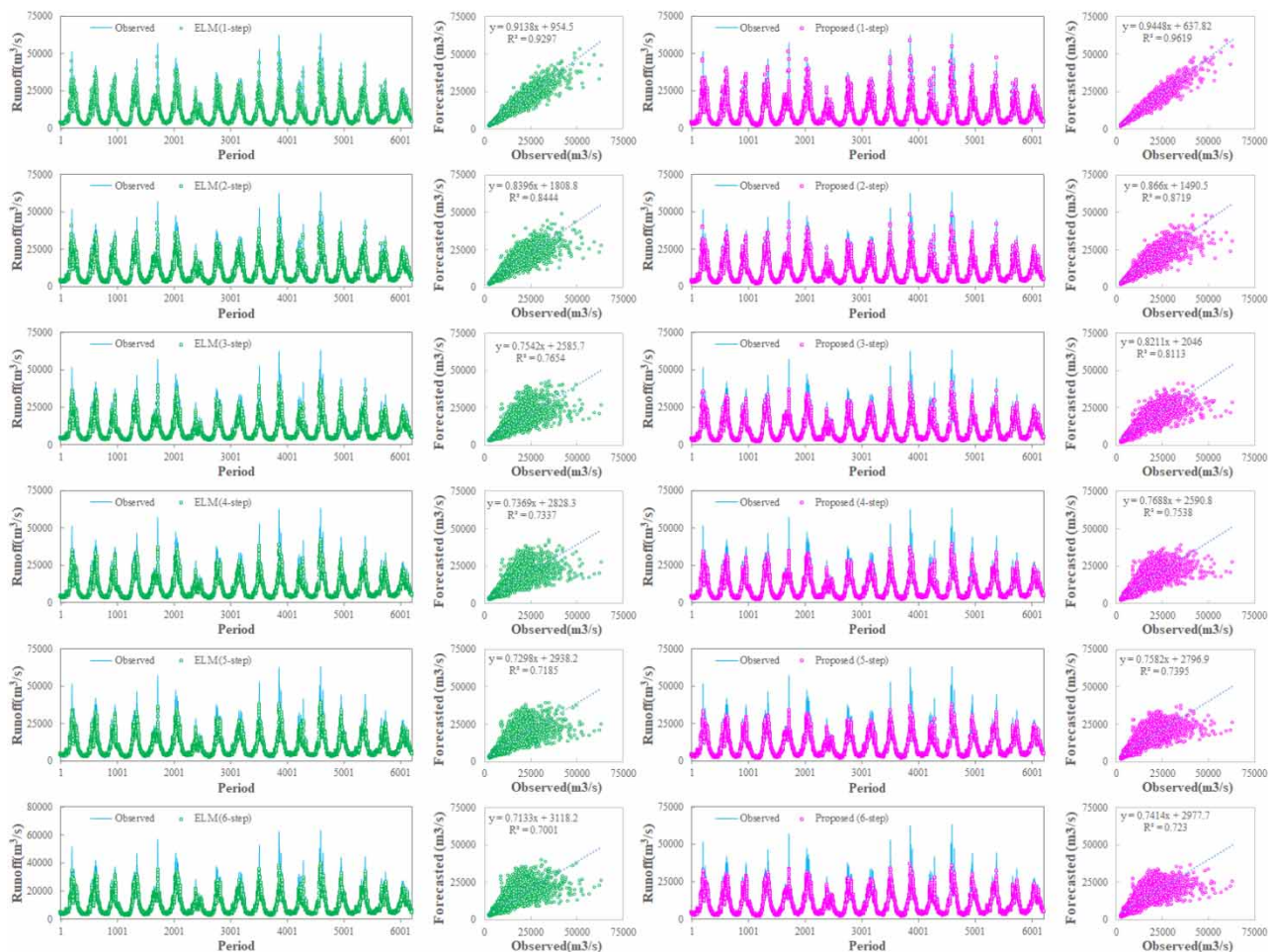
Figure 5 | Relative forecasting errors of the two methods for station A at the testing phase.

