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

Wenchuan Wang (wangwen1621@163.com)
North China University of Water Resources and Electric Power https://orcid.org/0000-0003-1367-5886
Yu-jin Du
North China University of Water Resources and Electric Power
Kwok-wing Chau
The Hong Kong Polytechnic University
Dong-mei Xu
North China University of Water Resources and Electric Power
Chang-jun Liu
China Institute of Water Resources and Hydropower Research
Qiang Ma
China Institute of Water Resources and Hydropower Research

Research Article

Keywords: annual runoff prediction, two-phase decomposition, Long Short-Term Memory, extreme-point symmetric mode decomposition, wavelet packet decomposition, sample entropy

Posted Date: May 28th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-269127/v1

An ensemble hybrid forecasting model for annual runoff based on sample entropy, secondary decomposition, and long short-term memory neural network

Wen-chuan Wang
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
Corresponding author, E-mail: wangwen1621@163.com; wangwenchuan@ncwu.edu.cn

Yu-jin Du
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
E-mail: 2366036835@qq.com

Kwok-wing Chau
Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, People's Republic of China
E-mail: cekwchau@polyu.edu.hk

Dong-mei Xu
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
E-mail: xudongmei@ncwu.edu.cn

Chang-jun Liu
China Institute of Water Resources and Hydropower Research, Beijing 100081, People's Republic of China
E-mail:lcj2005@iwhr.com

Qiang Ma
China Institute of Water Resources and Hydropower Research, Beijing 100081, People's Republic of China
E-mail: maqiang@iwhr.com

Abstract: Accurate and consistent annual runoff prediction in regions is a hot topic in the management, optimization, and monitoring of water resources. A novel prediction model (ESMD-SE-WPD-LSTM) is presented in this study. Firstly, the 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, the sample entropy (SE) method is employed to measure the complexity of each IMF. Thirdly, we adopt wavelet packet decomposition (WPD) to further decompose the IMF with the maximum SE into several appropriate components and detailed components. Then the LSTM model, a deep learning algorithm based recurrent approach, is employed to predict all components obtained in the previous step. Finally, the forecasting results of all components are aggregated to generate the final prediction. The proposed model, which is applied to five annual series from different areas in China, is evaluated based on four quantitative indexes (R, NSEC, MAPE and RMSE). The results indicate that the ESMD-SE-WPD-LSTM outperforms other benchmark models in terms of four quantitative indexes. Hence the proposed model can provide higher accuracy and consistency for annual runoff prediction, making 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 the optimal management of hydro-resources (Evsukoff et al. 2012; Meng et al. 2019), 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 of researchers (Kisi and Sanikhani 2015; Wang et al. 2009; Xiang et al. 2020; Zamani Sabzi et al. 2017). Numerous models have been presented to improve the prediction accuracy of annual hydrologic time series (Feng et al. 2020c; Tan et al. 2018; Wang et al. 2015). Generally, these models can be divided into two types (Chau et al. 2005; He et al. 2014), namely physical-based and data-driven models. Physical-based models require detailed multi-source information, powerful mathematical calculation tools and sophisticated parameter optimization process (Aqil et al. 2007) while data-driven models are an efficient alternative by building direct relationships between input and output data without understanding the complex physical mechanisms of the system. Recently, many data-driven models (such as LSTM, ANFIS, ANN, etc.) have been adopted in hydrological forecasting field (Ehteram et al. 2019; Parisouj et al. 2020; Sahoo et al. 2019; Song et al. 2020). Hence, this study 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 (Liu et al. 2020; Tao et al. 2017; Yen et al. 2019). Recurrent neural network (RNN), a deep learning algorithm, is capable of modeling complex temporal dynamics. Numerous improvement methods have been undertaken to overcome the problems of vanishing gradients and gradient explosion. As a representative of them, LSTM has been used in the field of signal recognition and forecast. The LSTM model and its principles, proposed by Hochreiter and Schmidhuber (1997), has been adopted in hydrology field as well. The 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 forecasting monthly runoff in Astor River Basin by integrating LSTM and ant lion optimizer algorithm. Hu et al. (2019) constructed a water quality forecasting model based on LSTM using preprocessed data. Zhu et al. (2020) presented a probabilistic LSTM coupled with Gaussian process model for probabilistic daily runoff forecasting. Srinivas et al. (2020) developed a method using RNN and LSTM to improve the rainfall prediction performance. Saeed et al. (2020) proposed a method for wind speed prediction by using a bidirectional LSTM model and automatic encoder. Gao et al. (2020) used LSTM and gated recurrent unit networks for shore-term flood forecasting. These studies prove the competitiveness and high stability of LSTM hydrological time series prediction.

Recently, many hybrid forecasting models for hydrological time series prediction have been developed and studied, which mainly include the forecast modeling and data preprocessing. The decomposition algorithm can significantly enhance the forecasting ability of the model by decomposing the raw hydrological series into more clean sub-series. The emergence of multi-resolution decomposition tools, i.e., principal component analysis (PCA), singular spectrum analysis (SSA), empirical mode decomposition (EMD), extreme-point symmetric mode decomposition (ESMD), ensemble empirical mode decomposition (EEMD), CEEMDAN (complete EEMD with adaptive noise), wavelet packet decomposition (WPD), wavelet transform (WT), and the variational mode decomposition (VMD), have further stimulated researchers to make in-depth research on data preprocessing. 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 involves certain subjective factors in the process of noise reduction and is subject to the restriction of r matrix perturbation. Wang and Zhou (2020) examined streamflow prediction at each hydrological station in the mainstream of the Yellow River in China by coupling PCA with time series analysis method, but it can only be utilized to linear dimensionality reduction of data. Discrete wavelet transform (DWT), as a classical data analysis method, is capable of helping a forecasting model to extract useful information (Feng et al. 2015; Tayyab et al. 2019), but it may suffer from signal loss. Zuo et al. (2020) proposed a single-model forecasting (SF) framework namely SF-VMD-LSTM, to forecast daily streamflow. However, the drawback of VMD is that the optimal parameter combination needs to be set in advance artificially. EMD and EEMD techniques lack an accurate mathematical theory. To overcome the weaknesses of them, a further improvement of EMD called ESMD, proposed by Wang and Li (2013), was adopted to reduce the uncertainty and noise of signal. The main idea of ESMD is to identify the large-scale cycle and nonlinear trend of the data using the internal extreme-point symmetry interpolation according to the characteristics of the data itself. ESMD replaces the traditional integral transformation with direct interpolation, and the residual is optimized by the least square approach. Therefore, ESMD is capable of more intuitively reflecting the time-varying characteristics of frequency and amplitude of each component. Hence, ESMD is very suitable to analyze nonlinear and non-stationary series. While this method has been successfully used in broad fields such as economics, medicine, atmospheric science, and hydrology (Li et al. 2017; Lin et al. 2017; Zhou et al. 2019a), few attempts have been made to use the latest advance in ESMD to solve the problem of hydrological time series prediction. Therefore, an objective of this article is to explore the efficiency of ESMD in capturing hydrological time series characteristics.

