Water Resources Management (2021) 35:4695–4726
https://doi.org/10.1007/s11269-021-02920-5

An Ensemble Hybrid Forecasting Model for Annual Runoff Based on Sample Entropy, Secondary Decomposition, and Long Short-Term Memory Neural Network

Wen-chuan Wang[1] · Yu-jin Du[1] · Kwok-wing Chau[2] · Dong-mei Xu[1] · Chang-jun Liu[3] · Qiang Ma[3]

Received: 23 February 2021 / Accepted: 10 August 2021 / Published online: 5 October 2021

Abstract
Accurate and consistent annual runoff prediction in a region is a hot topic in management, optimization, and monitoring of water resources. A novel prediction model (ESMD-SE-WPD-LSTM) is presented in this study. Firstly, extreme-point symmetric mode decomposition (ESMD) is used to produce several intrinsic mode functions (IMF) and a residual (Res) by decomposing the original runoff series. Secondly, sample entropy (SE) method is employed to measure the complexity of each IMF. Thirdly, wavelet packet decomposition (WPD) is adopted to further decompose the IMF with the maximum SE into several appropriate components. Then long short-term memory (LSTM) model, a deep learning algorithm based recurrent approach, is employed to predict all components. Finally, forecasting results of all components are aggregated to generate the final prediction. The proposed model, which is applied to seven annual series from different areas in China, is evaluated based on four evaluation indexes (R, MAE, MAPE and RMSE). Results indicate that ESMD-SE-WPD-LSTM outperforms other benchmark models in terms of four evaluation indexes. Hence the proposed model can provide higher accuracy and consistency for annual runoff prediction, rendering it an efficient instrument for scientific management and planning of water resources.

Keywords Annual runoff prediction · Two-phase decomposition · Long short-term memory · Extreme-point symmetric mode decomposition · Wavelet packet decomposition · Sample entropy

1 Introduction

Long-term runoff forecasting is essential for optimal management of hydro-resources (Reddy et al. 2021), ecological restoration (Feng et al. 2020a), flood mitigation (He et al. 2020), power generation (Feng et al. 2020b), irrigation scheduling (Poul et al. 2019), etc. The problem has received extensive attention globally (Xiang et al. 2020). Numerous

✉ Wen-chuan Wang
wangwen1621@163.com; wangwenchuan@ncwu.edu.cn

Extended author information available on the last page of the article

models have been applied to the prediction accuracy of annual hydrologic time series (Al-Juboori 2021), which can be divided into two types (Chau et al. 2005): physical-based and data-driven models. Physical-based models require detailed multi-source information and powerful computational tools, while data-driven models are an efficient alternative by building direct relationships between input and output data, without involving the complex internal physical mechanisms. Recently, many data-driven models (such as LSTM, ANFIS (Adaptive Network-based Fuzzy Inference System), ANN (Artificial Neural Network), etc.) have been adopted in the field (Parisouj et al. 2020; Sahoo et al. 2019). Hence, this paper focuses on developing an appropriate data-driven model for annual runoff prediction.

In the last few decades, deep learning algorithm has been gradually employed for hydrological field with fruitful research results (Tao et al. 2017; Yen et al. 2019). Recurrent neural network (RNN) is capable of modeling complex temporal dynamics. Numerous improvement methods have been undertaken to overcome problems of vanishing gradients and gradient explosion of RNN. As a representative of them, LSTM has been used in signal recognition and forecast. LSTM model was employed to predict monthly water table depth in agricultural field (Zhang et al. 2018). Kratzert et al. (2018) investigated the potential of LSTM for daily streamflow prediction. Akbari Asanjan et al. (2018) developed a rainfall prediction method by extrapolating cloud-top brightness temperature utilizing LSTM. Yuan et al. (2018) examined the accuracy of a hybrid method for monthly runoff forecasting by integrating LSTM and ant lion optimizer algorithm. Saeed et al. (2020) proposed a method for wind speed prediction by using a bidirectional LSTM model and automatic encoder. Bai et al. (2021) proposed a two-level cascade LSTM (C-LSTM) model for daily runoff forecasting, and the C-LSTM yielded better performance than single LSTM. These studies proved the competitiveness of LSTM hydrological time series prediction.

Recently, many hybrid forecasting models for hydrological time series prediction have been developed, which include forecast modeling and data preprocessing. A decomposition algorithm can enhance the forecasting ability of a model by decomposing the raw hydrological series into more clean sub-series. The emergence of multi-resolution decomposition tools, namely singular spectrum analysis (SSA), ESMD, ensemble empirical mode decomposition (EEMD), CEEMDAN (complete EEMD with adaptive noise), WPD, wavelet transform (WT), and variational mode decomposition (VMD), have further stimulated researchers to make in-depth research on data preprocessing. Meng et al. (2021) proposed a hybrid VMD-SVM (support vector machine) coupling innovative input selection framework and stepwise decomposition sampling strategy for practical hydrological prediction. Bojang et al. (2020) examined the reliability of combining SSA with random forest (RF) and least-squares support vector regression (LS-SVR) for monthly precipitation prediction. However, SSA involved certain subjective factors in the process of noise reduction and was subject to the restriction of r matrix perturbation. Yuan et al. (2021) coupled two methods, namely group by month (GM) and EEMD, with LSTM to enhance the forecasting accuracy of daily runoff. Discrete wavelet transform (DWT) was capable of helping forecasting models to extract useful information (Tayyab et al. 2019), yet it might suffer from signal loss. Zuo et al. (2020) proposed a single-model forecasting (SF) framework, termed SF-VMD-LSTM, to forecast daily streamflow. However, the drawback of VMD is that the optimal parameter combination needs to be artificially set in advance. EEMD lacks accurate mathematical theory. To overcome their weaknesses, termed ESMD proposed by Wang and Li (2013), was adopted to attain more linear signal. The main idea of ESMD is to identify large-scale cycle and nonlinear trend of the data using internal extreme-point symmetry interpolation according to the characteristics of data itself. ESMD method replaces traditional integral transformation with direct interpolation, and the residual is optimized by a

least square approach. Therefore, ESMD is capable of reflecting time-varying characteristics of frequency and amplitude of each component. ESMD has been successfully used in broad fields(Zhou et al. 2019), few attempts have tended to the latest advance of ESMD for hydrological prediction. Therefore, this paper is to explore the efficiency of ESMD in capturing hydrological time series characteristics.

WPD is another data decomposition technology that has gained numerous attentions. WPD, an improvement of DWT, decomposes the approximation value same as details of signals in each level of decomposition. Whilst DWT only decomposes the approximation coefficient, WPD has the capability of splitting both detail coefficient and approximation coefficient simultaneously; thereby WPD provides more possibilities for hydrological time series. Seo et al. (2016) combined three models, including SVM, ANFIS, and ANN, with WPD for daily river stage prediction. Sun et al. (2020) coupled WPD and FS (feature selection) with ELM (extreme learning machine) to predict multi-step wind speed. Although WPD has achieved fruitful results in many fields, it is of great significance to fill the research gap in mid- and long-term runoff forecasting.

Despite the above fruitful results, it should be noted out that single decomposition method might be hard to fully mitigate signal nonlinearity. To attain more linear series and higher forecasting accuracy, Liu et al. (2018) proposed a wind speed multistep prediction model by combining VMD, SSA, LSTM, and ELM. Sun and Huang (2020) combined secondary decomposition (SD) with sequence reconstruction to predict air pollutant concentration, and the model had excellent prediction performance. In summary, SD method can provide more linear signal and solve the limitation of single decomposition method to a certain extent. Therefore, this paper uses a secondary decomposition framework (ESMD-SE-WPD) to attain more linear series. Then, LSTM model is adopted for annual runoff forecasting. Finally, forecasted results of all sub-series are summed to generate the final prediction. The performance of the developed model is then compared with several benchmarking prediction models (LSTM, ANFIS, ANN, ESMD-LSTM, ESMD-SE-SSA-LSTM, ESMD-SE-CEEMDAN-LSTM).

The contribution of this paper includes three parts. First, an ESMD-SE-WPD-LSTM hybrid model, which can provide reasonable forecasting accuracy for annual runoff forecasting practice, is proposed. Second, an data preprocessing method based on secondary decomposition technology, which can provide more linear sub-series than single decomposition method, is proposed. Finally, three machine learning models are investigated with applications on seven basins, and LSTM model is combined with four preprocessing methods for examining the performance of data preprocessing technologies.

The remainder of the paper is arranged as follows: Sect. 2 is literature review. Section 3 depicts data source and evaluation indexes. Section 4 introduces the empirical forecasting experiments and discussion. Finally, Sect. 5 summarizes the paper.

2 Methodology

2.1 ESMD

ESMD, proposed by Wang and Li (2013), is a new adaptive data processing method and can be applied to analyze non-stationary and nonlinear signal. ESMD uses internal extreme-point symmetry interpolation, instead of external envelope interpolation, and optimizes the residual mode using least square approach, which overcomes shortcomings of

modal aliasing and screening termination in EMD. Detailed steps of ESMD are shown in (Sun et al. 2018).

2.2 Sample Entropy

Sample entropy (SE), proposed by Alcaraz and Rieta (2010), is a novel approach to describe the complexity of series. The computation steps are as follows:

(1) Recombine $X = (x(1),x(2), \ldots ,x(n))$ into a matrix:

$$\ddot{X} = \begin{bmatrix} x(1), x(2), \cdots, x(n-m+1) \\ x(2), x(3), \cdots, x(n-m+2) \\ \vdots \\ x(m), x(m+1), \cdots, x(n) \end{bmatrix} \tag{1}$$

(2) The distance between vector $x(j)$ and $x(i)$ can be defined as $d[x(i), x(j)]$:

$$d\left[x(i), x(j)\right] = \max(|x(i+l) - x(j+l)|), (1 \le l \le m-1; 1 \le i \ne j \le n-m+1) \tag{2}$$

where $l = 0, 1, 2, \cdots m-1$;

(3) For $x(i)$, r denotes the threshold. Compute their number meeting the threshold $d[x(i), x(j)] \le r$ as B_i. Then, compute ratio $B_i^m(r)$:

$$B_i^{\mathrm{m}}(r) = \frac{B_i}{n-m+1} \tag{3}$$

(4) Compute the average value $B^m(r)$ of $B_i^m(r)$:

$$B^{\mathrm{m}}(r) = \frac{1}{n-m} \sum_{i=1}^{n-m} B_i^m(r) \tag{4}$$

(5) Increase m by 1 and repeat steps 1 to 3, then compute $B^{m+1}(r)$:

$$B^{\mathrm{m}+1}(r) = \frac{1}{n-m} \sum_{i-1}^{n-m} B_i^{m+1}(r) \tag{5}$$

(6) SE is defined as follows:

$$SE(m, r) = \lim_{n \to \infty} \left\{ -\ln \left[\frac{B^m(r)}{B^{m+1}(r)} \right] \right\} \tag{6}$$

2.3 WPD

WPD is identical to wavelet decomposition, except that the former extends the abilities of the latter (Alickovic et al. 2018). The three-layer binary trees of WPD are illustrated in Fig. 1. WPD splits the signal into approximation coefficients and detail coefficients by a mother wavelet function. The decomposition levels and mother wavelet function have a deep influence on the performance of WPD. WPD includes DWT and CWT (continuous wavelet transform). CWT is as follows:

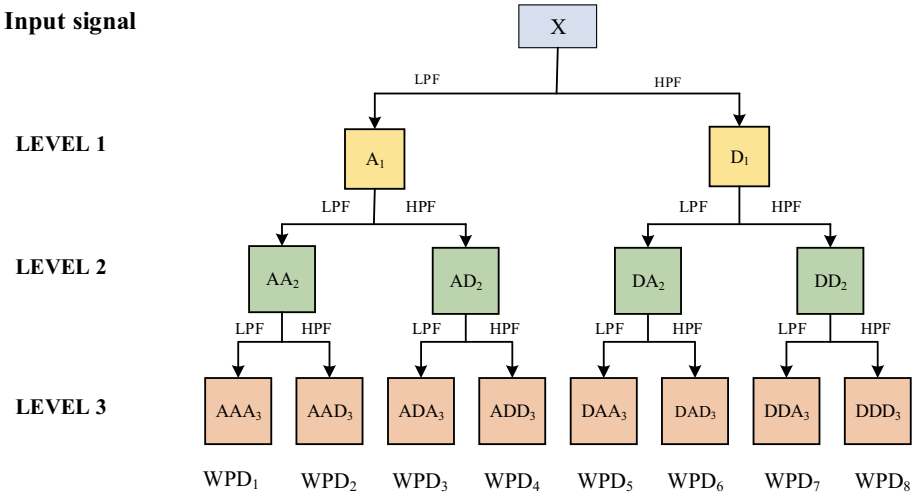

Fig. 1 Sketch map of WPD method

$$CWT_x(a,b) = \left\langle x(t), \psi_{a,b}(t) \right\rangle = \int x(t)\psi^*((t-b)/a)/\sqrt{a}dt \tag{7}$$

where $x(t)$ is input, $*$ complex conjugate, b translation parameter, a scale parameter, and $\psi(t)$ mother wavelet function. a and b in DWT are:

$$\begin{cases} a = 2^i \\ b = j2^i \end{cases} \tag{8}$$

where i and j denotes the scale and translation parameters, respectively.