Table 2 | Statistical indexes of multi-step-ahead forecasting results of various forecasting models with different inputs at station B

Forecasting period	Model	Training				Testing			
		RMSE	MAPE	R	CE	RMSE	MAPE	R	CE
$\tau = 1$	ELM	2,524.510	9.206	0.964	0.929	2,156.296	8.104	0.964	0.929
	Proposed	1,848.003	6.521	0.981	0.962	1,591.868	6.081	0.981	0.962
$\tau = 2$	ELM	3,733.924	14.214	0.919	0.844	3,204.370	12.996	0.919	0.844
	Proposed	3,391.754	12.544	0.934	0.872	2,905.577	11.244	0.934	0.872
$\tau = 3$	ELM	4,475.360	24.004	0.881	0.777	3,931.205	19.754	0.875	0.765
	Proposed	4,078.582	15.834	0.903	0.814	3,530.244	14.615	0.901	0.811
$\tau = 4$	ELM	4,750.504	24.434	0.865	0.748	4,188.029	20.857	0.857	0.734
	Proposed	4,584.702	20.083	0.875	0.766	4,032.888	18.434	0.868	0.753
$\tau = 5$	ELM	4,868.630	24.747	0.858	0.736	4,308.272	21.549	0.848	0.718
	Proposed	4,722.937	21.117	0.867	0.751	4,154.468	19.970	0.860	0.738
$\tau = 6$	ELM	5,025.318	26.161	0.848	0.718	4,447.283	22.855	0.837	0.700
	Proposed	4,858.154	22.350	0.858	0.737	4,284.134	21.268	0.850	0.721

3.4. Comparison with the ELM method at station B

Table 2 lists the multiple-step-ahead forecasting results of the ELM method and the hybrid method at both training and testing phases for station B. It can be found that the hybrid method outperforms the ELM methods in terms of various

**Figure 6** | Multi-step-ahead forecasting results of various forecasting models for station B at the testing phase.

statistical indexes at both training and testing phases, proving that the SSA method can effectively improve the network's compactness. For instance, as the forecasting period is increased from 1 to 3, our method better the ELM method with approximately 26.18%, 9.32% and 10.20% improvements in the RMSE value at the testing phase. Thus, the hybrid method is an effective hydrological time series forecasting tool that can provide better performance than the standard ELM method.

Figure 6 shows the multiple-step-ahead forecasting results of various methods for daily runoff of station B in the testing period. For the two methods, the variation tendency of runoff time series is well simulated in the testing phase, while the standard ELM method is inferior to the proposed method due to its smaller correlation coefficient value. Therefore, it can be concluded that the generalization ability of the standard ELM method can be sharply improved by the SSA method.

Figure 7 shows the peak flows of the various methods for station B at the testing phase. It can be clearly seen that the proposed method can obtain better forecasting results than the control methods. For the first forecasting period, the hybrid method makes about 14.95% underestimation in the forecasting error of the peak flow, smaller than the 26.27% of ELM. Hence, our method can effectively yield reliable hydrological forecasting information to provide strong technical support for the decision-making process of the water resources system.

3.5. Comparison with different evolutionary algorithms at station C

Table 3 gives multiple-step-ahead forecasting results of the ELM method and the hybrid method at both training and testing phases for station C. Figure 8 shows the multiple-step-ahead forecasting results of the various methods for daily runoff of station C in the testing period. The hybrid method provides better results than the ELM methods with respect to various statistical indexes. Thus, the hybrid method combining the ELM and SSA method is an effective tool for hydrological time series forecasting.