WPD is another data decomposition technology that has gained numerous attentions recently. The main idea of WPD is using multiple filters to split the original signal into several sub-series with different frequency characteristics. WPD, an improvement of DWT, decomposes the approximation value same as details of signals in each level of decomposition. DWT only decomposes the approximation coefficient, while WPD has the capability for splitting both detail coefficient and approximation coefficient simultaneously, thereby the later provides more possibilities for hydrological time series. Seo et al. (2016) combined three models, including SVM (support vector machine), ANFIS, and ANN, with WPD for daily stream segment 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 still needs to be further studied to fill an important research gap in mid- and long-term runoff forecasting.

Despite the above fruitful results of decomposition technology, it should be pointed out that single decomposition may be difficult to fully mitigate signal nonlinearity. To attain a smoother series and higher forecasting accuracy than one-phase decomposition, Liu et al. (2018b) proposed a wind speed multistep prediction model by combining VMD, SSA, LSTM, and extreme learning machine (ELM). Sun and Huang (2020) combined secondary decomposition (SD) with sequence reconstruction to predict air pollutant concentration, and the model had excellent and robust prediction performance. Li et al. (2020) proposed a novel hybrid air cargo forecasting method by coupling a new SD ensemble method with cuckoo search algorithm. In summary, SD method can further mitigate signal nonlinearity and solve the limitation of single decomposition method to a certain extent. Meanwhile, there are few attempts to use combined decomposition approach to solve hydrological time series prediction. There, this paper firstly uses the secondary decomposition framework (ESMD-SE-WPD) to attain a more linear series. Then, LSTM is adopted for annual runoff forecasting. Finally, the forecasted results of all sub-series are summed to generate the final prediction. The proposed model performs runoff forecasting for two real-world runoff series at Liaoning and Henan Provinces, China, respectively. The performance of the developed model is compared with several benchmarking prediction models (LSTM, ANFIS, ANN, ESMD-LSTM, ESMD-SE-SSA-LSTM, ESMD-SE-CEEMDAN-LSTM).

The innovation contribution of this paper can be generalized as follows: (a) On the basis of complexity judgment of SE, the proposed model adopts an ensemble hybrid method (ESMD-SE-WPD-LSTM) to preprocess the annual runoff time series, which efficiently mitigates the non- stationarity of series, and greatly reduces the forecasting difficulties. (b) For different series, seven models, namely LSTM, ANFIS, ANN, ESMD-LSTM, ESMD-SE-SSA-LSTM, ESMD-SE-CEEMDAN-LSTM and ESMD-SE-WPD-LSTM, are employed for benchmark comparison to study the prediction performance of the proposed framework. (c) A comprehensive evaluation of the forecasting accuracy of the presented framework by combining four quantitative indexes (R, NSEC, RMSE, and MAPE), a depth analysis of the forecasting indexes of each method, the efficiency of the presented model, is comprehensively verified.

The remainder of this article is arranged as follows: Section 2 provides the basic theories of the relevant methods. After depicting the data source and quantitative indexes in Section 3, Section 4 introduces the content of empirical forecasting experiments and discussion. Finally, Section 5 summarizes the study.

2 Methodology

2.1 ESMD

ESMD, proposed by Wang and Li (2013), is a new adaptive data processing method and can be used to analyze a non-stationary and nonlinear signal. ESMD uses the internal extreme-point symmetry interpolation in place of external envelope interpolation, and optimizes the residual mode using least square approach, which overcomes the shortcomings of modal aliasing and screening termination in EMD. The detailed steps of ESMD are as follows:

(1) Find all the poles of series Y and record them as $x_i (i = 1, 2, \cdots, n)$.

(2) Link the adjacent x_i with lines and mark the midpoints by $B_i (i = 1, 2, \cdots, n-1)$.

(3) Add the boundary midpoints F_0 and F_n using direct interpolation.

(4) Construct p bar differential curves, $L_p(p = 1, 2, \cdots, n)$, using the obtained midpoints, and compute the mean curve by $\bar{L} = (L_1 + \cdots L_p)/p$.

(5) Repeat steps 1-4 on $Y - \bar{L}$ until $|\bar{L}| \leq \theta$ (θ denotes the permitted error), and then the first mode $M_1$is obtained.

(6) Repeat steps 1 to 5 on $Y - M_1$ to obtain $M_2, \cdots,$ M_n and a residual (R) until R only has a certain number of poles.

(7) Change K within the interval $[K_{min}, K_{max}]$ and repeat steps 1 to 6, then compute the standard variance σ of $Y - \mathrm{R}$。.

(8) Select K_0 corresponding to a minimum σ, then K_0 is adopted to repeat steps 1 to 6 and output the final decomposition results.

After decomposition, the original series can generate a series of intrinsic mode functions (IMF) and a residual (R).

2.2 Sample entropy

Sample entropy (SE), proposed by Alcaraz and Rieta (2010), is a novel approach to describe the complexity of series. It has been used in many fields, such as the battery health monitoring (Widodo et al. 2011), wind speed forecasting (Liu et al. 2018a) and electroencephalography (Tsai et al. 2012). The computation steps are as follows:

(1) Recombine X=(x(1), x(2),⋯x(n)) into a matrix:

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

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

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

(2)

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

(3) For $x(i)$, r is the threshold, compute the number meeting the threshold d[$x(i), x(j)$]$\le r$ as B_i. Then, compute the 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) The 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 WD, except that the former extends the abilities of the later (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 mother wavelet function. The decomposition levels and mother wavelet function have a deep influence on the performance of WPD. WPD includes DWT (discrete wavelet transform) and CWT (continuous wavelet transform). CWT is as follows:

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

where $x(t)$ is the input, $*$ is the complex conjugate, b is the translation parameter, a is the scale parameter, b is the translation parameter, and $\psi(t)$ is the 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.

Insert Fig. 1

2.4 LSTM

The LSTM model, proposed by Hochreiter and Schmidhuber (1997), is improved from RNN and 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 can be adopted to memorize the temporal state. Each memory cell encompasses three gates, namely input, output, and forget gates. These gates conduct as filters in playing different roles, solving the exploding, and vanishing gradient problem of RNN. The input gate determines what input singles are going to be accumulated to the cell. The forget gate decides whether previous cell state needs to be remembered or forgotten. The output gate determines which information of the cell state can be output. The framework of LSTM is shown in Fig. 2.

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

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

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

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

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

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

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

$$y_t = \sigma(W_{yx} \cdot x_t + W_{yh} \cdot h_{t-1} + b_y) \tag{15}$$

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

where x_t is the input, y_t is the output, f_t is the forget gate, o_t is the output gate, i_t is the input gate, c_t is the activation vectors, *W* is the weight matrices, *b* is the bias vectors, σ_x is the nonlinear activation function, *c* is the cell state vectors, and h_t is the hidden state.