2.4 LSTM

LSTM model is capable of solving the dependency problems of short-term and long-term time series. The memory cell of LSTM is a critical parameter, which contributes to memorize the temporal state. Each memory cell encompasses three gates, namely input, output, and forget gates. These gates perform as filters in playing different roles, solving exploding and vanishing gradient problems of RNN. The framework of LSTM is shown in Fig. 2.

The implementation of cell state update and computation of LSTM output are:

$$f_t = \sigma(W_{fx} \cdot x_t + W_{fh} \cdot h_{t-1} + b_f) \tag{9}$$

$$i_t = \sigma(W_{ix} \cdot x_t + W_{ih} \cdot h_{t-1} + b_i) \tag{10}$$

$$\overline{c}_t = \tanh(W_{cx} \cdot x_t + W_{ch} \cdot h_{t-1} + b_c) \tag{11}$$

$$c_t = f_t * c_{t-1} + i_t * \overline{c}_t \tag{12}$$

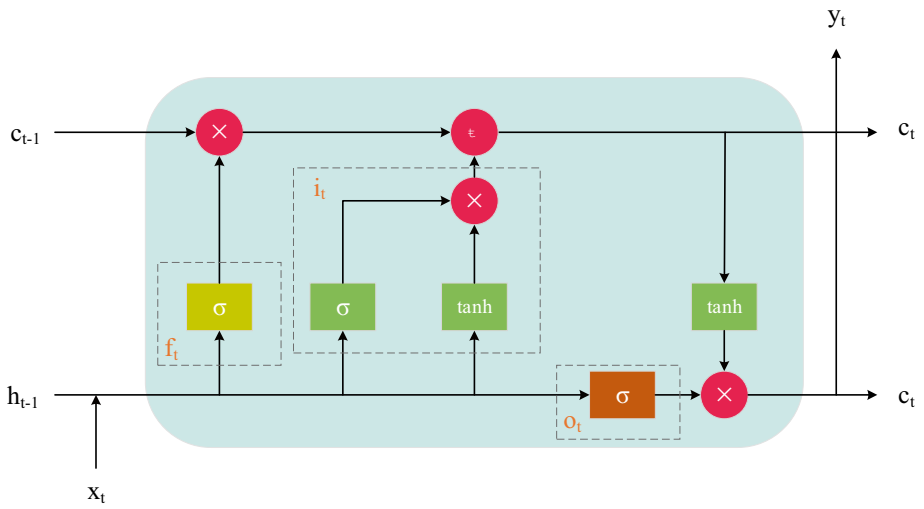

Fig. 2 The basic LSTM architecture

$$o_t = \sigma(W_{ox} \cdot x_t + W_{oh} \cdot h_{t-1} + b_o) \tag{13}$$

$$h_t = o_t * \tanh(c_t) \tag{14}$$

$$y_t = \sigma(W_{yx} \cdot h_t + b_y) \tag{15}$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{16}$$

where $x(t)$ is input, $y(t)$ output, f_t forget gate, o_t output gate, i_t input gate, c_t cell state at time t, W weight matrix, b bias vector, $\sigma(x)$ nonlinear activation function, and h_t activation vectors for each memory block.

2.5 Model Construction

The basic framework of the runoff forecasting system proposed in this paper is shown in Fig. 3. The modeling processes are as follows:

Step 1: ESMD. ESMD is adopted to split the observed runoff series into several IMFs and a Res.

Step 2: Sample entropy. Compute SE of each subsequence obtained in the previous step.

Step 3: Two-phase decomposition. ESMD-SE-WPD is adopted to attain more linear subseries.

Step 4: Normalize all data between [0, 1] by:

$$x_i' = \frac{x_i - \min\limits_{1 \le i \le n} \{x_i\}}{\max\limits_{1 \le i \le n} \{x_i\} - \min\limits_{1 \le i \le n} \{x_i\}} \tag{17}$$

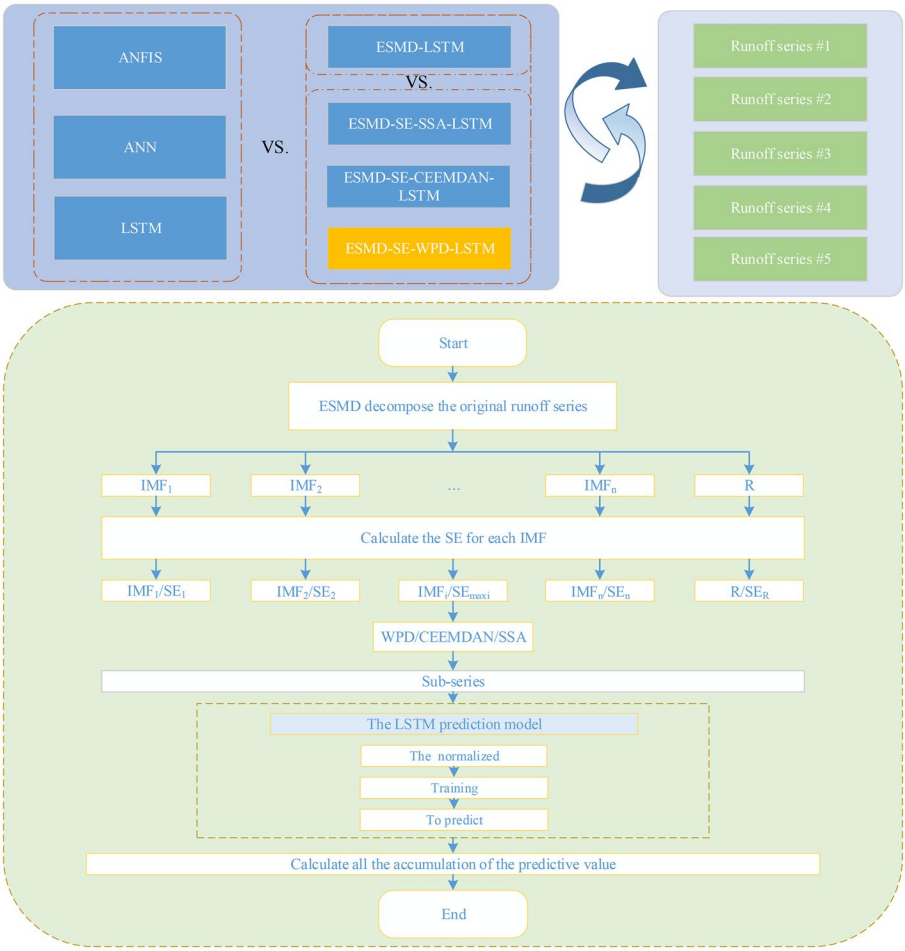

Fig. 3 Framework of the proposed model

Step 5: Select input variables. PACF (partial autocorrelation function) and precipitation knowledge are used to screen the number of input variables.

Step 6: Training and prediction. All model components are input to LSTM for training and prediction.

3 Data Description and Evaluation Indicators

The reliability of data is an important factor affecting the accuracy of mid- and long-term runoff prediction. The data in this paper are from seven areas in China, namely Mopanshan reservoir, Dahuofang reservoir, Biliuhe reservoir, Changshui hydrological station, Hongjiadu reservoir, Jiayuguan station and Yingluoxia station. Mopanshan Reservoir is located in Heilongjiang Province, Northeast China. The water source area of the reservoir is 1151 km^2, the average annual precipitation is about 750 mm, and the average annual runoff

is 5.60 billion m^3. Dahuofang reservoir is located in Fushun City, Northeast China, with a watershed area of 5437 km^2, annual average discharge of 52.3 m^3/s and annual average precipitation of 812 mm. Biliuhe reservoir is located in Liaoning Province of China. The drainage area is 2085 km^2, and the average annual precipitation is 742.8 mm. Hongjiadu hydropower is located on the main stream of Wujiang River in northwest of Guizhou Province, China, with a drainage area of 9900 km^2 and an average annual runoff of 4.89 billion m^3. Changshui hydrological station is located in Henan Province, China. It is a national basic hydrological station with a drainage area of 874 km^2, annual average rainfall of 530 mm and annual average runoff of 8.17 billion m^3. TaoLai River is a tributary of Heihe River system in China's inland river basin. Jiayuguan hydrological station is a national first-class streamflow control station for monitoring changes in TaoLai river regime, with a catchment area of 7095 km^2 and average runoff of 6.36 billion m^3. Yingluoxia hydrological station is a control station and boundary section between upper and middle reaches of Heihe River, with a catchment area of 10,009 km^2 and average annual discharge of 51.5 m^3/s. The observed annual data for seven stations are shown in Fig. 4. Their statistical descriptions are listed in Table 1, where data for Mopanshan hydropower, Dahuofang hydropower, Biliuhe hydropower Changshui station, Hongjiadu hydropower, Jiayuguan station and Yingluoxia station, run from 1952–2004, 1953–2008, 1951–2007, 1961–2016, 1951–2005, 1956–2009 and 1956–2009, respectively. These data variations of seven stations are quite different, implying the high modeling difficulty of these regions. For these seven stations, approximately 90% of the data are used for training and the remainder are used for testing.

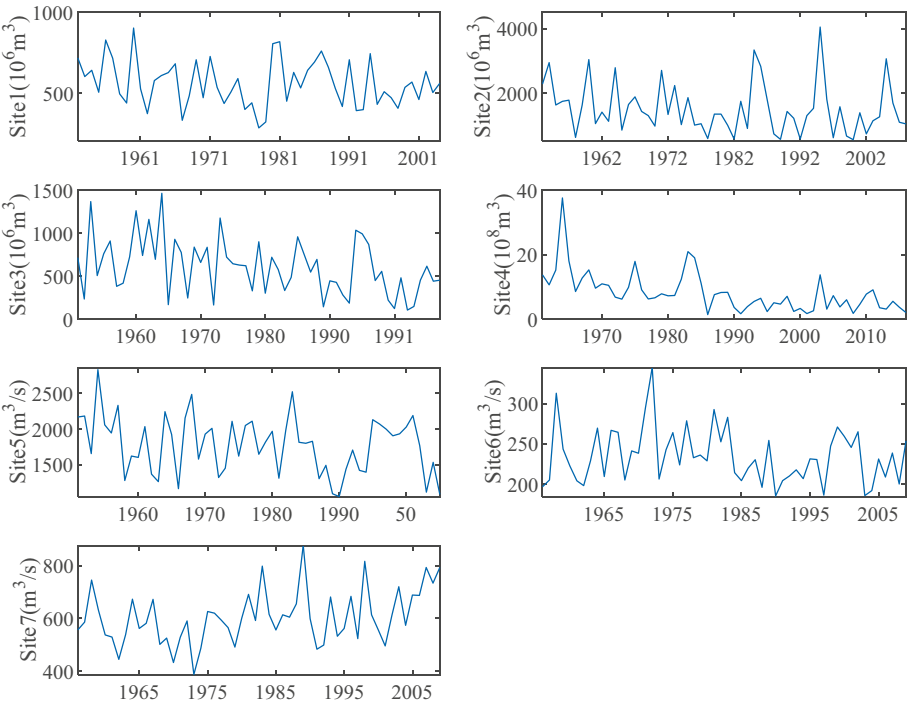

Fig. 4 Original runoff series

Table 1 Statistical description of runoff series at seven stations

Data	Station	Size	Max	Min	Mean	Std	Skewness
Site1	Mopanshan	53	903.00	281.00	558.70	143.40	7.50
Site2	Dahuofang	56	4043.00	536.30	1501.00	793.90	-7.70
Site3	Biliuhe	57	1466.00	100.50	607.90	328.00	-7.76
Site4	Changshui	56	37.62	1.32	8.17	6.25	-7.70
Site5	Hongjiadu	55	2831.00	1052.00	1778.00	402.10	-7.63
Site6	Jiayuguan	54	344.70	184.80	234.50	34.20	-7.56
Site7	Yingluoxia	54	874.70	387.20	604.60	102.20	-7.56

Where Std is the standard deviation

Results of the models are evaluated based on four evaluation indicators. These indexes include coefficient of correlation (R), mean absolute percentage error (MAPE), mean absolute error (MAE), and root mean square errors (RMSE). Their equations are as follows:

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_e(i)-y_o(i))^2} \tag{18}$$

$$MAPE=\frac{1}{N}\sum_{i=1}^{n}\left|\frac{y_e(i)-y_o(i)}{y_e(i)}\right|\times 100 \tag{19}$$

$$MAE=\frac{1}{n}\sum_{i=1}^{n}\left|y_e(i)-y_o(i)\right| \tag{20}$$

$$R=\frac{\sum_{i=1}^{n}(y_o(i)-\overline{y_o})(y_e(i)-\overline{y_e})}{\sqrt{\sum_{i=1}^{n}(y_o(i)-\overline{y_o})^2\sum_{i=1}^{n}(y_e(i)-\overline{y_e})^2}} \tag{21}$$

where $y_e(i)$, $y_o(i)$, $\overline{y_e}$ and $\overline{y_o}$ are estimated, observed, mean estimated, and mean observed precipitation values, respectively.