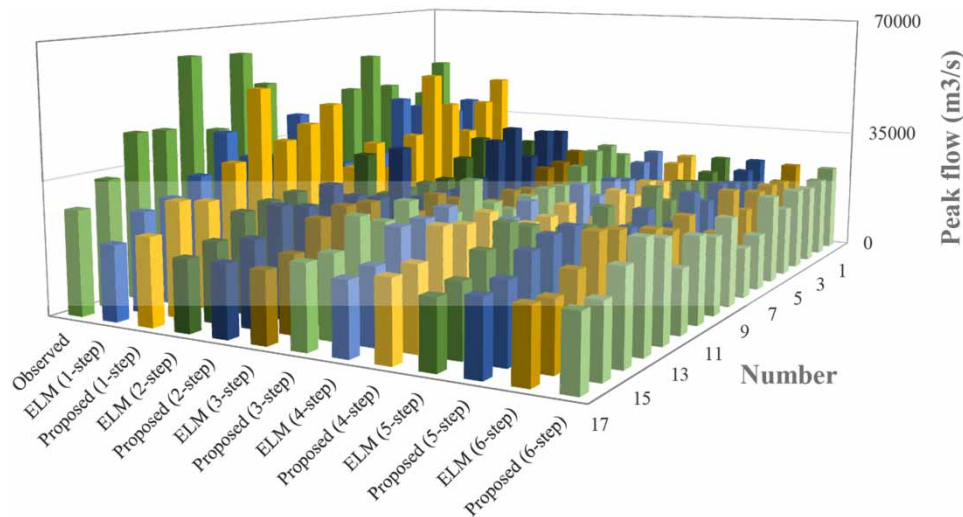


Figure 7 | Detailed results of various forecasting models for station B at the testing phase.

Table 3 | Statistical indexes of one-step-ahead forecasting results of various forecasting models with different inputs at station C

Method	Training				Testing			
	RMSE	MAPE	R	CE	RMSE	MAPE	R	CE
ELM	2,618.2055	8.8309	0.9718	0.9440	2,321.1800	8.7469	0.9737	0.9476
ELM-GA	1,575.7859	6.9297	0.9899	0.9797	1,457.6769	6.5555	0.9897	0.9793
ELM-DE	1,787.2459	8.0147	0.9871	0.9739	1,642.8204	7.5818	0.9870	0.9738
ELM-PSO	1,462.1992	5.7864	0.9913	0.9825	1,381.9364	5.7689	0.9908	0.9814
ELM-GSA	1,393.0435	5.5704	0.9922	0.9842	1,321.0378	5.5150	0.9916	0.9830
Proposed	1,332.3079	5.3211	0.9929	0.9855	1,296.9795	5.4292	0.9920	0.9836