Insert Fig. 2.

2.5 Model construction

The basic framework of the runoff forecasting system proposed in this study is shown in Fig. 3. The modeling process is illustrated in the following steps:

Step 1: ESMD. ESMD is adopted to split the original 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 mitigate non-stationarity and non-linearity of annual runoff series.

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

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

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

Step 6: Parameter setting. Set values of basic model parameters, such as the number of hidden layers in ANN, ANFIS, and LSTM.

Step 7: 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 study are from five areas in China, namely Mopanshan reservoir, Dahuofang reservoir, Biliuhe reservoir, Hongjiadu reservoir and Changshui hydrological station. Mopanshan Reservoir is located in Wuchang City, 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 the middle reaches of Biliuhe River in Liaoning Province of China. The drainage area is 2085 km^2, and the average annual precipitation is 742.8 mm. Hongjiadu hydropower station is located on the main stream of Wujiang River in the 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 Luoning County, 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. A total of 5 groups of data are shown in Fig 4, and their statistical descriptions are listed in Table 1.

The results of the models are evaluated based on four numerical indicators. These indexes include coefficient of correlation (R), mean absolute percentage error (MAPE), Nash-Sutcliffe efficiency coefficient (NSEC), and root mean square errors (RMSE). Their equations are provided below.

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

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

$$NSEC=1-\frac{\sum_{i=1}^{n}(y_e(i)-y_o(i))^2}{\sum_{i=1}^{n}(y_o(i)-\overline{y_o})^2} \tag{20}$$

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

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

Insert Fig. 3.

Insert Fig. 4.

Insert Table 1

4 Case study

4.1 Series decomposition

The first stage of the runoff prediction framework presented in this study is to decompose the original data using ESMD. Before decomposing the runoff series, the best screening number should be determined by repeating tests and comparisons. In this article, the number of iterations is 100, and the numbers of remaining extreme points of the five runoff datasets are 5, 7, 5, 6 and 7, respectively. The results at Site 1 after the decomposition are shown in Fig 5-7, whilst the decomposition results at all other sites are not presented here. One can see that each IMF split by ESMD is independent, the fluctuation of sub-series from IMF1 to R steadily decreases and the stability gradually becomes stronger. That is to say, IMFs are steadier than the original runoff series and hence are more conducive to capture signal features and predict non-stationary and non-linear sequences.

Insert Fig. 5.

Insert Fig. 6.

Insert Fig. 7.

4.2 Sample entropy computation and two-phase decomposition

SE of each subsequence obtained in the previous step is computed. Then we adopt three decomposition methods to further decompose IMF with the maximum SE. As shown in Fig. 8, SE of all sub-series present a similar trend, and one can clearly see that SE of IMF1 in each dataset is higher than other subseries, which means that IMF1 is more difficult to analyze and predict. To mitigate the high complexity of IMF1, we use three kinds of decomposition algorithm, namely WPD, CEEMDAN and SSA, to decompose IMF1.

Insert Fig. 8.

The selection of the 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. Therefore, we adopt a three-scale and fourth order Symlet wavelet as the wavelet basis function of WPD in this study.

SSA is a traditional and powerful non-parametric decomposition algorithm for signal identification and analysis, which can capture noise component, trend, periodic, oscillatory, quasi-periodic from an input signal (Dong et al. 2017). SSA can decompose a time series into some decipherable and simpler components. SSA contains two steps, namely decomposition and reconstruction. Reconstruction comprises diagonal averaging and grouping, and decomposition incorporates singular value decomposition (SVD) and embedding. In SSA, the window size (L) and eigenvalue grouping (EVG) are key parameters. Before decomposing the IMF1, the best L and EVG value should be determined by repeating tests and comparisons. In this study, the number of L and EVG are set to 12 and 6, respectively.

CEEMDAN (complete EEMD with adaptive noise), proposed by Colominas et al. (2012), is an improvement progress on EEMD and can attain better separation and accurately reconstruct the raw signal. CEEMDAN obtains the modes by adding the white Gaussian noise and computing the unique residue to reduce the EEMD deficiency.

This method can overcome the mode mixing problem, since the procedure of CEEMDAN in decomposition and reconstruction are complete. In the application of CEEMDAN, too many modes may cause extra computational costs and complex training process. Hence, we determine to reconstruct five IMFs and a residual.

Then all IMF1 are split by WPD, SSA and CEEMDAN. The re-decomposition results of IMF1 obtained at Site 1 are shown in Fig 9, where y is the reconstructed time series by SSA, whilst the results at other sites are not presented here. One can see that IMF1 with the maximum SE is decomposed into 16 subsequences with more regular fluctuation by three methods, which means that this procedure may further improve the forecasting accuracy.

Insert Fig 9.

4.3 Number of input variables

The determination of input variables is an important procedure for the final prediction results. In this paper, two methods are utilized to select the 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 the information of PACF and trial-and-error method, while input variables and PACF value for other sites are not shown here.

Insert Table 2.

4.4 Model development

To verify the proposed model, seven models, that is, LSTM, ANN, ANFIS, ESMD-LSTM, ESMD-SE-WPD-LSTM, ESMD-SE-SSA-LSTM, ESMD-SE-

CEEMDAN-LSTM, are employed for comparison. Detailed information relating to these models are presented in the following section.

(1) ANN

ANNs are capable of providing promising results in different aspects of hydrological modeling such as groundwater level modeling (Iqbal et al. 2020; Mirarabi et al. 2019), climate change study (Sabbaghi et al. 2020), rainfall forecasting (Liu et al. 2019), wind speed forecasting (Liu et al. 2018c), streamflow prediction (Ba et al. 2018), etc. Qiu et al. (2020) developed a novel hydrological implementation of emotional ANN model for daily rainfall-runoff modeling. Li et al. (2019) studied the performance of several preprocessing techniques based on ANN for long-term streamflow forecasting. ANNs have been demonstrated to provide fruitful results in hydrology field. In this paper, the standard three-layer feed forward ANN is adopted for annual runoff prediction. The number of input and output layer nodes are equal to the number of input variables and one, respectively. Levenberg-Marquardt (LM) method, sigmoid function and perelin formula are adopted as the training function, transfer function and output function, respectively. The best number of hidden nodes is determined as eight by trial-and-error method, and the training epochs are determined as 500.

(2) ANFIS

ANFIS, proposed by Jang (1993), is capable of identifying nonlinearity and uncertainty between variables and has become popular in parameter estimation and forecasting. Awan and Bae (2014) evaluated the applicability of classified precipitation prediction using ANFIS model to improve the performance of monthly dam inflow forecasting. Bartoletti et al. (2018) presented a simple and effective streamflow forecasting method by integrating ANFIS and principal component analysis techniques. Zhou et al. (2019b) explored a recurrent ANFIS embedded with least square estimator for modelling multi-step-ahead forecasts. Three methods, namely genfis1, genfis2 and genfis3, are available to initialize the data structure of ANFIS. Of the three methods, genfis3 provides the most robust fuzzy inference system 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.