4 Case Studies

4.1 Series Decomposition

The first stage of the runoff prediction framework is to decompose the observed data using ESMD. Before decomposing the runoff series, the best screening number should be determined by repeating tests and comparisons. In this paper, the number of iterations is 100, and the numbers of remaining extreme points of the seven runoff datasets are 5, 6, 7, 5, 7, 5 and 5, respectively. The results at Site 1 after the decomposition are shown in Figs. 5, 6 and 7, whilst the decomposition results of other sites are not presented here. In Fig. 7, A1 and

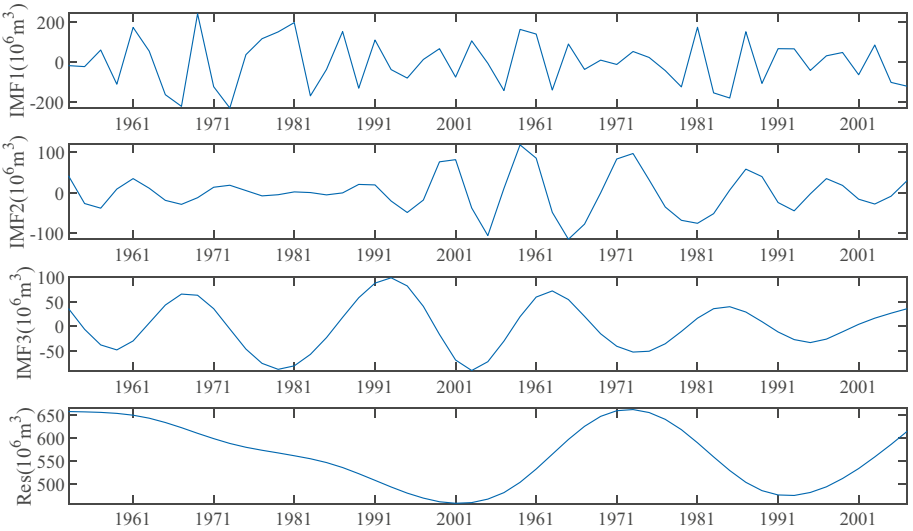

Fig. 5 Decomposition at Mopanshan hydropower by ESMD

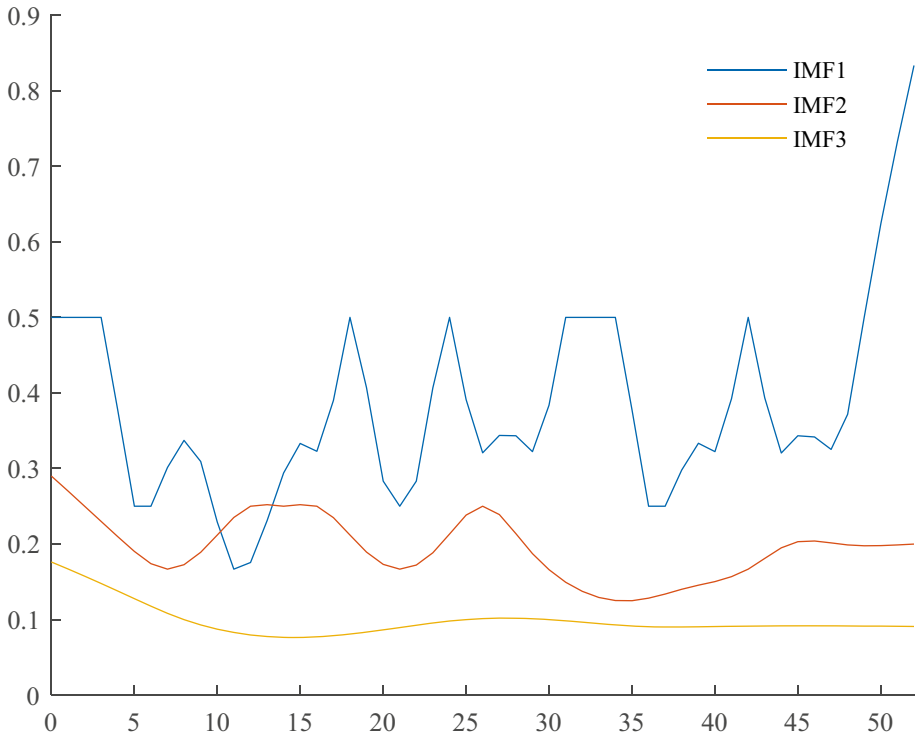

Fig. 6 Frequency distribution of each IMF at Site 1

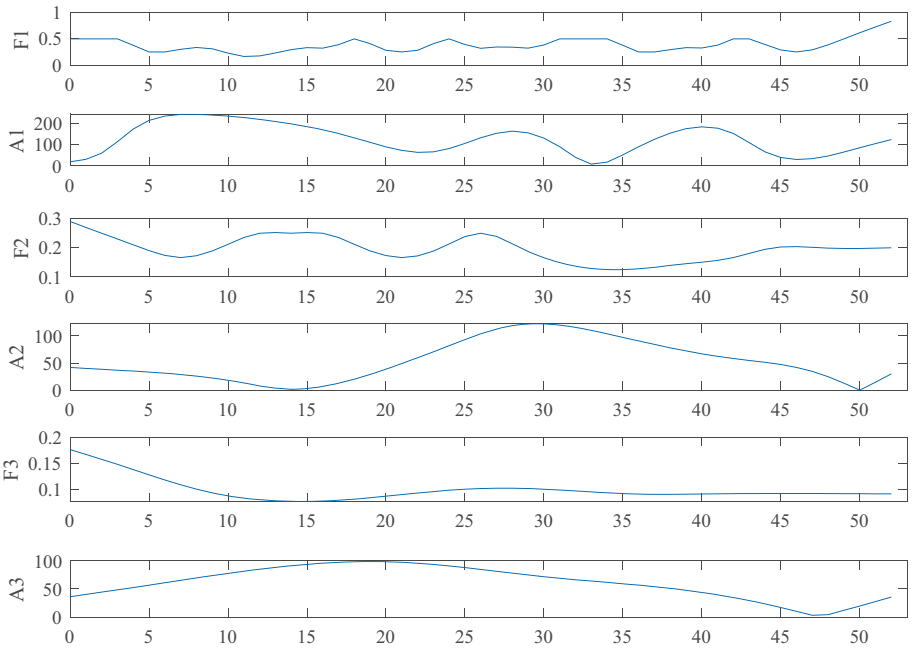

Fig. 7 Amplitude of each IMF at Site 1

F1denote the amplitude and frequency of IMF1, A2 and F2 the amplitude and frequency of IMF2, and A3 and F3 the amplitude and frequency of IMF3, respectively. It can be seen from Fig. 5 that each IMF split by ESMD is independent, the fluctuation of sub-series from IMF1 to Res decreases steadily and the stability becomes stronger gradually. Therefore, IMFs are steadier than original data and more conducive to capture signal features and predict non-linear sequences.

4.2 Sample Entropy Computation and Two-Phase Decomposition

SE of each subsequence obtained in the previous step is computed. Three decomposition methods are then adopted to further decompose IMF with the maximum SE. As shown in Fig. 8, SE of all sub-series present a similar trend, and it can be noted that SE of IMF1 in each dataset is higher than that of other subseries, which means that it is more difficult to predict IMF1. To mitigate the complexity of IMF1, three decomposition algorithms, namely WPD, CEEMDAN and SSA, are used to decompose IMF1.

The selection of an appropriate wavelet basis function is very important to WPD. Symlet wavelet is an improved approximate symmetric wavelet function based on Daubechies wavelet, which can avoid signal distortion during decomposition and reconstruction(Yin et al. 2019). Therefore, a three-scale and fourth order Symlet wavelet is adopted as the wavelet basis function of WPD.

SSA is a traditional and powerful non-parametric decomposition algorithm for signal identification and analysis, which can capture noise component, trend and periodic from input signal(Dong et al. 2017). SSA can decompose a time series into some decipherable

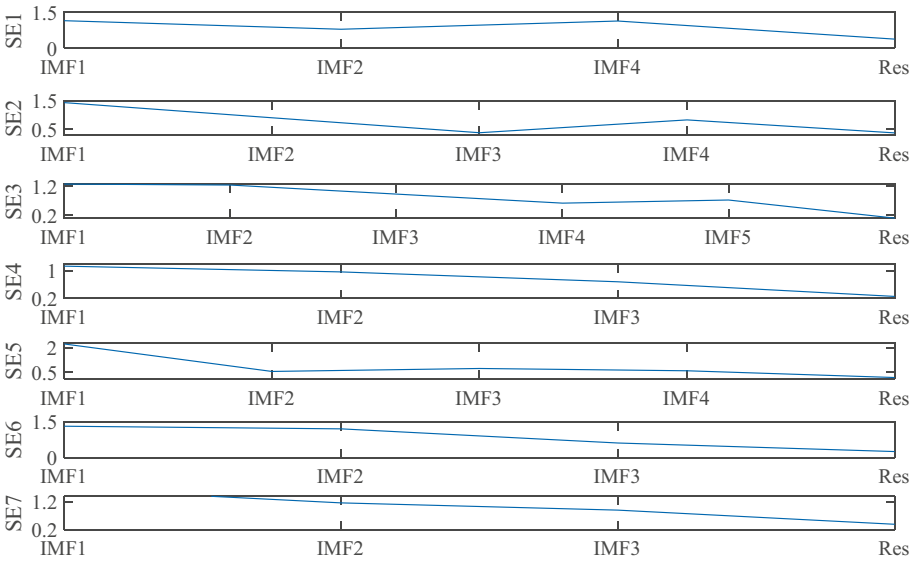

Fig. 8 SE of each sequence decomposed by ESMD

and simpler components, and contains two steps, namely decomposition and reconstruction. Decomposition incorporates singular value decomposition (SVD) and embedding whilst reconstruction comprises diagonal averaging and grouping. In SSA, window size (L) and eigenvalue grouping (EVG) are key parameters. Before decomposing IMF1, the best L and EVG value should be determined by repeating tests and comparisons. In this paper, L and EVG are set to 12 and 6, respectively.

CEEMDAN proposed by Colominas et al. (2012), is an enhancement on EEMD and can attain better separation and accurately reconstruct the raw signal. CEEMDAN obtains the modes by adding white Gaussian noise and computing a unique residue to reduce EEMD deficiency. The method overcomes the mode mixing problem since the procedure of CEEMDAN in decomposition and reconstruction are complete. In CEEMDAN application, too many modes may cause extra computational costs and complex training process. Hence, five IMFs and a residual are reconstructed in this paper.

All IMF1 are then split by WPD, SSA and CEEMDAN. The re-decomposition results of IMF1 at Site 1 are shown in Fig. 9, where y is the reconstructed series by SSA, whilst results at other sites are not presented here, and IMF1 with the maximum SE is decomposed into 16 subsequences with more regular fluctuation by three methods.