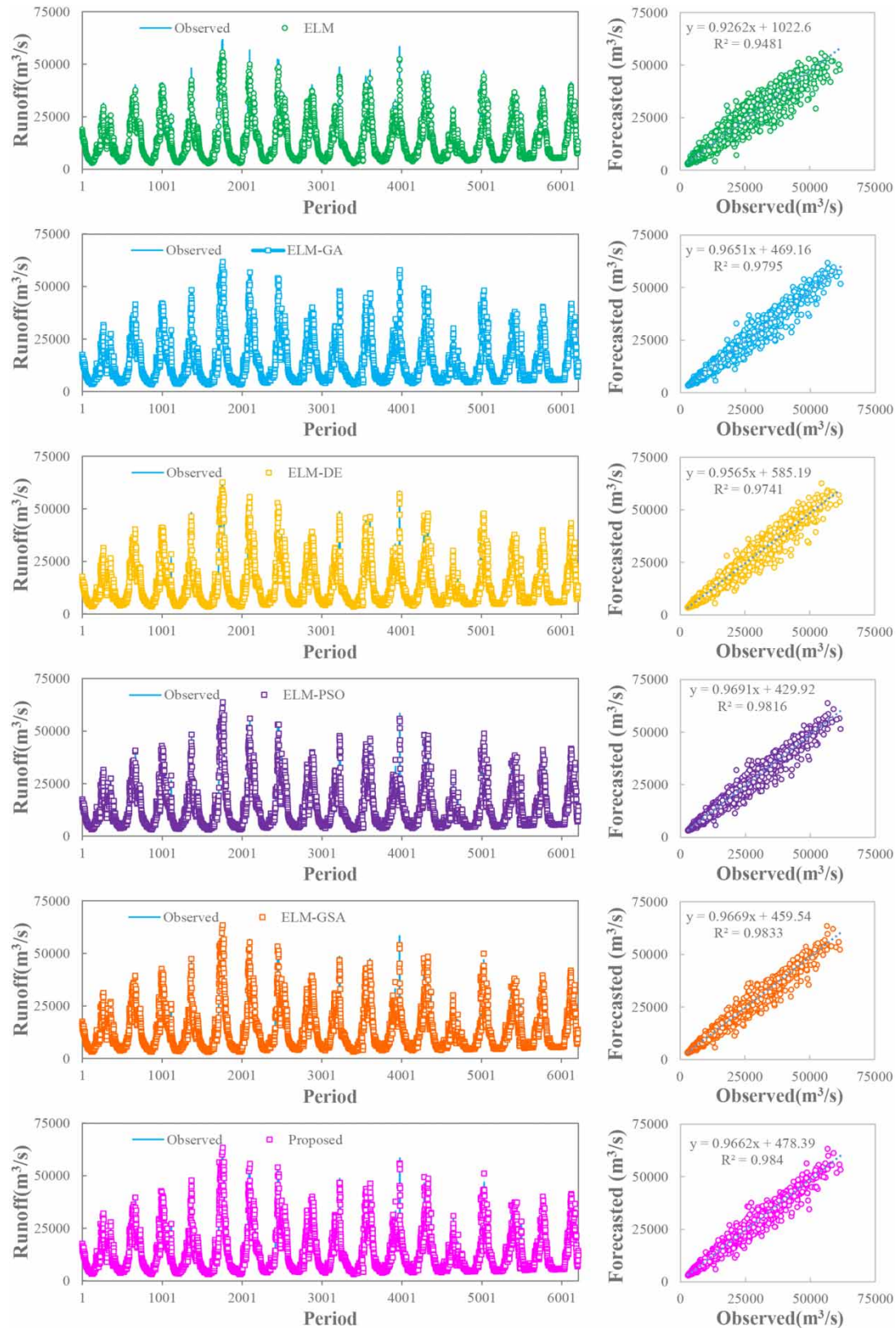


Figure 8 | One-step-ahead forecasting results of various forecasting models for station C at the testing phase.

4. CONCLUSION

In this paper, an effective evolutionary extreme learning machine based on the sparrow search algorithm is developed to accurately forecast hydrological time series. Specifically, the sparrow search algorithm is used to search for satisfying combinations of the input-hidden weights as well as hidden biases, while the Moore–Penrose generalized inverse method is chosen to analytically obtain the hidden-output weights. In this way, the developed method has a higher probability to find a better network structure than the standard ELM model without any parameter tuning. The developed method successfully forecasts the runoff time series of three hydrological stations in China. The experimental results show that the developed method is superior to several traditional methods with respect to various performance evaluation indexes. Thus, a novel and practical evolutionary extreme learning machine model using swarm intelligence is developed to handle the complex hydrological forecasting task.

On the other hand, the application of the proposed method may be limited in practice since it is rather difficult to determine the optimal parameters of the neural network model in theory while there are also big differences in the characteristic information of the hydrological elements at different places. Considering the rapid development of computer technology, the prospects of the proposed method can be improved by introducing more effective soft computing methods or developing more robust modified strategies. In addition, combinations of the newly developed signal processing technique can also be used to improve the performance of the proposed method.

ACKNOWLEDGEMENTS

This paper is supported by the National Key Research and Development Program of China (no. 2018YFC1508002) and Scientific Research Projects of China Three Gorges Group Co. (0799251).