Insert Table 3.

(3) LSTM

The selection of hyper-parameters is a difficult task for LSTM model construction. Adaptive moment estimation can be employed to optimize the parameters. The model structure of LSTM, i.e., hyper-parameters and the number of hidden units are determined by trial-and-error method. In this article, the number of hidden units is 100. The maximum number of epochs is 2000. The size of the mini-batch used for each training iteration is 100. The initial learning rate is adjusted as 0.01. Other parameters are determined to default values used by adaptive moment estimation. Finally, RMSE is adopted as the loss function.

(4) ESMD-LSTM model

For ESMO-LSTM model, the observed 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

For ESMD-SE-WPD-LSTM, ESMD-SE-SSA-LSTM and ESMD-SE-CEEMDAN-LSTM, ESMD is adopted to decompose the raw runoff series into a series of IMF and a Res. Then SE method is employed to measure the complexity of each composition. Thirdly, we adopt three decomposition algorithms, namely WPD, SSA and CEEMDAN, to further decompose IMF with the maximum SE. Then, LSTM model is employed to predict each subseries obtained in the previous step.

4.5 Results and discussion

In this section, forecasting results of seven methods for five experiments are presented. Tables 4-8 show error estimation results of different methods for five runoff time series. Table 9 presents forecasting results of ESMD-LSTM at different sites. Figures 10-14 present forecasting results of five stations. The following should be noted before analyzing the results. The forecasting results of the testing phase plays 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.

4.5.1. Experiment 1: Comparison of several single prediction models

In this section, we analyze prediction results of three single models at five sites. As seen from Table 4, when forecasting the annual runoff in Site 1, LSTM provides the best prediction effect during the testing phase, with the minimum error indexes (MAPE and RMSE) and the maximum fitting indexes (NSEC and R), which are respectively

18.1835%, 115.7743, -2.7495 and 0.7046. Similarly, for Sites 2-5, it can be clearly seen that LSTM outperforms other two single models. In addition, the 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 five datasets in terms of four considered 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 the method combined with one-phase decomposition

This section compares the performance between single LSTM model and ESMD-LSTM hybrid model. Taking Site 1 as an example, the one-phase decomposition method significantly improves the forecasting accuracy of single LSTM model. In the testing period, ESMD-LSTM outperforms LSTM by 20.14% and 115.65% improvement in R and NSEC, respectively, and 61.02% and 67.92% reduction in RMSE and MAPE, respectively. For Site 2, in the testing phase, ESMD-LSTM outperforms LSTM by 20.97% and 99.63% improvement in R and NSEC, respectively, and 44.46% and 30.83% reduction in RMSE and MAPE, respectively. According to the values in Tables 6-8, we can reaffirm ESMD-LSTM model is able to provide better results than LSTM model with substantial improvement in terms of four considered indexes. Table 9 lists the prediction results obtained by ESMD-LSTM for five datasets. One can clearly see that the forecasting results of IMF1 are inferior to those of the other subseries. For IMF1, in the testing phase, R at Site 1, Site 2, Site 3, Site 4 and Site 5 are 0.8718, 0.9175, 0.5734, 0.5044 and 0.8488, respectively, with an average value of 0.7432. However, for IMF2, in the testing phase, R at Site 1, Site 2, Site 3, Site 4 and Site 5 are 9864, 0.8732, 0.8098, 0.9602 and 0.8428, respectively, with an average value of 0.8945, and the average values of IMF3, IMF4, IMF5 and Res are 0.9682, 0.9990, 0.9994 and 0.9998, respectively.

In general, this analysis illustrates that ESMD is suitable to decompose the annual runoff series, and can significantly improve forecasting accuracy. In addition, we can confirm from Table 9 that results of IMF1 are inferior to those of other subsequence in terms of RMSE, NSEC and MAPE. These analyses also illustrate that single decomposition may be difficult to fully mitigate sequence nonlinearity. Then, we attempt to use the secondary decomposition method to further mitigate nonlinearity of sub-series and solve the limitation of 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, and ESMD-SE-CEEMDAN-LSTM, and ESMD-LSTM are treated as the benchmark methods. Tables 4-8 list composite results at Sites 1-5. One can see that when forecasting annual runoff at Site 1, the proposed ESMD-SE-WPD-LSTM exhibits the best results in terms of all indexes. For Site 1, in the testing phase, R of ESMD-LSTM, ESMD-SE-SSA-LSTM, ESMD-SE-CEEMDAN-LSTM, and ESMD-SE-WPD-LSTM are 0.8465, 0.8521, 0.8796, and 0.9163, respectively. RMSE of ESMD-LSTM, ESMD-SE-SSA-LSTM, ESMD-SE-

CEEMDAN-LSTM, and ESMD-SE-WPD-LSTM are 34.1646, 48.0671 and 28.102, respectively. It is suggested that three two-phase decomposition hybrid methods, namely ESMD-SE-SSA-LSTM, ESMD-SE-CEEMDAN-LSTM and ESMD-SE-WPD-LSTM, outperform ESMD-LSTM model with 0.66%, 3.91% and 8.25% improvements in R, respectively, and 24.30%, -6.50% and 37.73% reduction in RMSE, respectively. Compared with ESMD-LSTM, MAPE of three two-phase decomposition hybrid methods are reduced by 19.38%, -9.02% and 33.11%, respectively, while NSEC of three two-phase decomposition hybrid methods are improved by 56.56%, -17.78% and 81.10%, respectively. For Site 2, in the testing phase, R of ESMD-LSTM, ESMD-SE-SSA-LSTM, ESMD-SE-CEEMDAN-LSTM, and ESMD-SE-WPD-LSTM are 0.9166, 0.9386, 0.9397, and 0.9817, respectively. Compared with ESMD-LSTM, R of three secondary decomposition hybrid methods are improved by 2.40%, 2.52% and 7.10%, respectively; RMSE of three secondary decomposition hybrid methods are reduced by 13.81%, 2.18% and 43.44%, respectively; NSEC of three secondary decomposition hybrid methods are improved by 5.72%, 0.95% and 15.14%, respectively. From the results of these analysis, we can draw the following conclusions:

(1) It is confirmed that IMF1 is highly nonlinear and difficult to forecast, which will affect the overall prediction accuracy of the model.

(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 can exhibit the best performance for all forecasting sites. From

Tables 4-8, it can be seen that the prediction performance of ESMD-SE-SSA-LSTM is not stable, and the forecasting accuracy of ESMD-SE-CEENDAN-LSTM is slightly inferior to the proposed model. Compared with SSA and CEEMDAN, WPD is more suitable to extract the significant features of IMF1. ESMD-SE-WPD-LSTM outperforms all comparing methods. The possible reason is 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, which provide more distinctive features and better forecasting accuracy.