4.3 Number of Input Variables

The determination of input variables is an important procedure for prediction results. In this paper, two methods are utilized to select input combinations: (a) trial-and-error method; (b) PACF statistical approach. We conduct twelve ANN models with different input combinations. Table 2 lists input variables for Site 1 with respect to PACF and trial-and-error method, while input variables and PACF value for other sites are not shown here.

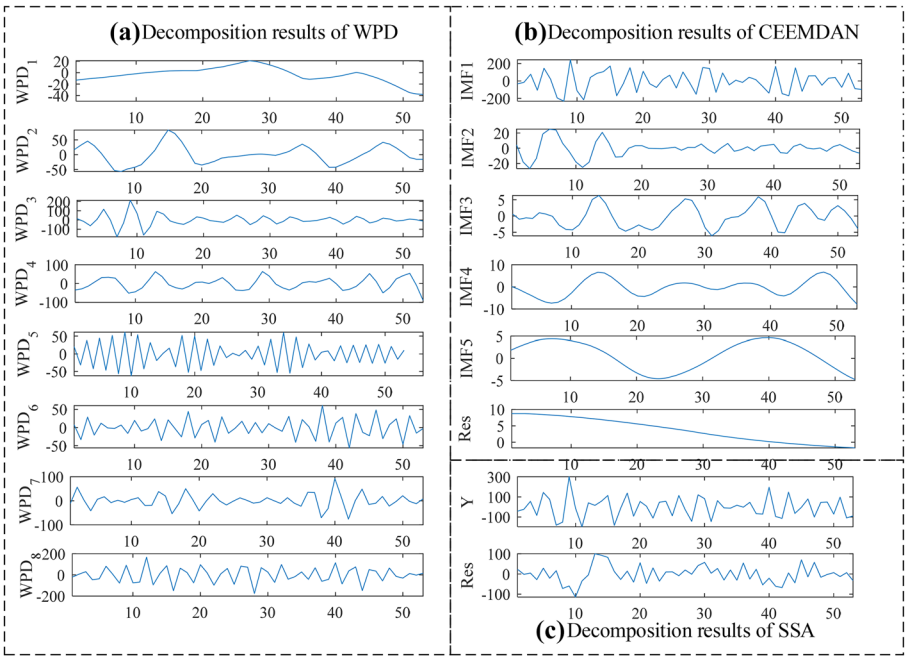

Fig. 9 Decomposition results of IMF1 at Site 1

4.4 Model Development

To verify the proposed model, seven models, namely LSTM, ANN, ANFIS, ESMD-LSTM, ESMD-SE-WPD-LSTM, ESMD-SE-SSA-LSTM, ESMD-SE-CEEMDAN-LSTM, are employed as benchmark for comparison. Detailed information relating to these models are presented in the following section.

(1) ANN

A standard three-layer feed forward ANN is adopted for annual runoff prediction. The numbers of input and output layer nodes are equal to that of input variables and one, respectively. Levenberg–Marquardt (LM) method, sigmoid function and purelin formula are adopted as the training, transfer and output functions, respectively. The best number of hidden nodes is determined as eight by trial-and-error method, and the training epochs are 500.

(2) ANFIS

Three methods, namely genfis1, genfis2 and genfis3, are available to initialize the data structure of ANFIS. Of three methods, genfis3 provides the most robust results in terms of generalization and stability in runoff modeling, and hence is employed throughout the processes. The specific parameter settings are shown in Table 3.

(3) LSTM

The selection of hyper-parameters is a difficult task for LSTM model construction. Adaptive moment estimation is to optimize the parameters. The model structure of LSTM, i.e., hyper-parameters and the number of hidden units are determined by trial-

Table 2 Input variables for Site 1

Series		LSTM	ANFIS	ANN
	Original	$r_{(t-1)} \sim r_{(t-3)}$	$r_{(t-1)} \sim r_{(t-6)}$	$r_{(t-1)} \sim r_{(t-3)}$
ESMD	IMF1	$r_{(t-1)} \sim r_{(t-10)}$	/	/
	IMF2	$r_{(t-1)} \sim r_{(t-4)}$	/	/
	IMF3	$r_{(t-1)} \sim r_{(t-6)}$	/	/
	R	$r_{(t-1)} \sim r_{(t-6)}$	/	/
WPD	WPD_1	$r_{(t-1)} \sim r_{(t-3)}$	/	/
	WPD_2	$r_{(t-1)} \sim r_{(t-10)}$	/	/
	WPD_3	$r_{(t-1)} \sim r_{(t-7)}$	/	/
	WPD_4	$r_{(t-1)} \sim r_{(t-3)}$	/	/
	WPD_5	$r_{(t-1)} \sim r_{(t-6)}$	/	/
	WPD_6	$r_{(t-1)} \sim r_{(t-6)}$	/	/
	WPD_7	$r_{(t-1)} \sim r_{(t-4)}$	/	/
	WPD_8	$r_{(t-1)} \sim r_{(t-5)}$	/	/
SSA	Y	$r_{(t-1)} \sim r_{(t-4)}$	/	/
	Res	$r_{(t-1)} \sim r_{(t-7)}$	/	/
CEEMDAN	IMF1	$r_{(t-1)} \sim r_{(t-10)}$	/	/
	IMF2	$r_{(t-1)} \sim r_{(t-6)}$	/	/
	IMF3	$r_{(t-1)} \sim r_{(t-6)}$	/	/
	IMF4	$r_{(t-1)} \sim r_{(t-6)}$	/	/
	IMF5	$r_{(t-1)} \sim r_{(t-6)}$	/	/
	Res	$r_{(t-1)} \sim r_{(t-6)}$	/	/

Where $r_{(t)}$ represents the estimated value of runoff and $r_{(t-p)}$ is the runoff at time t-p

and-error method. The number of hidden units is 200. The maximum number of epochs is 2000. The mini-batch size in each training iteration is 100. The initial learning rate is set to 0.01. Other parameters are determined to default values used by adaptive moment estimation. RMSE is adopted as the loss function.

(4) ESMD-LSTM model

For ESMD-LSTM model, the runoff datasets are first decomposed into a certain number of sub-series by ESMD. Each component is then modeled using LSTM, and the input variables for each composition are shown in Table 2.

(5) Two-phase decomposition methods combined with LSTM

Table 3 Parameters of ANFIS

Parameter	Value	Parameter	Value
Number of clusters	10	Maximum epochs	100
Partition matrix index	2	Expected error	0
Maximum iteration	500	Initial Step Size	default
Minimum improvement	1.00E-05	Step decrease rate	default
		Step increase rate	default

For ESMD-SE-WPD-LSTM, ESMD-SE-SSA-LSTM and ESMD-SE-CEEMDAN-LSTM, ESMD is adopted to decompose the raw data into several IMFs and a Res. Then SE method is employed to measure the complexity of each composition. Three decomposition algorithms, namely WPD, SSA and CEEMDAN, are adopted to further decompose the IMF with the maximum SE. Then, LSTM model is employed to predict each subseries obtained in the previous step.

4.5 Results and Discussion

Forecasting results of different methods for seven runoff datasets are presented in this section. Tables 4, 5, 6, 7, 8, 9 and 10 show error estimation results of different methods for seven runoff time series. Table 11 presents forecasting results of ESMD-LSTM of different sites. Figures 10, 11, 12, 13, 14, 15 and 16 present forecasting results of seven stations. The following should be noted before analyzing the results. The forecasting results of the testing phase play a greater role than those of the training phase. It is because the training period is utilized to train the model, and its performance is measured by data related to modeling. Since the testing dataset does not participate in modeling, its performance can truly reflect the model application efficiency.

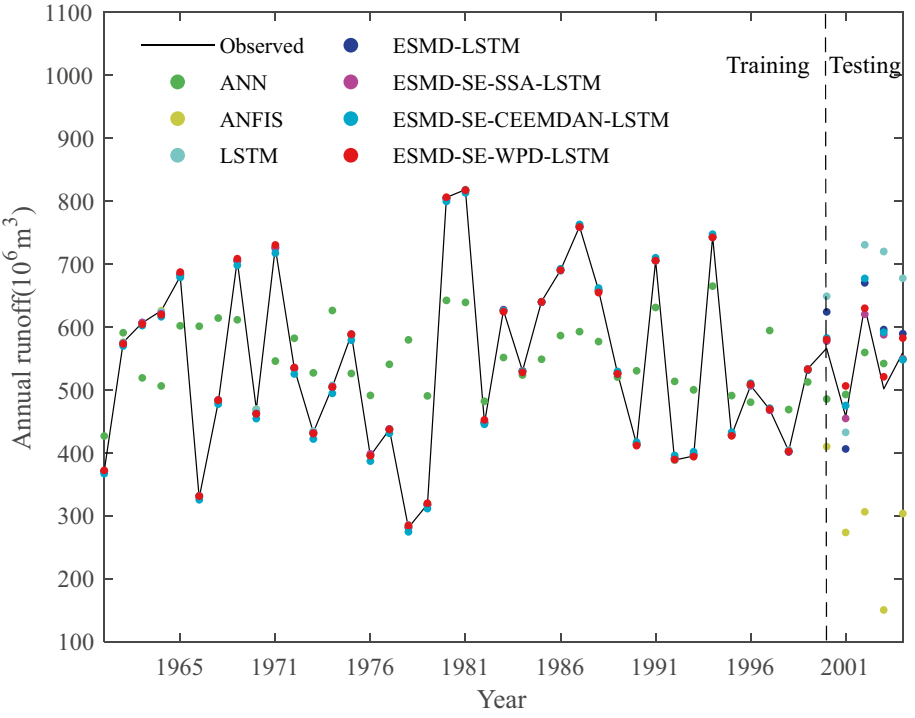

Fig. 10 Forecasting results of Mopanshan Hydropower

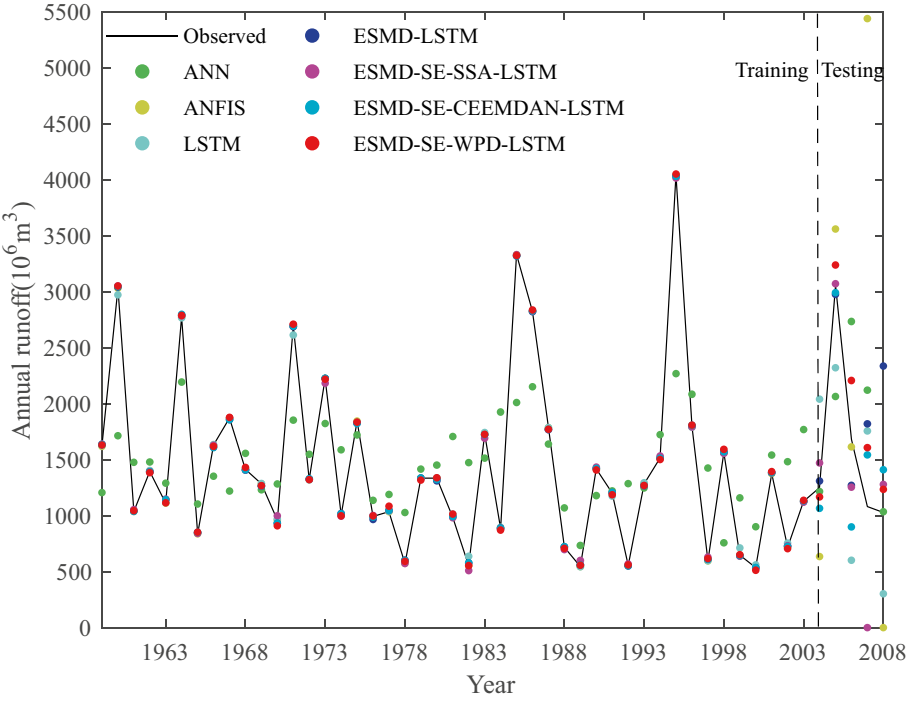

Fig. 11 Forecasting results of Dahuofang Hydropower

4.5.1 Experiment 1: Comparison of Several Single Prediction Models

In this section, we analyze prediction results of three single models at seven sites. As seen from Tables 4, 5, 6, 7, 8, 9 and 10, in the testing period, LSTM obtain the best average R in seven sites, the lowest average MAE, RMSE, MAPE in Hongjiadu station and Yingluoxia station, and similar MAE, RMSE, and MAPE with ANN in Biliuhe and Changshui station. Meanwhile, the differences between the best and worst value of four evaluation indicators of ANN and ANFIS are significantly higher than those of LSTM model. Besides, results of LSTM and ANFIS in the training period are clearly better than those in the testing period. ANN is in the lower level during the training period but can provide middle level result during the testing period.