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

REFERENCES

- Adib, A., Kisi, O., Khoramgah, S., Gafouri, H. R., Liaghat, A., Lotfird, M. & Moayyeri, N. 2021 [A new approach for suspended sediment load calculation based on generated flow discharge considering climate change](#). *Water Supply* **21**, 2400–2413.
- Birbal, P., Azamathulla, H., Leon, L., Kumar, V. & Hosein, J. 2021 [Predictive modelling of the stage–discharge relationship using Gene-Expression Programming](#). *Water Supply* **21**, 3503–3514.
- Cambria, E., Huang, G. B., Kasun, L. L. C., Zhou, H., Vong, C. M., Lin, J., Yin, J., Cai, Z., Liu, Q., Li, K., Leung, V. C. M., Feng, L., Ong, Y. S., Lim, M. H., Akusok, A., Lendasse, A., Corona, F., Nian, R., Miche, Y., Gastaldo, P., Zunino, R., Decherchi, S., Yang, X., Mao, K., Oh, B. S., Jeon, J., Toh, K. A., Teoh, A. B. J., Kim, J., Yu, H., Chen, Y. & Liu, J. 2013 [Extreme learning machines](#). *IEEE Intell. Syst.* **28** (6), 30–59.
- Cao, J., Lin, Z. & Huang, G. B. 2012 [Self-adaptive evolutionary extreme learning machine](#). *Neural Process Lett.* **36**, 285–305.
- Chen, Z., Fang, L. & Wang, H. 2019 [Internal incentives and operations strategies for the water-saving supply chain with cap-and-trade regulation](#). *Front. Eng.* **6**, 87–101.
- Chen, X. Y. & Chau, K. W. 2016 [A hybrid double feedforward neural network for suspended sediment load estimation](#). *Water Resour. Manag.* **30**, 2179–2194.
- Chen, S., Wei, Q., Zhu, Y., Ma, G., Han, X. & Wang, L. 2020 [Medium- and long-term runoff forecasting based on a random forest regression model](#). *Water Supply* **20**, 3658–3664.
- Chua, L. H. C. & Wong, T. S. W. 2011 [Runoff forecasting for an asphalt plane by Artificial Neural Networks and comparisons with kinematic wave and autoregressive moving average models](#). *J. Hydrol.* **397**, 191–201.
- Dalkılıç, H. Y. & Hashimi, S. A. 2020 [Prediction of daily streamflow using artificial neural networks \(ANNs\), wavelet neural networks \(WNNs\), and adaptive neuro-fuzzy inference system \(ANFIS\) models](#). *Water Supply* **20**, 1396–1408.
- Du, J. & Weng, F. 2021 [Construction management and technology innovation for main projects of Quanzhou Bay Bridge](#). *Front. Eng.* **8**, 151–155.
- Feng, Z. K. & Niu, W. J. 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.
- Feng, Z. K., Niu, W. J. & Liu, S. 2021a [Cooperation search algorithm: a novel metaheuristic evolutionary intelligence algorithm for numerical optimization and engineering optimization problems](#). *Appl. Soft Comput.* **98**, 106734.
- Feng, Z. K., Niu, W. J., Tang, Z. Y., Xu, Y. & Zhang, H. R. 2021b [Evolutionary artificial intelligence model via cooperation search algorithm and extreme learning machine for multiple scales nonstationary hydrological time series prediction](#). *J. Hydrol.* **595**, 126062.