4.5.4 Comparison of all involved models

The performances of all comparing models developed in this study are shown in Figures 10-14. From the figures, one can clearly see that the forecasting performance of six models (except ANN) in the training phase are slightly overestimated. Meanwhile, in the testing phase, it is obvious that the forecasting accuracy of all sites can be significantly improved, and the performances of different models are uneven. The proposed two-phase decomposition hybrid prediction model provides the best performance as the trend line is very close to the original data line, and the method can capture abrupt changes in annual runoff series.

Insert Table 4.

Insert Table 5.

Insert Table 6.

Insert Table 7.

Insert Table 8.

Insert Table 9.

Insert Fig. 10.

Insert Fig 11.

Insert Fig 12.

Insert Fig. 13.

Insert Fig14.

5 Conclusion

Long-term runoff forecasting plays a critical role in the management and monitoring of water resources. To attain a 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 compute SE (sample entropy) of all sub-series. Secondly, the sub-series with the maximum SE is adopted for secondary decomposition using WPD, which can fully mitigate non-linearity of the sub-series. Next, we employ LSTM 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. Forecasting errors of all compared models are evaluated based on four numerical indicators. According to the forecasting results and corresponding analysis, the following conclusions can be drawn:

First, the proposed hybrid model provides the most robust performance and excellent forecasting accuracy among all compared models. This demonstrates that the proposed model can significantly improve the prediction accuracy of long-term runoff time series.

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

Third, EMSD and WPD, as two signal processing methods with high efficiency, can complement each other. After screening by sample entropy, the original single 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 study combines data preprocessing technology, sample entropy, forecasting model, and error analysis to develop runoff forecasting model, which is more conducive to be a useful and efficient soft computing model to forecast runoff time series.

Acknowledgements

The authors are grateful to acknowledge the funding support of Project of key science and technology of the Henan province (No: 202102310259; No: 202102310588), Henan province university scientific and technological innovation

team (No: 18IRTSTHN009).

Authors' Contributions

Wen-chuan Wang: Conceptualization, Methodology, Writing-original draft. **Yu-jin Du:** Methodology, data curation, Writing - original draft preparation. **Kwok-wing Chau**: Writing and editing-original draft. **Dong-mei Xu**: Formal analysis and data collection. **Chang-jun Liu:** Formal analysis. **Qiang Ma:** Investigation.

Availability of data and materials

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

Ethics 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.

Conflict 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 Journal of Geophysical Research: Atmospheres 123:12,543-512,563 doi:https://doi.org/10.1029/2018JD028375

Alcaraz R, Rieta JJ (2010) A review on sample entropy applications for the non-invasive analysis of atrial fibrillation electrocardiograms Biomedical Signal Processing and Control 5:1-14 doi: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 Biomedical Signal Processing and Control 39:94-102 doi:https://doi.org/10.1016/j.bspc.2017.07.022

Aqil M, Kita I, Yano A, Nishiyama S (2007) A comparative study of artificial neural networks and neuro-fuzzy in continuous modeling of the daily and hourly behaviour of runoff Journal of Hydrology

337:22-34 doi:https://doi.org/10.1016/j.jhydrol.2007.01.013

Awan JA, Bae DH (2014) Improving ANFIS Based Model for Long-term Dam Inflow Prediction by Incorporating Monthly Rainfall Forecasts Water Resources Management 28:1185-1199 doi:10.1007/s11269-014-0512-7

Ba HH, Guo SL, Wang Y, Hong XJ, Zhong YX, Liu ZJ (2018) Improving ANN model performance in runoff forecasting by adding soil moisture input and using data preprocessing techniques Hydrology Research 49:744-760 doi:10.2166/nh.2017.048

Bartoletti N, Casagli F, Marsili-Libelli S, Nardi A, Palandri L (2018) Data-driven rainfall/runoff modelling based on a neuro-fuzzy inference system Environmental Modelling & Software 106:35-47 doi:10.1016/j.envsoft.2017.11.026

Bojang PO, Yang TC, Pham QB, Yu PS (2020) Linking Singular Spectrum Analysis and Machine Learning for Monthly Rainfall Forecasting Applied Sciences-Basel 10 doi:10.3390/app10093224

Chau KW, Wu CL, Li YS (2005) Comparison of Several Flood Forecasting Models in Yangtze River Journal of Hydrologic Engineering 10:485-491 doi:10.1061/(ASCE)1084-0699(2005)10:6(485)

Colominas MA, Schlotthauer G, Torres ME, Flandrin P (2012) NOISE-ASSISTED EMD METHODS IN ACTION Advances in Adaptive Data Analysis 04:1250025 doi: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 Renewable Energy 102:241-257 doi:https://doi.org/10.1016/j.renene.2016.10.030

Ehteram M et al. (2019) Assessing the Predictability of an Improved ANFIS Model for Monthly Streamflow Using Lagged Climate Indices as Predictors Water 11 doi:10.3390/w11061130

Evsukoff AG, Cataldi M, de Lima BSLP (2012) A multi-model approach for long-term runoff modeling using rainfall forecasts Expert Systems with Applications 39:4938-4946 doi:https://doi.org/10.1016/j.eswa.2011.10.023

Feng Q, Wen X, Li J (2015) Wavelet Analysis-Support Vector Machine Coupled Models for Monthly Rainfall Forecasting in Arid Regions Water Resources Management 29:1049-1065 doi:10.1007/s11269-014-0860-3

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

Feng Z-k, Niu W-j, Cheng X, Wang J-y, Wang S, Song Z-g (2020b) An effective three-stage hybrid optimization method for source-network-load power generation of cascade hydropower reservoirs serving multiple interconnected power grids Journal of Cleaner Production 246 doi:10.1016/j.jclepro.2019.119035

Feng Z-k, Niu W-j, Tang Z-y, Jiang Z-q, Xu Y, Liu Y, Zhang H-r (2020c) Monthly runoff time series prediction by variational mode decomposition and support vector machine based on quantum-behaved particle swarm optimization Journal of Hydrology 583 doi:10.1016/j.jhydrol.2020.124627

Gao S, Huang Y, Zhang S, Han J, Wang G, Zhang M, Lin Q (2020) Short-term runoff prediction with GRU and LSTM networks without requiring time step optimization during sample generation Journal of Hydrology 589 doi:10.1016/j.jhydrol.2020.125188

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 Resources Management 34:865-884 doi:10.1007/s11269-020-02483-x

He Z, Wen X, Liu H, Du J (2014) A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region Journal of Hydrology 509:379-386 doi:https://doi.org/10.1016/j.jhydrol.2013.11.054

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

Hu ZH, Zhang YR, Zhao YC, Xie MS, Zhong JZ, Tu ZG, Liu JT (2019) A Water Quality Prediction Method Based on the Deep LSTM Network Considering Correlation in Smart Mariculture Sensors 19 doi:10.3390/s19061420

Iqbal M, Naeem UA, Ahmad A, Habib ur R, Ghani U, Farid T (2020) Relating groundwater levels with meteorological parameters using ANN technique Measurement 166 doi:10.1016/j.measurement.2020.108163