Overall, LSTM model can provide optimal results for seven datasets in terms of four evaluation indexes. This analysis also demonstrates that there is still room to improve the forecasting accuracy of LSTM.

4.5.2 Experiment 2: Comparison of LSTM and ESMD-LSTM

This section compared the performance between single LSTM model and ESMD-LSTM hybrid model. Taking Site 1 as an example, ESMD-LSTM significantly improves the forecasting accuracy of LSTM model. In the testing period, ESMD-LSTM model improves LSTM model with 50.02%, 53.31% and 48.58% reduction in average MAE, RMSE and MAPE, respectively, and the improvement of prediction accuracy regarding average R is

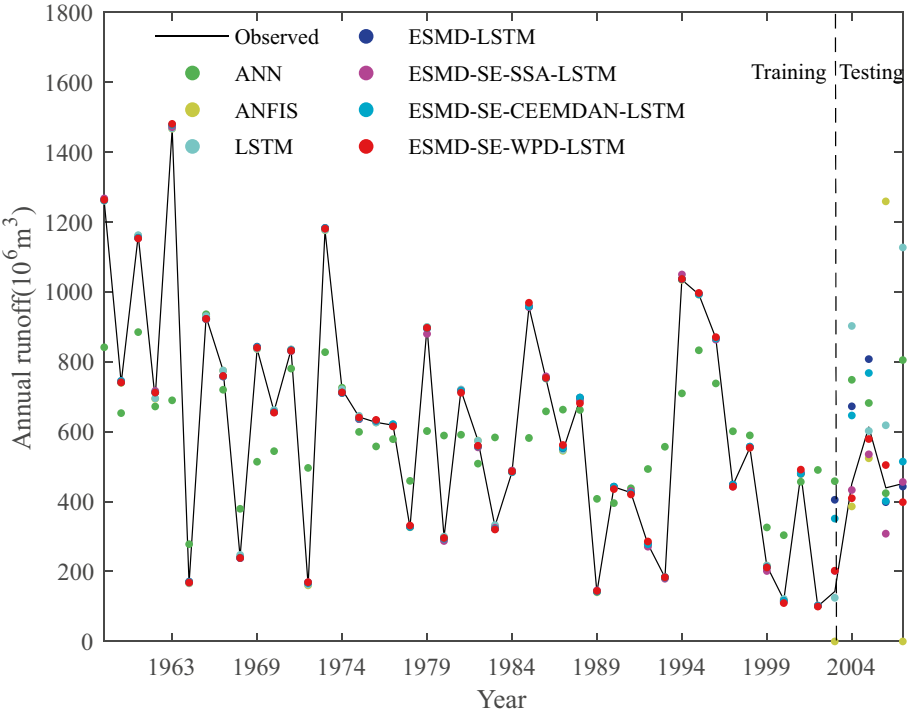

Fig. 12 Forecasting results of Biliuhe Hydropower

21.24%. According to results in Tables 5, 6, 7, 8, 9 and 10, ESMD-LSTM model is able to provide better results than LSTM model with substantial improvement in terms of four evaluation indexes. Table 11 lists the prediction results of different subseries obtained by ESMD-LSTM for seven datasets, the forecasting results of IMF1 are inferior to those of the other subseries. For IMF1, in the testing phase, the average R value of Sites 1–7 are 0.841, 0.654, 0.795, 0.573, 0.841, and 0.828, respectively, with an average of 0.735. However, in the testing period, the average R value of IMF2, IMF3 and Res are 0.904, 0.972 and 0.996, respectively. These results demonstrate that there is still room for improvement in the prediction performance of IMF1.

In general, this analysis illustrates that ESMD is suitable to decompose the annual runoff series and can improve forecasting accuracy. In addition, it can be confirmed from Table 11 that results of IMF1 are inferior to those of other subsequences. These analyses also illustrate that single decomposition method may have difficulty in fully capturing the frequency characteristics of the original data. The secondary decomposition method is then attempted to attain more linear sub-series and overcome the limitation of the single decomposition method to a certain extent.

4.5.3 Experiment 3: Comparison of Several Re-decomposition Hybrid Models

In this section, composite methods including ESMD-LSTM, ESMD-SE-WPD-LSTM, ESMD-SE-SSA-LSTM, ESMD-SE-CEEMDAN-LSTM, and ESMD-LSTM are treated

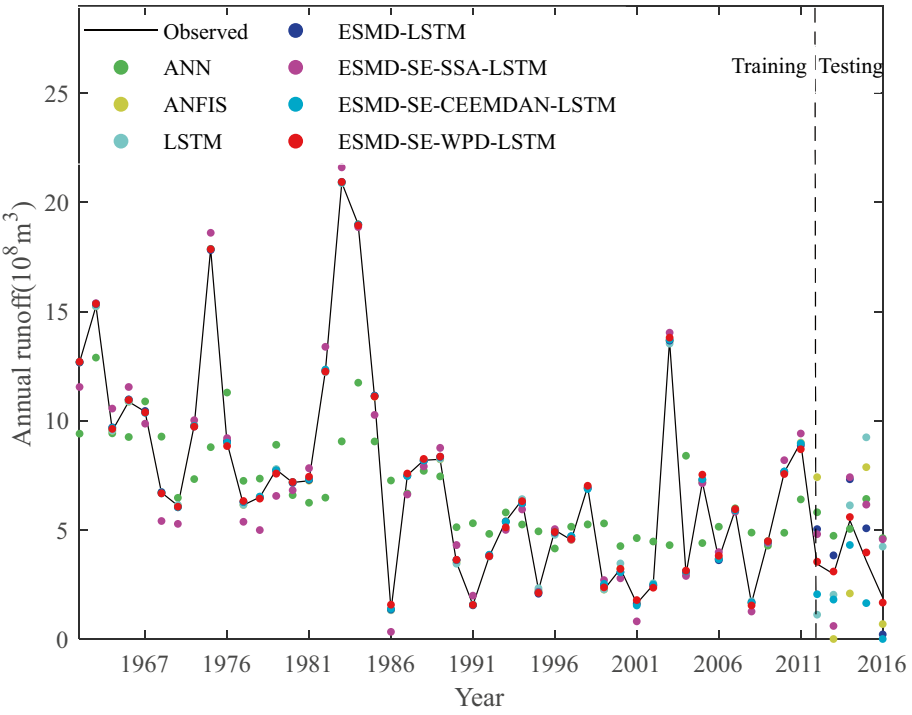

Fig. 13 Forecasting results of Changshui station

as the benchmark methods. Tables 4, 5, 6, 7, 8, 9 and 10 list the composite results for Sites 1–7. When forecasting annual runoff in seven stations, ESMD-SE-WPD-LSTM model exhibits the best results in terms of all evaluation indexes. Table 5 lists the forecasting results of Dahuofang reservoir in the testing period and average R values of ESMD-LSTM, ESMD-SE-SSA-LSTM, ESMD-SE-CEEMDAN-LSTM, and ESMD-SE-WPD-LSTM are 0.620, 0.838, 0.821, and 0.954, respectively. For Dahuofang reservoir in the testing period, compared with average R value of ESMD-LSTM model, ESMD-SE-SSA-LSTM, ESMD-SE-CEEMDAN-LSTM and ESMD-SE-WPD-LSTM yield improvements of 32.25%, 32.48% and 93.55%, respectively. Compared with average RMSE value of ESMD-LSTM, the three two-phase decomposition prediction methods yield reductions of 8.48%, 35.61% and 48.90% in the testing period, respectively. Compared to ESMD-LSTM, the three two-phase decomposition prediction methods exhibit average MAE value reductions of 14.02%, 27.86% and 41.42% in the testing period, respectively. Compared to ESMD-LSTM, the three two-phase decomposition prediction methods exhibit average MAPE value reductions of 16.41%, 37.03% and 50.52% in the testing period, respectively. To verify the performance of the presented model, seven datasets are used to test the model. The forecasting results of remainder stations reaffirm the superior performance of ESMD-SE-WPD-LSTM model for annual runoff forecasting. It can be seen from Tables 4, 5, 6, 7, 8, 9 and 10 that forecasting performances of seven prediction methods (except ANN) have little difference. The running time of the LSTM-based models is significantly longer than