- Ghasemlounia, R. & Sagheblian, S. M. 2021 [Uncertainty assessment of kernel based approaches on scour depth modeling in downstream of ski-jump bucket spillways](#). *Water Supply* **21**, 2333–2346.
- Guo, J., Zhou, J., Zou, Q., Liu, Y. & Song, L. 2013 [A novel multi-objective shuffled complex differential evolution algorithm with application to hydrological model parameter optimization](#). *Water Resour. Manag.* **27**, 2923–2946.
- 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**, 1571–1590.
- Huang, G. B., Zhu, Q. Y., Mao, K. Z., Siew, C. K., Saratchandran, P. & Sundararajan, N. 2006a [Can threshold networks be trained directly?](#) *IEEE Transactions on Circuits and Systems II: Express Briefs* **53**, 187–191.
- Huang, G. B., Zhu, Q. Y. & Siew, C. K. 2006b [Extreme learning machine: theory and applications](#). *Neurocomputing* **70**, 489–501.
- Huang, G., Song, S., Gupta, J. N. D. & Wu, C. 2014 [Semi-supervised and unsupervised extreme learning machines](#). *IEEE T. Cybernetics* **44**, 2405–2417.
- Huang, G., Huang, G. B., Song, S. & You, K. 2015 [Trends in extreme learning machines: a review](#). *Neural Networks* **61**, 32–48.
- Ji, Y., Dong, H., Xing, Z., Sun, M., Fu, Q. & Liu, D. 2021 [Application of the decomposition–prediction–reconstruction framework to medium- and long-term runoff forecasting](#). *Water Supply* **21**, 696–709.
- Jiang, Z., Li, R., Li, A. & Ji, C. 2018 [Runoff forecast uncertainty considered load adjustment model of cascade hydropower stations and its application](#). *Energy* **158**, 693–708.
- Kaluarachchi, Y. 2021 [Potential advantages in combining smart and green infrastructure over silo approaches for future cities](#). *Front. Eng.* **8**, 98–108.
- Leon, L. P., Chaplot, B. & Solomon, A. 2020 [Water consumption forecasting using soft computing – a case study, Trinidad and Tobago](#). *Water Supply* **20**, 3576–3584.
- Luo, J., Chen, H., Hu, Z., Huang, H., Wang, P., Wang, X., Lv, X. E. & Wen, C. 2019 [A new kernel extreme learning machine framework for somatization disorder diagnosis](#). *IEEE Access* **7**, 45512–45525.
- Mei, W., Liu, Z., Cheng, Y. & Zhang, Y. 2018 [A MDPSO-based constructive ELM approach with adjustable influence value](#). *IEEE Access* **6**, 60757–60768.
- Moghayedi, A. & Windapo, A. 2019 [Key uncertainty events impacting on the completion time of highway construction projects](#). *Front. Eng.* **6**, 275–298.
- Niu, W. J., Feng, Z. K., Li, S. S., Wu, H. J. & Wang, J. Y. 2021a [Short-term electricity load time series prediction by machine learning model via feature selection and parameter optimization using hybrid cooperation search algorithm](#). *Environ. Res. Lett.* **16**, 055032.
- Niu, W. J., Feng, Z. K. & Liu, S. 2021b [Multi-strategy gravitational search algorithm for constrained global optimization in coordinative operation of multiple hydropower reservoirs and solar photovoltaic power plants](#). *Appl. Soft. Comput.* **107**, 107315.
- Niu, W. J., Feng, Z. K., Liu, S., Chen, Y. B., Xu, Y. S. & Zhang, J. 2021c [Multiple hydropower reservoirs operation by hyperbolic grey wolf optimizer based on elitism selection and adaptive mutation](#). *Water Resour. Manag.* **35**, 573–591.
- Niu, W. J., Feng, Z. K., Xu, Y. S., Feng, B. F. & Min, Y. W. 2021d [Improving prediction accuracy of hydrologic time series by least-squares support vector machine using decomposition reconstruction and swarm intelligence](#). *J. Hydrol. Eng.* **26**, 04021030.
- Niu, W. J., Feng, Z. K., Jiang, Z. Q., Wang, S., Liu, S., Guo, W. & Song, Z. G. 2021e [Enhanced harmony search algorithm for sustainable ecological operation of cascade hydropower reservoirs in river ecosystem](#). *Environ. Res. Lett.* **16**, 055013.
- Niu, W. J., Feng, Z. K., Li, Y. R. & Liu, S. 2021f [Cooperation search algorithm for power generation production operation optimization of cascade hydropower reservoirs](#). *Water Resour. Manag.* **35**, 2465–2485.
- Peng, T., Zhou, J., Zhang, C. & Zheng, Y. 2017 [Multi-step ahead wind speed forecasting using a hybrid model based on two-stage decomposition technique and AdaBoost-extreme learning machine](#). *Energ. Convers. Manage.* **153**, 589–602.
- Peng, A., Zhang, X., Peng, Y., Xu, W. & You, F. 2019 [The application of ensemble precipitation forecasts to reservoir operation](#). *Water Supply* **19**, 588–595.
- Pérez Lespier, L., Long, S., Shoberg, T. & Corns, S. 2019 [A model for the evaluation of environmental impact indicators for a sustainable maritime transportation systems](#). *Front. Eng.* **6**, 368–383.
- Sun, L., Li, Z., Zhang, K. & Jiang, T. 2021 [Impacts of precipitation and topographic conditions on the model simulation in the north of China](#). *Water Supply* **21**, 1025–1035.
- Tamura, T., Miyajima, Y. & Nishiyama, K. 1990 [Expert system for water purification plant control](#). *Water Supply* **8**, 568–577.
- Tian, Z. 2020 [A combined prediction approach based on wavelet transform for crop water requirement](#). *Water Supply* **20**, 1016–1034.
- Truchet, C., Arbelaez, A., Richoux, F. & Codognot, P. 2016 [Estimating parallel runtimes for randomized algorithms in constraint solving](#). *J. Heuristics* **22**, 613–648.
- Tuerxun, W., Chang, X., Hongyu, G., Zhijie, J. & Huajian, Z. 2021 [Fault diagnosis of wind turbines based on a support vector machine optimized by the sparrow search algorithm](#). *IEEE Access* **9**, 69307–69315.
- Viccione, G., Guarnaccia, C., Mancini, S. & Quartieri, J. 2020 [On the use of ARIMA models for short-term water tank levels forecasting](#). *Water Supply* **20**, 787–799.
- Wang, H., Wu, X. & Gholinia, F. 2021 [Forecasting hydropower generation by GFDL-CM3 climate model and hybrid hydrological-Elman neural network model based on Improved Sparrow Search Algorithm \(ISSA\)](#). *Concurrency Computation* **33**, e6476.
- Won, J. & Kim, S. 2020 [Future drought analysis using SPI and EDDI to consider climate change in South Korea](#). *Water Supply* **20**, 3266–3280.