Jang JR (1993) ANFIS: adaptive-network-based fuzzy inference system IEEE Transactions on Systems, Man, and Cybernetics 23:665-685 doi:10.1109/21.256541

Kisi O, Sanikhani H (2015) Prediction of long-term monthly precipitation using several soft computing methods without climatic data International Journal of Climatology 35:4139-4150 doi:10.1002/joc.4273

Kratzert F, Klotz D, Brenner C, Schulz K, Herrnegger M (2018) Rainfall-runoff modelling using Long Short-Term Memory (LSTM) networks Hydrology and Earth System Sciences 22:6005-6022 doi:10.5194/hess-22-6005-2018

Li FF, Wang ZY, Qiu J (2019) Long-term streamflow forecasting using artificial neural network based on preprocessing technique Journal of Forecasting 38:192-206 doi:10.1002/for.2564

Li G, Zhang S, Yang H (2017) A Deep Learning Prediction Model Based on Extreme-Point Symmetric Mode Decomposition and Cluster Analysis Mathematical Problems in Engineering 2017 doi:10.1155/2017/8513652

Li H, Bai J, Cui X, Li Y, Sun S (2020) A new secondary decomposition-ensemble approach with cuckoo search optimization for air cargo forecasting Applied Soft Computing 90 doi:10.1016/j.asoc.2020.106161

Lin Q, Wu Z, Singh VP, Sadeghi SHR, He H, Lu G (2017) Correlation between hydrological drought, climatic factors, reservoir operation, and vegetation cover in the Xijiang Basin, South China Journal of Hydrology 549:512-524 doi:https://doi.org/10.1016/j.jhydrol.2017.04.020

Liu DR, Jiang WC, Mu L, Wang S (2020) Streamflow Prediction Using Deep Learning Neural Network: Case Study of Yangtze River Ieee Access 8:90069-90086 doi:10.1109/access.2020.2993874

Liu H, Duan Z, Han F-z, Li Y-f (2018a) Big multi-step wind speed forecasting model based on secondary decomposition, ensemble method and error correction algorithm Energy Conversion and Management 156:525-541 doi:10.1016/j.enconman.2017.11.049

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

Liu H, Mi XW, Li YF (2018c) Comparison of two new intelligent wind speed forecasting approaches based on Wavelet Packet Decomposition, Complete Ensemble Empirical Mode Decomposition with Adaptive Noise and Artificial Neural Networks Energy Conversion and Management 155:188-200 doi:10.1016/j.enconman.2017.10.085

Liu Y, Zhao QZ, Yao WQ, Ma XW, Yao YB, Liu LL (2019) Short-term rainfall forecast model based on the improved BP-NN algorithm Scientific Reports 9 doi:10.1038/s41598-019-56452-5

Meng E, Huang S, Huang Q, Fang W, Wu L, Wang L (2019) A robust method for non-stationary streamflow prediction based on improved EMD-SVM model Journal of Hydrology 568:462-478 doi:https://doi.org/10.1016/j.jhydrol.2018.11.015

Mirarabi A, Nassery HR, Nakhaei M, Adamowski J, Akbarzadeh AH, Alijani F (2019) Evaluation of data-driven models (SVR and ANN) for groundwater-level prediction in confined and unconfined systems Environmental Earth Sciences 78 doi:10.1007/s12665-019-8474-y

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 Resources Management 34:4113-4131 doi: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 Resources Management 33:2907-2923 doi:10.1007/s11269-019-02273-0

Qiu RJ, Wang YK, Wang D, Qiu WJ, Wu JC, Tao YW (2020) Water temperature forecasting based on modified artificial neural network methods: Two cases of the Yangtze River Science of the Total Environment 737 doi:10.1016/j.scitotenv.2020.139729

Sabbaghi MA, Nazari M, Araghinejad S, Soufizadeh S (2020) Economic impacts of climate change on water resources and agriculture in Zayandehroud river basin in Iran Agricultural Water Management 241 doi:10.1016/j.agwat.2020.106323

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 doi: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 Geophysica 67:1471-1481 doi: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 Resources Management 30:4011-4035 doi:10.1007/s11269-016-1409-4

Song P et al. (2020) Annual Runoff Forecasting Based on Multi-Model Information Fusion and Residual Error Correction in the Ganjiang River Basin Water 12 doi:10.3390/w12082086

Srinivas AST, Somula R, Govinda K, Saxena A, Reddy PA (2020) Estimating rainfall using machine learning strategies based on weather radar data International Journal of Communication Systems 33 doi:10.1002/dac.3999

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 Transactions of the Institute of Measurement and Control 42:3-21 doi:10.1177/0142331218771141

Sun W, Huang C (2020) A hybrid air pollutant concentration prediction model combining secondary decomposition and sequence reconstruction Environmental Pollution 266 doi:10.1016/j.envpol.2020.115216

Tan QF, Lei XH, Wang X, Wang H, Wen X, Ji Y, Kang AQ (2018) An adaptive middle and long-term runoff forecast model using EEMD-ANN hybrid approach Journal of Hydrology 567:767-780 doi:10.1016/j.jhydrol.2018.01.015

Tao YM, Gao XG, Ihler A, Sorooshian S, Hsu KL (2017) Precipitation Identification with Bispectral Satellite Information Using Deep Learning Approaches Journal of Hydrometeorology 18:1271-1283 doi: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 Meteorology and Atmospheric Physics 131:115-125 doi:10.1007/s00703-017-0546-5

Tsai P-H, Lin C, Tsao J, Lin P-F, Wang P-C, Huang NE, Lo M-T (2012) Empirical mode decomposition based detrended sample entropy in electroencephalography for Alzheimer's disease Journal of Neuroscience Methods 210:230-237 doi:https://doi.org/10.1016/j.jneumeth.2012.07.002

Wang J-L, Li Z-J (2013) EXTREME-POINT SYMMETRIC MODE DECOMPOSITION METHOD FOR DATA ANALYSIS Advances in Adaptive Data Analysis 05:1350015 doi:10.1142/S1793536913500155

Wang T, Zhou Y (2020) Forecasting the Yellow River runoff based on functional data analysis methods Environmental and Ecological Statistics doi:10.1007/s10651-020-00469-x

Wang W-C, Chau K-W, Cheng C-T, Qiu L (2009) A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series Journal of Hydrology 374:294-306 doi:https://doi.org/10.1016/j.jhydrol.2009.06.019

Wang WC, Chau KW, Xu DM, Chen XY (2015) Improving Forecasting Accuracy of Annual Runoff Time Series Using ARIMA Based on EEMD Decomposition Water Resources Management 29:2655-2675 doi:10.1007/s11269-015-0962-6

Widodo A, Shim M-C, Caesarendra W, Yang B-S (2011) Intelligent prognostics for battery health monitoring based on sample entropy Expert Systems with Applications 38:11763-11769 doi:https://doi.org/10.1016/j.eswa.2011.03.063