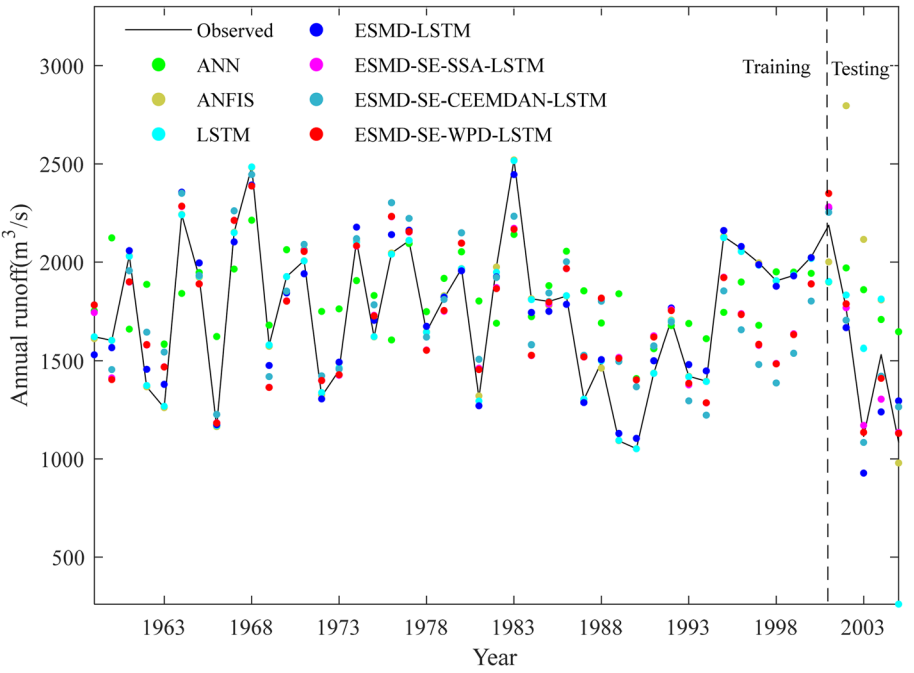

Fig. 14 Forecasting results of Hongjiadu Hydropower

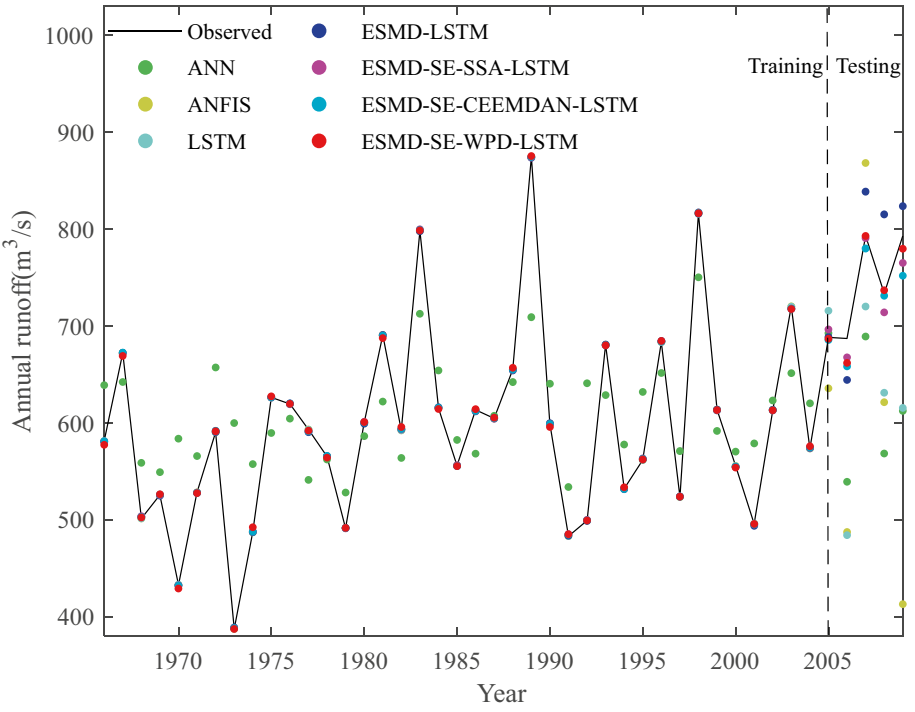

Fig. 15 Forecasting results of Yingluoxia station

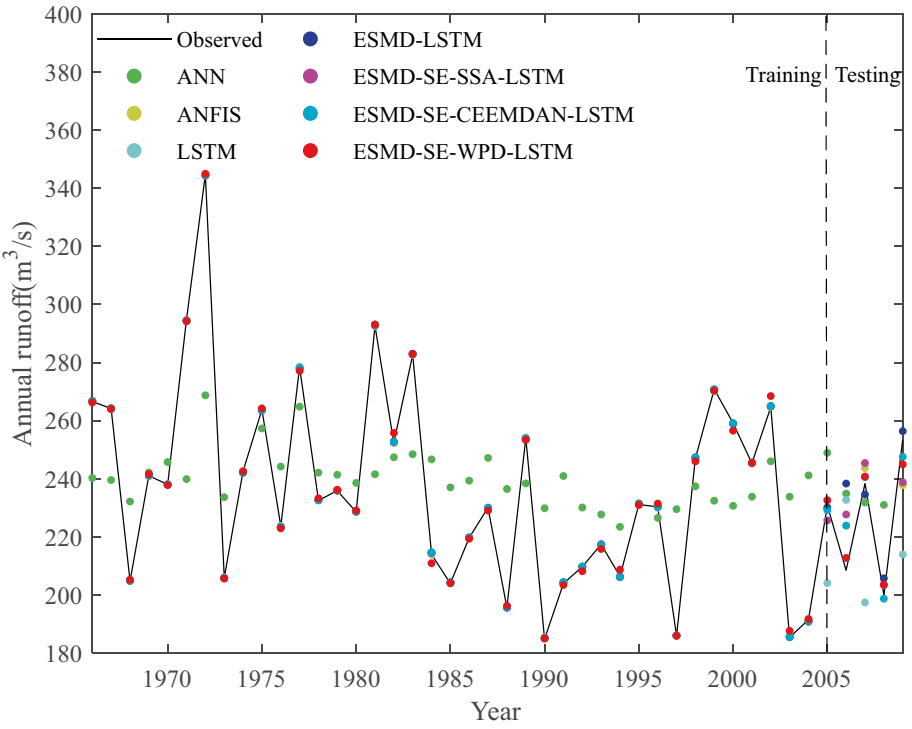

Fig. 16 Forecasting results of Jiayuguan station

those of ANN and ANFIS models. In addition, in the testing period, compared to three two-phase decomposition prediction methods, the deviation between the best and worst value of four indexes of ESMD-LSTM is clearer. Therefore, the following conclusions can be drawn:

(1) IMF1 is highly nonlinear and difficult to forecast, which can affect the overall prediction accuracy of models.
(2) The two-phase decomposition can capture important features better than the conventional single decomposition method. Besides, when comparing ESMD-SE-WPD-LSTM model with ESMD-SE-SSA-LSTM, and ESMD-SE-CEEMDAN-LSTM, the proposed model exhibits the best performance for all forecasting sites because the results in testing period are better. From Tables 4, 5, 6, 7, 8, 9 and 10 indicate that the prediction performance of ESMD-SE-SSA-LSTM is not stable, and the forecasting accuracy of ESMD-SE-CEEMDAN-LSTM is slightly inferior to the proposed model. Therefore, compared with SSA and CEEMDAN methods, WPD is more suitable to extract the significant features of IMF1. Overall, ESMD-SE-WPD-LSTM model outperforms all other methods. The reason may be that the model can make full use of the time–frequency positioning ability of WPD, the auto-adapted feature extraction properties of ESMD and the long-term memory function of LSTM.

Table 4 Errors of different models for Mopanshan Hydropower

Model		Training				Testing				Time(s)
		MAE	RMSE	MAPE	R	MAE	RMSE	MAPE	R	
ANN	Best	103.6335	127.966	20.156	0.556	47.624	53.144	8.481	0.653	0.185
	Worst	154.1805	175.339	30.661	0.047	59.637	69.562	11.567	0.094	0.231
	Average	100.1215	122.087	19.878	0.643	47.863	54.433	8.672	0.508	0.356
ANFIS	Best	0.0028	0.035	0.006	1.000	250.419	262.325	46.311	0.480	4.170
	Worst	0.004	0.051	0.008	1.000	258.676	270.953	47.818	0.451	4.100
	Average	0.0024	0.035	0.005	1.000	254.587	265.809	47.036	0.468	3.912
LSTM	Best	1.0406	1.459	0.211	1.000	93.794	102.520	17.646	0.755	20.457
	Worst	1.6592	2.042	0.322	0.999	138.112	168.590	25.502	0.681	20.281
	Average	1.0802	1.475	0.218	1.000	108.375	125.403	20.045	0.719	19.999
ESMD-LSTM	Best	1.3813	1.826	0.269	0.999	43.549	50.836	7.988	0.923	101.987
	Worst	1.9832	2.729	0.387	0.999	51.296	59.116	9.727	0.838	102.217
	Average	1.3519	1.943	0.276	0.999	54.164	58.545	10.305	0.872	100.572
ESMD-SE-SSA-LSTM	Best	1.4669	2.078	0.294	0.999	21.323	34.007	4.089	0.874	127.666
	Worst	1.7542	2.273	0.373	0.999	38.983	47.804	7.376	0.704	128.256
	Average	1.4591	2.047	0.296	0.999	24.529	39.424	4.707	0.802	127.465
ESMD-SE-CEEMDAN-LSTM	Best	5.049	6.068	1.028	0.993	31.801	41.921	6.003	0.889	225.190
	Worst	5.2551	6.481	1.088	0.991	42.564	53.490	8.042	0.829	224.970
	Average	5.0659	6.062	1.029	0.993	35.903	46.420	6.743	0.856	225.907
ESMD-SE-WPD-LSTM	Best	2.4857	2.888	0.494	0.998	12.421	15.279	2.247	0.979	225.010
	Worst	3.9869	4.624	0.800	0.995	30.306	36.667	5.981	0.949	224.970
	Average	2.094	2.828	0.416	0.998	21.674	26.499	4.329	0.988	225.907

Table 5 Errors of different models for Dahuofang Hydropower

Model		Training				Testing				Time(s)
		MAE	RMSE	MAPE	R	MAE	RMSE	MAPE	R	
ANN	Best	314.662	574.998	22.781	0.695	532.792	641.719	31.335	0.528	0.197
	Worst	772.085	1052.600	49.940	0.154	1801.400	2118.800	130.419	0.013	0.216
	Average	462.468	600.333	40.097	0.691	624.617	796.492	38.892	0.395	0.219
ANFIS	Best	0.396	0.507	0.035	1.000	1753.800	2407.100	157.585	0.365	4.376
	Worst	0.855	1.057	0.080	1.000	1937.700	2793.700	178.630	0.290	4.824
	Average	0.370	0.560	0.033	1.000	1758.700	2457.200	157.734	0.343	4.891
LSTM	Best	11.397	22.204	1.109	0.997	600.833	645.431	38.468	0.572	22.828
	Worst	22.969	29.982	2.401	0.996	1146.000	1320.200	84.255	0.109	21.098
	Average	12.792	23.436	1.278	0.997	801.349	814.195	56.921	0.493	20.995
ESMD-LSTM	Best	8.973	11.328	0.833	0.999	387.819	446.825	32.236	0.846	128.435
	Worst	35.041	43.941	2.910	0.993	671.176	965.175	58.630	0.325	127.452
	Average	10.556	13.934	0.987	0.998	519.415	697.446	45.451	0.620	125.891
ESMD-SE-SSA-LSTM	Best	12.171	16.483	1.084	0.998	256.839	287.340	19.248	0.970	153.437
	Worst	18.672	27.613	1.998	0.999	632.650	734.981	45.807	0.735	153.176
	Average	14.501	19.570	1.443	0.997	446.588	638.324	37.994	0.838	152.763
ESMD-SE-CEEMDAN-LSTM	Best	8.988	11.694	0.818	0.999	342.341	418.761	25.922	0.841	276.886
	Worst	20.900	27.884	1.828	0.995	427.193	491.578	33.766	0.776	276.411
	Average	10.483	13.181	0.946	0.999	374.680	449.113	28.619	0.821	277.994
ESMD-SE-WPD-LSTM	Best	14.328	16.827	1.286	0.998	282.188	320.126	21.292	0.963	303.327
	Worst	19.652	26.090	1.788	0.995	368.814	439.514	26.182	0.925	303.796
	Average	13.648	17.106	1.245	0.998	304.274	356.422	22.491	0.954	304.481

Table 6 Errors of different models for Biliuhe Hydropower

Model		Training				Testing				Time(s)
		MAE	RMSE	MAPE	R	MAE	RMSE	MAPE	R	
ANN	Best	115.658	251.415	21.846	0.710	121.903	142.065	38.818	0.604	0.227
	Worst	409.831	520.739	109.485	− 0.075	331.069	414.673	127.331	− 0.040	0.198
	Average	171.460	230.690	48.636	0.794	210.198	252.251	75.927	0.533	0.203
ANFIS	Best	0.036	0.047	0.010	1.000	476.179	725.913	243.555	0.875	4.501
	Worst	0.045	0.060	0.011	1.000	831.726	941.068	286.133	0.472	4.705
	Average	0.032	0.044	0.009	1.000	633.087	788.771	255.059	0.662	4.007
LSTM	Best	0.485	0.631	0.131	1.000	261.213	360.397	59.632	0.640	20.134
	Worst	12.240	14.855	3.290	0.997	275.128	379.293	62.292	0.583	20.019
	Average	1.483	1.895	0.449	1.000	267.131	372.326	61.103	0.616	20.039
ESMD-LSTM	Best	4.145	6.048	0.927	0.998	138.876	174.102	52.916	0.727	151.527
	Worst	6.885	8.720	1.506	0.997	171.403	195.796	59.820	0.651	151.734
	Average	4.589	6.239	0.996	0.998	145.547	177.607	55.083	0.692	152.218
ESMD-SE-SSA-LSTM	Best	6.656	8.332	1.448	0.997	46.079	55.712	14.674	0.962	179.674
	Worst	12.715	15.195	2.937	0.995	77.244	91.059	24.898	0.858	180.125
	Average	6.774	8.654	1.597	0.997	57.780	73.320	17.645	0.913	179.748
ESMD-SE-CEEMDAN-LSTM	Best	4.469	6.254	0.987	0.998	121.864	138.613	44.772	0.853	274.115
	Worst	4.475	6.351	0.983	0.998	138.671	155.693	50.215	0.768	274.705
	Average	4.475	6.286	0.981	0.998	132.088	149.257	47.422	0.817	276.385
ESMD-SE-WPD-LSTM	Best	5.740	7.382	1.370	0.998	43.903	48.246	16.200	0.963	322.568
	Worst	7.940	10.739	1.814	0.995	59.722	63.390	16.701	0.924	325.200
	Average	6.134	7.721	1.398	0.997	50.250	51.593	16.466	0.948	323.916

Table 7 Errors of different models for Changshui station

Model		Training				Testing				Time(s)
		MAE	RMSE	MAPE	R	MAE	RMSE	MAPE	R	
ANN	Best	2.741	3.898	50.029	0.606	1.920	2.576	63.697	0.430	0.342
	Worst	5.127	7.093	62.850	− 0.063	3.133	3.670	104.327	− 0.302	0.336
	Average	2.556	3.644	54.522	0.649	2.011	2.202	71.429	0.244	0.241
ANFIS	Best	0.001	0.001	0.002	1.000	2.392	2.795	71.938	0.607	4.101
	Worst	0.002	0.027	0.042	1.000	4.113	4.936	108.942	− 0.273	4.008
	Average	0.002	0.002	0.007	1.000	3.228	3.401	94.484	0.199	4.500
LSTM	Best	0.040	0.057	0.948	0.999	2.879	3.238	87.308	0.467	21.989
	Worst	0.056	0.083	1.208	0.999	3.269	3.731	108.413	0.250	20.949
	Average	0.040	0.063	0.995	0.999	2.392	2.975	78.637	0.326	20.928
ESMD-LSTM	Best	0.063	0.078	1.351	0.999	0.979	1.192	30.366	0.989	106.768
	Worst	0.109	0.136	2.076	0.997	1.513	1.993	61.629	0.706	105.818
	Average	0.069	0.087	1.471	0.998	1.502	1.542	48.034	0.946	105.096
ESMD-SE-SSA-LSTM	Best	0.511	0.640	10.242	0.992	1.223	1.570	48.562	0.527	132.742
	Worst	0.512	0.644	10.302	0.992	1.328	1.791	55.026	0.251	133.115
	Average	0.511	0.641	10.279	0.992	2.195	2.246	73.706	0.585	132.991
ESMD-SE-CEEMDAN-LSTM	Best	0.067	0.088	1.556	0.998	1.853	1.994	66.801	0.961	244.059
	Worst	0.111	0.135	2.330	0.996	2.060	2.252	73.591	0.909	248.916
	Average	0.070	0.090	1.579	0.998	1.703	1.827	61.345	0.962	246.112
ESMD-SE-WPD-LSTM	Best	0.093	0.117	2.118	0.997	0.186	0.218	7.066	0.990	326.525
	Worst	0.127	0.155	3.182	0.994	0.366	0.415	12.537	0.951	326.424
	Average	0.100	0.129	2.315	0.996	0.198	0.224	6.519	0.992	325.925