- Wu, Y., Wang, Q., Li, G. & Li, J. 2020 Data-driven runoff forecasting for Minjiang River: a case study. *Water Supply* **20**, 2284–2295.
- Yadav, B., Sudheer, C., Mathur, S. & Adamowski, J. 2016 Estimation of in-situ bioremediation system cost using a hybrid Extreme Learning Machine (ELM)–particle swarm optimization approach. *J. Hydrol.* **543**, 373–385.
- Yang, L., Li, Z., Wang, D., Miao, H. & Wang, Z. 2021 Software defects prediction based on hybrid particle swarm optimization and sparrow search algorithm. *IEEE Access* **9**, 60865–60879.
- Yuan, X., Chen, C., Lei, X., Yuan, Y. & Muhammad Adnan, R. 2018 Monthly runoff forecasting based on LSTM–ALO model. *Stoch. Env. Res. Risk A.* **32**, 2199–2212.
- Yuan, J., Zhao, Z., Liu, Y., He, B., Wang, L., Xie, B. & Gao, Y. 2021 DMPPT control of photovoltaic microgrid based on improved sparrow search algorithm. *IEEE Access* **9**, 16623–16629.
- Zhang, P. & Ariaratnam, S. T. 2021 Life cycle cost savings analysis on traditional drainage systems from low impact development strategies. *Front. Eng.* **8**, 86–97.
- Zhang, C. & Ding, S. 2021 A stochastic configuration network based on chaotic sparrow search algorithm. *Knowl. Based Syst.* **220**, 106924.
- Zhou, J., Peng, T., Zhang, C. & Sun, N. 2018 Data pre-analysis and ensemble of various artificial neural networks for monthly streamflow forecasting. *Water* **10**, 628.
- Zhou, J. & Wang, S. 2021 A carbon price prediction model based on the secondary decomposition algorithm and influencing factors. *Energies* **14**, 1328.

First received 9 October 2021; accepted in revised form 23 November 2021. Available online 7 December 2021