Xiang ZR, Yan J, Demir I (2020) A Rainfall-Runoff Model With LSTM-Based Sequence-to-Sequence Learning Water Resources Research 56 doi: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 Scientific Reports 9 doi:10.1038/s41598-019-49242-6

Yuan XH, Chen C, Lei XH, Yuan YB, Adnan RM (2018) Monthly runoff forecasting based on LSTM-ALO model Stochastic Environmental Research and Risk Assessment 32:2199-2212 doi:10.1007/s00477-018-1560-y

Zamani Sabzi H, King JP, Abudu S (2017) Developing an intelligent expert system for streamflow prediction, integrated in a dynamic decision support system for managing multiple reservoirs: A case study Expert Systems with Applications 83:145-163 doi:https://doi.org/10.1016/j.eswa.2017.04.039

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 Journal of Hydrology 561:918-929 doi:10.1016/j.jhydrol.2018.04.065

Zhou J, Xu X, Huo X, Li Y (2019a) Forecasting Models for Wind Power Using Extreme-Point Symmetric Mode Decomposition and Artificial Neural Networks Sustainability 11 doi:10.3390/su11030650

Zhou YL, Guo SL, Chang FJ (2019b) Explore an evolutionary recurrent ANFIS for modelling multi-step-ahead flood forecasts Journal of Hydrology 570:343-355 doi:10.1016/j.jhydrol.2018.12.040

Zhu S, Luo XG, Yuan XH, Xu ZY (2020) An improved long short-term memory network for streamflow forecasting in the upper Yangtze River Stochastic Environmental Research and Risk Assessment 34:1313-1329 doi:10.1007/s00477-020-01766-4

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 Journal of Hydrology 585 doi:10.1016/j.jhydrol.2020.124776

Figure captions

Fig. 1. Sketch map of WPD method.

Fig. 2. The basic LSTM architecture.

Fig. 3. Framework of the proposed model.

Fig. 4. Original runoff series.

Fig 5. Decomposition at Site 1 by ESMD.

Fig 6. Frequency distribution of each IMF at Site 1.

Fig. 7. Amplitude of each IMF at Site 1.

Fig. 8. SE of each sequence decomposed by ESMD.

Fig. 9. Decomposition results of IMF1 at Site 1.

Fig. 10. Forecasting results at Site 1.

Fig. 11. Forecasting results at Site 2.

Fig. 12. Forecasting results at Site 3.

Fig. 13. Forecasting results at Site 4.

Fig. 14. Forecasting results at Site 5.

Table captions

Table 1 Statistical description of runoff series at five stations.

Table 2 Input variables for Site 1.

Table 3 Parameters of ANFIS.

Table 4 Errors of different models at Site 1.

Table 5 Errors of different models at Site 2.

Table 6 Errors of different models at Site 3.

Table 7 Errors of different models at Site 4.

Table 8 Errors of different models at Site 5.

Table 9. Errors of ESMD-LSTM at different sites.

Figures

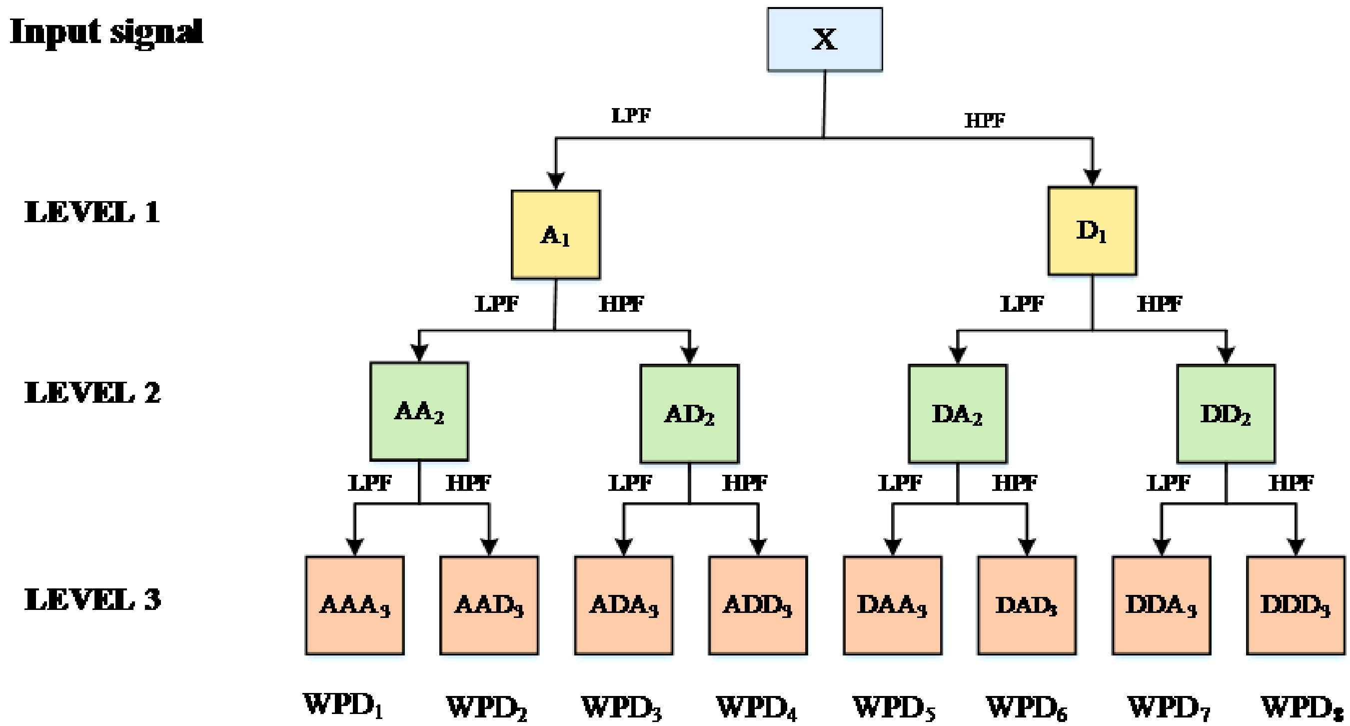

Figure 1

Sketch map of WPD method.