Table 8 Errors of different models for Hongjiadu Hydropower

Model		Training				Testing				Time (s)
		MAE	RMSE	MAPE	R	MAE	RMSE	MAPE	R	
ANN	Best	257.476	331.208	15.178	0.544	280.528	370.161	20.994	0.502	0.146
	Worst	408.180	522.754	27.035	0.255	510.326	611.734	41.693	0.412	0.186
	Average	260.127	326.858	16.314	0.586	395.283	453.279	31.005	0.594	0.190
ANFIS	Best	3.584	8.125	0.229	0.998	493.267	621.128	35.375	0.552	3.892
	Worst	5.547	12.374	0.362	0.999	610.115	714.938	43.747	0.312	4.504
	Average	3.764	8.850	0.242	0.997	519.370	660.670	36.834	0.495	4.316
LSTM	Best	4.788	7.190	0.293	0.999	221.386	282.623	16.581	0.850	21.587
	Worst	8.226	12.039	0.484	0.996	626.037	744.812	50.465	0.471	19.462
	Average	4.872	7.435	0.289	0.998	380.449	456.410	30.235	0.681	19.930
ESMD-LSTM	Best	50.188	59.236	2.948	0.987	167.567	188.646	11.722	0.906	121.830
	Worst	52.258	61.356	3.100	0.986	266.561	295.721	18.845	0.881	121.391
	Average	50.488	59.620	2.983	0.987	175.949	191.315	13.030	0.918	121.696
ESMD-SE-SSA-LSTM	Best	50.238	59.496	2.960	0.987	114.419	133.867	7.879	0.973	151.706
	Worst	50.908	60.651	2.995	0.986	159.189	181.265	11.800	0.900	152.309
	Average	50.244	59.682	2.958	0.987	85.682	114.782	5.781	0.965	151.553
ESMD-SE-CEEMDAN-LSTM	Best	52.735	62.286	3.099	0.985	81.580	105.497	5.616	0.978	254.192
	Worst	52.846	62.483	3.101	0.985	113.594	137.235	8.424	0.951	257.315
	Average	59.078	80.432	3.389	0.977	90.172	104.198	6.657	0.968	252.306
ESMD-SE-WPD-LSTM	Best	49.232	58.831	2.887	0.987	74.997	93.421	4.317	0.989	293.269
	Worst	50.491	60.122	2.929	0.987	93.849	117.755	6.094	0.976	300.550
	Average	49.744	59.126	2.915	0.987	74.221	93.800	4.531	0.984	297.442

Table 9 Errors of different models for Jiayuguan station

Model		Training				Testing				Time (s)
		MAE	RMSE	MAPE	R	MAE	RMSE	MAPE	R	
ANN	Best	21.536	30.456	9.135	0.529	14.813	19.590	6.262	0.470	0.210
	Worst	27.519	35.615	11.387	0.160	26.445	27.786	11.673	− 0.157	0.212
	Average	24.071	29.615	10.461	0.610	19.529	21.341	8.989	0.339	0.200
ANFIS	Best	0.001	0.002	0.004	1.000	66.552	85.160	30.348	0.618	5.274
	Worst	0.009	0.017	0.004	1.000	70.192	90.321	31.762	− 0.117	5.009
	Average	0.005	0.009	0.002	1.000	60.087	91.881	26.837	0.334	4.682
LSTM	Best	0.087	0.176	0.038	1.000	31.506	32.548	13.897	0.463	25.892
	Worst	0.848	0.951	0.370	0.996	33.238	35.443	15.142	0.315	25.304
	Average	0.130	0.130	0.057	1.000	36.865	38.314	16.493	0.422	24.363
ESMD-LSTM	Best	0.206	0.259	0.089	1.000	8.747	12.566	3.935	0.856	103.765
	Worst	0.369	0.463	0.158	0.999	8.791	15.488	4.215	0.730	103.704
	Average	0.225	0.281	0.097	1.000	8.640	13.805	4.067	0.792	104.787
ESMD-SE-SSA-LSTM	Best	0.240	0.282	0.104	1.000	8.615	10.514	3.868	0.859	79.492
	Worst	0.371	0.483	0.158	0.999	12.674	14.250	5.576	0.706	79.772
	Average	0.236	0.290	0.101	1.000	9.901	11.591	4.384	0.817	79.207
ESMD-SE-CEEMDAN-LSTM	Best	0.223	0.278	0.097	1.000	4.863	6.379	2.211	0.959	182.369
	Worst	0.280	0.338	0.120	1.000	9.386	12.012	4.207	0.790	181.375
	Average	0.233	0.284	0.100	1.000	5.219	7.479	2.369	0.932	182.245
ESMD-SE-WPD-LSTM	Best	0.810	1.203	0.351	0.999	3.979	4.244	1.744	0.984	239.897
	Worst	1.067	1.495	0.454	0.991	4.367	6.718	1.830	0.962	238.447
	Average	0.881	1.302	0.380	0.993	4.061	4.483	1.759	0.981	239.348

Table 10 Errors of different models for Yingluoxia station

Model		Training				Testing				Time(s)
		MAE	RMSE	MAPE	R	MAE	RMSE	MAPE	R	
ANN	Best	64.490	84.015	11.338	0.523	88.382	105.261	11.884	0.283	0.174
	Worst	40.518	89.334	7.993	0.645	226.942	257.476	31.122	0.146	0.155
	Average	54.663	70.502	9.858	0.754	120.372	136.057	16.110	0.281	0.164
ANFIS	Best	0.001	0.002	0.002	1.000	143.242	174.800	19.133	0.235	4.548
	Worst	0.002	0.003	0.004	1.000	167.948	220.976	22.441	0.126	4.379
	Average	0.001	0.009	0.001	1.000	164.049	202.745	21.897	0.227	4.631
LSTM	Best	0.128	0.168	0.022	1.000	120.730	146.530	16.493	0.406	26.550
	Worst	0.273	0.338	0.045	1.000	119.490	139.611	16.171	0.340	26.406
	Average	0.141	0.181	0.024	1.000	116.704	133.669	15.822	0.360	26.868
ESMD-LSTM	Best	0.713	0.932	0.123	1.000	34.891	40.074	4.632	0.956	96.493
	Worst	2.301	2.619	0.403	0.997	55.095	60.544	7.480	0.913	93.721
	Average	0.773	1.011	0.135	0.999	47.408	52.505	6.339	0.935	94.890
ESMD-SE-SSA-LSTM	Best	0.724	0.939	0.124	1.000	18.781	21.105	2.529	0.972	119.696
	Worst	1.671	2.002	0.286	0.998	20.414	34.048	2.649	0.761	119.435
	Average	0.785	0.984	0.135	1.000	15.463	17.914	2.097	0.960	119.845
ESMD-SE-CEEMDAN-LSTM	Best	0.692	0.901	0.121	1.000	38.084	39.165	5.158	0.981	215.225
	Worst	0.821	1.011	0.141	1.000	31.269	37.261	4.225	0.895	214.288
	Average	0.702	0.896	0.122	1.000	17.541	23.194	2.339	0.947	214.748
ESMD-SE-WPD-LSTM	Best	1.313	1.708	0.228	0.999	13.772	15.823	1.900	0.990	262.862
	Worst	2.347	2.883	0.410	0.996	17.610	21.808	2.403	0.928	267.305
	Average	1.386	1.877	0.241	0.998	8.626	12.823	1.198	0.980	267.121

Table 11 Errors of ESMD-LSTM for seven stations

Station	Subseries	Training				Testing			
		MAE	RMSE	MAPE	R	MAE	RMSE	MAPE	R
Mopanshan	IMF1	0.204	0.256	0.455	1.000	59.573	63.459	76.187	0.841
	IMF2	1.012	1.782	26.685	0.995	3.768	5.822	16.644	0.974
	IMF3	0.572	0.728	2.323	0.999	4.145	4.732	37.538	0.983
	Res	0.521	0.599	0.097	1.000	1.612	2.183	0.275	0.993
Dahuofang	IMF1	7.147	10.183	2.599	0.999	631.850	796.257	60.806	0.654
	IMF2	2.528	3.498	4.539	1.000	173.223	189.550	240.329	0.893
	IMF3	5.544	8.306	20.035	0.998	86.861	114.943	12.776	0.879
	IMF4	0.338	0.430	2.291	0.999	6.629	7.448	17.662	0.997
	Res	0.594	0.729	0.039	0.999	3.179	3.792	0.204	0.999
Biliuhe	IMF1	4.223	1.020	0.654	1.000	163.895	184.833	145.505	0.795
	IMF2	0.513	0.671	0.848	1.000	24.531	28.420	136.722	0.874
	IMF3	4.056	6.123	6.955	0.999	12.658	13.031	23.025	0.975
	IMF4	0.575	0.734	4.251	0.999	7.522	9.037	14.813	0.997
	IMF5	0.499	0.676	2.603	0.999	2.841	3.993	11.780	0.994
	Res	0.858	1.094	0.146	1.000	4.223	4.516	1.156	0.999
Changshui	IMF1	0.031	0.031	1.911	1.000	1.557	1.672	225.483	0.573
	IMF2	0.041	0.041	7.268	0.999	0.453	0.275	223.104	0.810
	IMF3	0.054	0.054	8.518	0.992	0.147	0.158	17.122	0.996
	Res	0.042	0.042	0.391	0.999	0.056	0.095	2.346	1.000
Hongjiadu	IMF1	2.276	2.835	3.419	1.000	106.845	167.810	263.110	0.841
	IMF2	1.244	2.632	1.878	0.999	25.535	32.088	8.222	0.973
	IMF3	1.720	2.331	7.123	0.997	13.307	15.248	17.712	0.999
	IMF4	0.899	2.646	11.965	0.994	3.501	4.691	6.942	0.998
	Res	0.608	0.821	0.035	0.999	0.650	0.653	0.039	1.000
Jiayuguan	IMF1	0.048	0.060	0.259	1.000	10.431	12.385	57.132	0.828
	IMF2	0.069	0.090	1.986	1.000	5.795	6.675	304.374	0.868
	IMF3	0.095	0.123	4.587	0.999	1.438	1.765	28.024	0.989
	Res	0.178	0.230	0.075	0.999	0.751	0.850	0.344	0.982
Yingluoxia	IMF1	0.556	0.675	1.217	1.000	43.375	51.254	144.202	0.616
	IMF2	0.374	0.621	2.457	0.999	5.906	8.227	61.913	0.938
	IMF3	0.411	0.583	4.060	0.998	3.080	3.576	38.698	0.983
	Res	0.379	0.465	0.063	1.000	7.578	8.842	1.024	0.996

4.5.4 Comparison of All Investigated Models

The performances of all investigated models developed in this study are shown in Figs. 10, 11, 12, 13, 14, 15 and 16, which imply that forecasting performances of six models (except ANN) in the training phase are slightly overestimated. Meanwhile, in the testing phase, the forecasting accuracy of all sites can be significantly improved, and performances of different models are uneven. The proposed hybrid prediction model with the secondary decomposition provides the best performance as the trend line is very close to observed data line, and the method can capture abrupt changes in annual runoff series.

4.6 Discussion of Results

Experimental results demonstrate these differences between LSTM, ANFIS and ANN, indicating the importance of choosing an appropriate forecasting method. ANFIS is greatly affected by the clustering parameters, which limits the performance of the model. The gradient-based training strategy of the traditional ANN may suffer from dimensionality and overfitting issues. LSTM discards or retains information to the cell state using unique gate structure. If information at a certain time is more important, the forget gate can keep the information transmission, which is one of the reasons why LSTM can process long sequences. The gate structure of LSTM overcomes the weaknesses of ANFIS and ANN to some extent and attains relatively better prediction results. However, runoff contains different frequency components due to influencing climate, underlying surface characteristic of river basin, human activities, etc. Therefore, it is difficult for a single prediction model to fully reflect the formation mechanism of runoff since only one resolution component is used to construct the forecasting model. In this paper, data preprocessing technologies are utilized to identify the resolution subcomponents. The characteristics of each component can be separated, which reduces the difficulty of modeling. Therefore, the modified LSTM models can provide better performance than standard LSTM model.

The followings are an analysis of possible reasons why the proposed hybrid model (ESMD-SE-WPD-LSTM) can improve the forecasting accuracy. Firstly, ESMD-SE-WPD decomposes the original data into several more linear sub-series, facilitating comprehensive identification of frequency features in original data. Secondly, LSTM is employed to model complex relationships of input–output variables in each subseries. Through specially designed model architecture, LSTM overcomes the shortcoming of RNN and provides an avenue for deeply exploring internal features of runoff time series. Finally, ESMD-SE-WPD-LSTM hybrid model overcomes shortcomings of a single LSTM method by generating synergistic effect in the prediction. Overall, the incorporation of data preprocessing and sample entropy into LSTM model can provide more accurate and reliable results for long-term runoff prediction.

5 Conclusion

Long-term runoff forecasting plays a critical role in the management and monitoring of water resources. To attain more accurate prediction of annual runoff, this paper presents a hybrid model for long-term runoff prediction, which couples two-phase decomposition and LSTM (ESMD-SE-WPD-LSTM). Firstly, ESMD is used to decompose the original time series, and SE (sample entropy) of all sub-series is computed. Secondly, the sub-series with the maximum SE is adopted for secondary decomposition using WPD method, which can provide more linear subseries. Next, LSTM model is employed to train and forecast the data. Finally, the forecasting accuracy of the proposed model is compared with ANN, ANFIS, ESMD-LSTM, ESMD-SE-SSA-LSTM, and ESMD-SE-CEEMDAN-LSTM. The forecasting errors of all investigated models are evaluated based on four evaluation indexes. According to the results, the following conclusions can be drawn:

Firstly, the proposed hybrid model with secondary decomposition provides the most robust performance and excellent forecasting accuracy among all investigated models.

This demonstrates that the proposed model can significantly improve the prediction accuracy of long-term runoff time series.

Secondly, the forecasting accuracies of hybrid methods (ESMD-LSTM, ESMD-SE-SSA-LSTM, ESMD-SE-CEEMDAN-LSTM, and ESMD-SE-WPD-LSTM) preprocessed by decomposition method are superior to those of ANN, ANFIS, and LSTM models, demonstrating high efficiency of data preprocessing technology in reducing non-linearity of runoff series.

Thirdly, EMSD and WPD, as two signal processing methods with high efficiency, can complement each other. After screening by sample entropy, the original single time series is re-decomposed by two-phase decomposition mode to attain a more linear annual time series, which reduces the complexity of forecasting, and mitigate the limitation of conventional single-phase decomposition method.

The hybrid model presented in this paper combines data preprocessing technology, sample entropy, and forecasting model to develop runoff forecasting model, which is more conducive to be a useful and efficient soft computing model to forecast runoff time series.

Authors' Contributions WW: Conceptualization, Methodology, Writing-original draft. YD: Methodology, data curation, Writing—original draft preparation. KC: Writing and editing-original draft. DX: Formal analysis and data collection. CL: Formal analysis. QM: Investigation.

Funding Project of key science and technology of the Henan province (No: 202102310259; No: 202102310588), Henan province university scientific and technological innovation team (No: 18IRTSTHN009).

Availability of Data and Materials All authors made sure that all data and materials support our published claims and comply with field standards.

Declarations

Ethics Approval All authors kept the 'Ethical Responsibilities of Authors'.

Consent to Participate All authors gave explicit consent to participate in this work.

Consent to Publish All authors gave explicit consent to publish this manuscript.

Competing of interest The authors declare that they have no conflict of interest.

References

Akbari Asanjan A, Yang T, Hsu K, Sorooshian S, Lin J, Peng Q (2018) Short-term precipitation forecast based on the PERSIANN system and LSTM recurrent neural networks. J Geophys Res Atmos 123:12543–12563. https://doi.org/10.1029/2018JD028375

Al-Juboori AM (2021) A hybrid model to predict monthly streamflow using neighboring rivers. Annu Flows Water Resour Manag 35:729–743. https://doi.org/10.1007/s11269-020-02757-4

Alcaraz R, Rieta JJ (2010) A review on sample entropy applications for the non-invasive analysis of atrial fibrillation electrocardiograms. Biomed Signal Process Control 5:1–14. https://doi.org/10.1016/j.bspc.2009.11.001

Alickovic E, Kevric J, Subasi A (2018) Performance evaluation of empirical mode decomposition, discrete wavelet transform, and wavelet packed decomposition for automated epileptic seizure detection and prediction. Biomed Signal Process Control 39:94–102. https://doi.org/10.1016/j.bspc.2017.07.022

Bai Y, Bezak N, Zeng B, Li C, Sapac K, Zhang J (2021) Daily runoff forecasting using a cascade long short-term memory model that considers different variables. Water Resour Manag 35:1167–1181. https://doi.org/10.1007/s11269-020-02759-2

Bojang PO, Yang TC, Pham QB, Yu PS (2020) Linking singular spectrum analysis and machine learning for monthly rainfall forecasting. Appl Sci. https://doi.org/10.3390/app10093224

Chau KW, Wu CL, Li YS (2005) Comparison of several flood forecasting models in Yangtze River. J Hydrol Eng 10:485–491. https://doi.org/10.1061/(ASCE)1084-0699(2005)10:6(485)

Colominas MA, Schlotthauer G, Torres ME, Flandrin P (2012) Noise-assisted EMD methods in action. Adv Adapt Data Anal 04:1250025. https://doi.org/10.1142/S1793536912500252

Dong Q, Sun Y, Li P (2017) A novel forecasting model based on a hybrid processing strategy and an optimized local linear fuzzy neural network to make wind power forecasting: a case study of wind farms in China. Renew Energy 102:241–257. https://doi.org/10.1016/j.renene.2016.10.030

Feng Z, Liu S, Niu W, Li S, Wu H, Wang J (2020a) Ecological operation of cascade hydropower reservoirs by elite-guide gravitational search algorithm with Lévy flight local search and mutation. J Hydrol 581:124425. https://doi.org/10.1016/j.jhydrol.2019.124425

Feng Z, Niu W, Cheng X, Wang J, Wang S, Song Z (2020b) An effective three-stage hybrid optimization method for source-network-load power generation of cascade hydropower reservoirs serving multiple interconnected power grids. J Clean Prod. https://doi.org/10.1016/j.jclepro.2019.119035

He XX, Luo JG, Li P, Zuo GG, Xie JC (2020) A hybrid model based on variational mode decomposition and gradient boosting regression tree for monthly runoff forecasting. Water Resour Manag 34:865–884. https://doi.org/10.1007/s11269-020-02483-x

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

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

Meng ER et al (2021) A hybrid VMD-SVM model for practical streamflow prediction using an innovative input selection framework. Water Resour Manag 35:1321–1337. https://doi.org/10.1007/s11269-021-02786-7

Parisouj P, Mohebzadeh H, Lee T (2020) Employing machine learning algorithms for streamflow prediction: a case study of four river basins with different climatic zones in the United States. Water Resour Manag 34:4113–4131. https://doi.org/10.1007/s11269-020-02659-5

Poul AK, Shourian M, Ebrahimi H (2019) A comparative study of MLR, KNN, ANN and ANFIS models with Wavelet transform in monthly stream flow prediction. Water Resour Manag 33:2907–2923. https://doi.org/10.1007/s11269-019-02273-0

Reddy BSN, Pramada SK, Roshni T (2021) Monthly surface runoff prediction using artificial intelligence: a study from a tropical climate river basin. J Earth Syst Sci. https://doi.org/10.1007/s12040-020-01508-8

Saeed A, Li C, Danish M, Rubaiee S, Tang G, Gan Z, Ahmed A (2020) Hybrid bidirectional LSTM model for short-term wind speed interval prediction. IEEE Access 8:182283–182294. https://doi.org/10.1109/access.2020.3027977

Sahoo BB, Jha R, Singh A, Kumar D (2019) Long short-term memory (LSTM) recurrent neural network for low-flow hydrological time series forecasting. Acta Geophys 67:1471–1481. https://doi.org/10.1007/s11600-019-00330-1

Seo Y, Kim S, Kisi O, Singh VP, Parasuraman K (2016) River stage forecasting using wavelet packet decomposition and machine learning models. Water Resour Manag 30:4011–4035. https://doi.org/10.1007/s11269-016-1409-4

Sun D, Zhang H, Guo Z (2018) Complexity Analysis of precipitation and runoff series based on approximate entropy and extreme-point symmetric mode. Decomposition 10:1388

Sun SZ, Fu JQ, Zhu F, Du DJ (2020) A hybrid structure of an extreme learning machine combined with feature selection, signal decomposition and parameter optimization for short-term wind speed forecasting. Trans Inst Meas Control 42:3–21. https://doi.org/10.1177/0142331218771141

Sun W, Huang C (2020) A hybrid air pollutant concentration prediction model combining secondary decomposition and sequence reconstruction. Environ Pollut. https://doi.org/10.1016/j.envpol.2020.115216

Tao YM, Gao XG, Ihler A, Sorooshian S, Hsu KL (2017) Precipitation identification with bispectral satellite information using deep learning approaches. J Hydrometeorol 18:1271–1283. https://doi.org/10.1175/jhm-d-16-0176.1

Tayyab M, Zhou JZ, Dong XH, Ahmad I, Sun N (2019) Rainfall-runoff modeling at Jinsha River basin by integrated neural network with discrete wavelet transform. Meteorol Atmos Phys 131:115–125. https://doi.org/10.1007/s00703-017-0546-5

Wang J-L, Li Z-J (2013) Extreme-point symmetric mode decomposition method for data analysis. Adv Adapt Data Anal 05:1350015. https://doi.org/10.1142/S1793536913500155

Xiang ZR, Yan J, Demir I (2020) A rainfall-runoff model with LSTM-based sequence-to-sequence learning. Water Resour Res. https://doi.org/10.1029/2019wr025326

Yen MH, Liu DW, Hsin YC, Lin CE, Chen CC (2019) Application of the deep learning for the prediction of rainfall in Southern Taiwan. Sci Rep. https://doi.org/10.1038/s41598-019-49242-6

Yin Y, Bai Y, Ge F, Yu H, Liu Y (2019) Long-term robust identification potential of a wavelet packet decomposition based recursive drift correction of E-nose data for Chinese spirits. Measurement 139:284–292. https://doi.org/10.1016/j.measurement.2019.03.011

Yuan RF et al (2021) Daily runoff forecasting using ensemble empirical mode decomposition and long short-term memory. Front Earth Sci. https://doi.org/10.3389/feart.2021.621780

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

Zhang JF, Zhu Y, Zhang XP, Ye M, Yang JZ (2018) Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. J Hydrol 561:918–929. https://doi.org/10.1016/j.jhydrol.2018.04.065

Zhou J, Xu X, Huo X, Li Y (2019) Forecasting Models for wind power using extreme-point symmetric mode decomposition and artificial neural networks. Sustainability. https://doi.org/10.3390/su11030650

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

Authors and Affiliations

Wen-chuan Wang[1] · Yu-jin Du[1] · Kwok-wing Chau[2] · Dong-mei Xu[1] · Chang-jun Liu[3] · Qiang Ma[3]

Yu-jin Du
2366036835@qq.com

Kwok-wing Chau
cekwchau@polyu.edu.hk

Dong-mei Xu
xudongmei@ncwu.edu.cn

Chang-jun Liu
lcj2005@iwhr.com

Qiang Ma
maqiang@iwhr.com

[1] College of Water Resources, Henan Key Laboratory of Water Resources Conservation and Intensive Utilization in the Yellow River Basin, North China University of Water Resources and Electric Power, Zhengzhou 450046, People's Republic of China

[2] Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, People's Republic of China

[3] China Institute of Water Resources and Hydropower Research, Beijing 100081, People's Republic of China