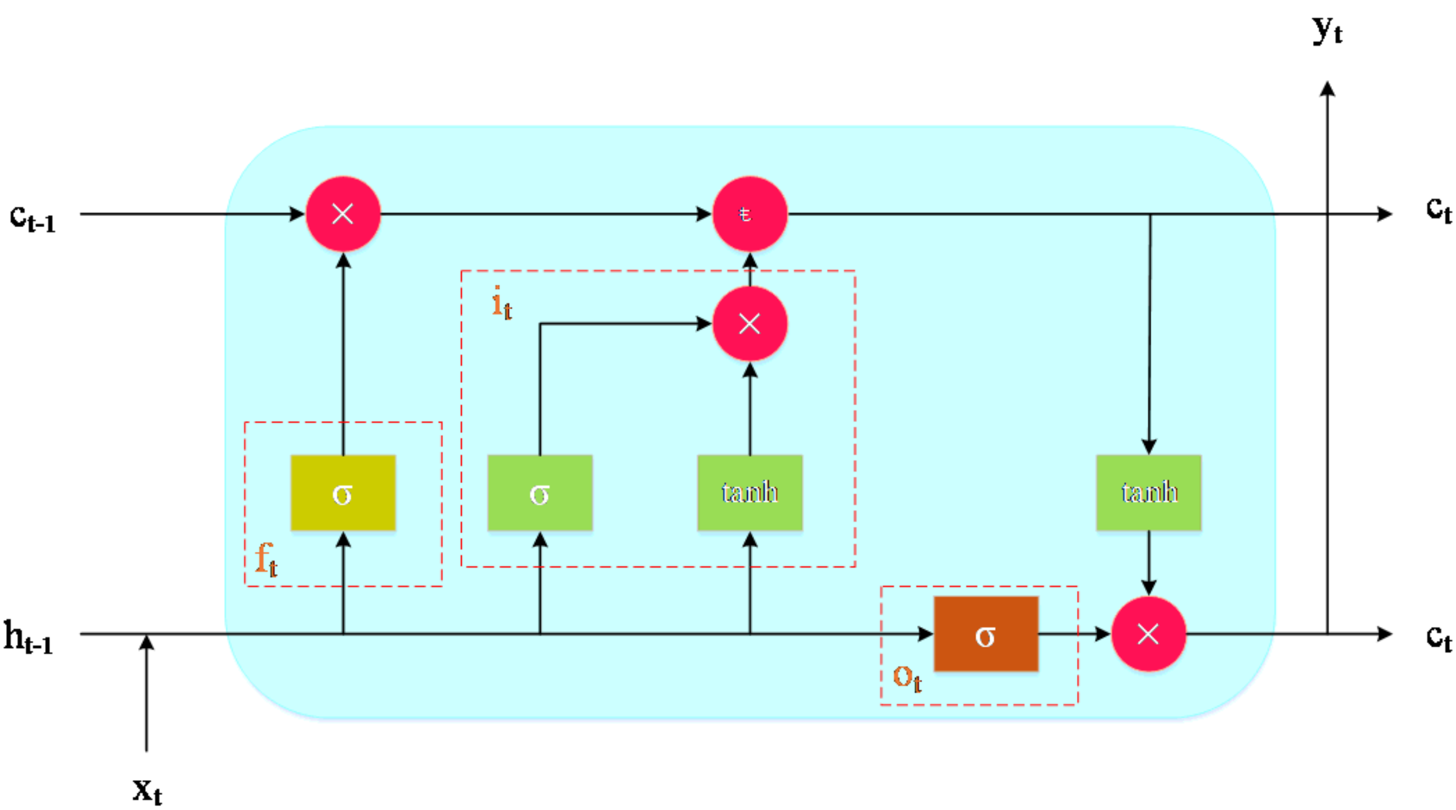

Figure 2

The basic LSTM architecture.

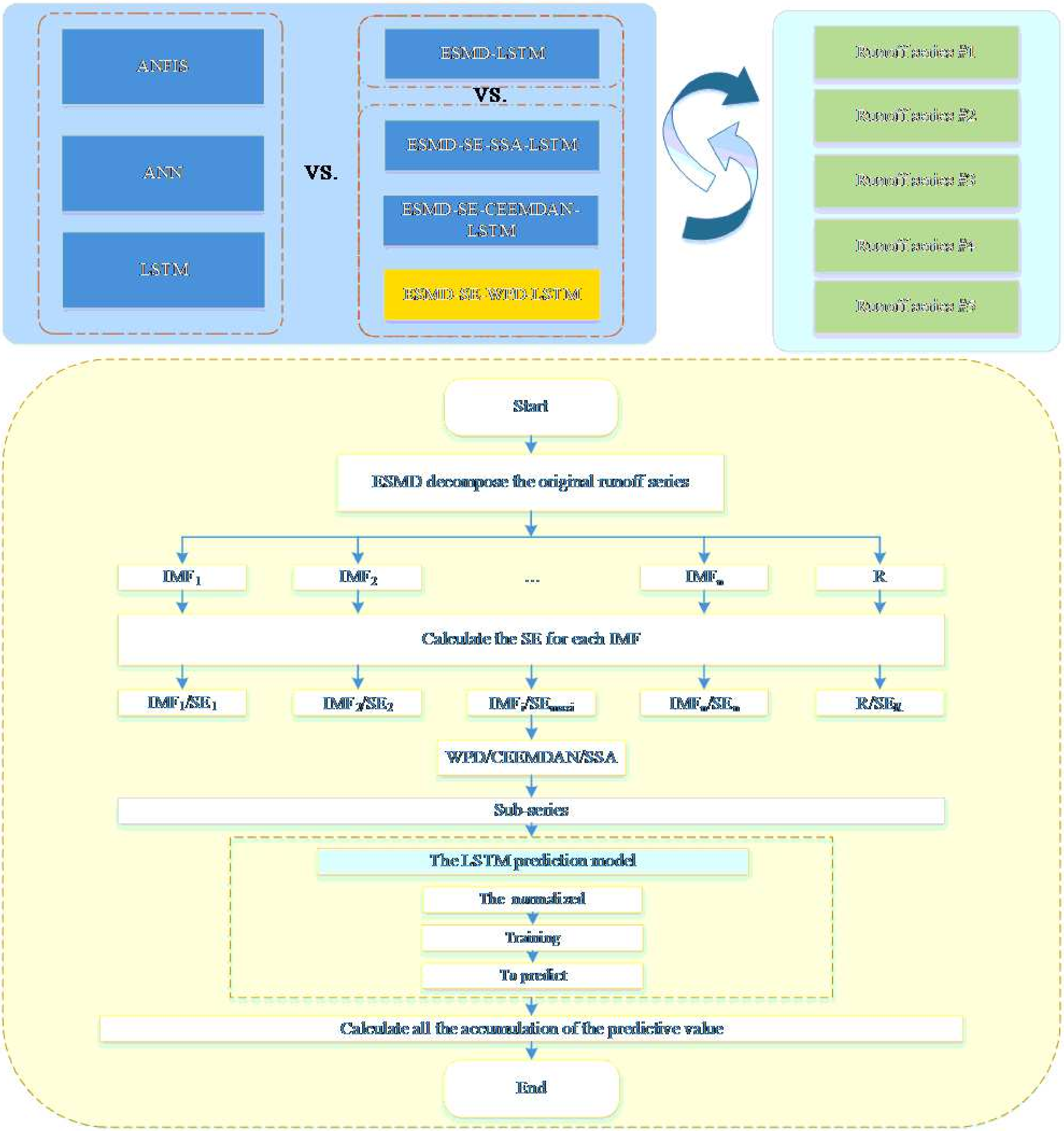

Figure 3

Framework of the proposed model.

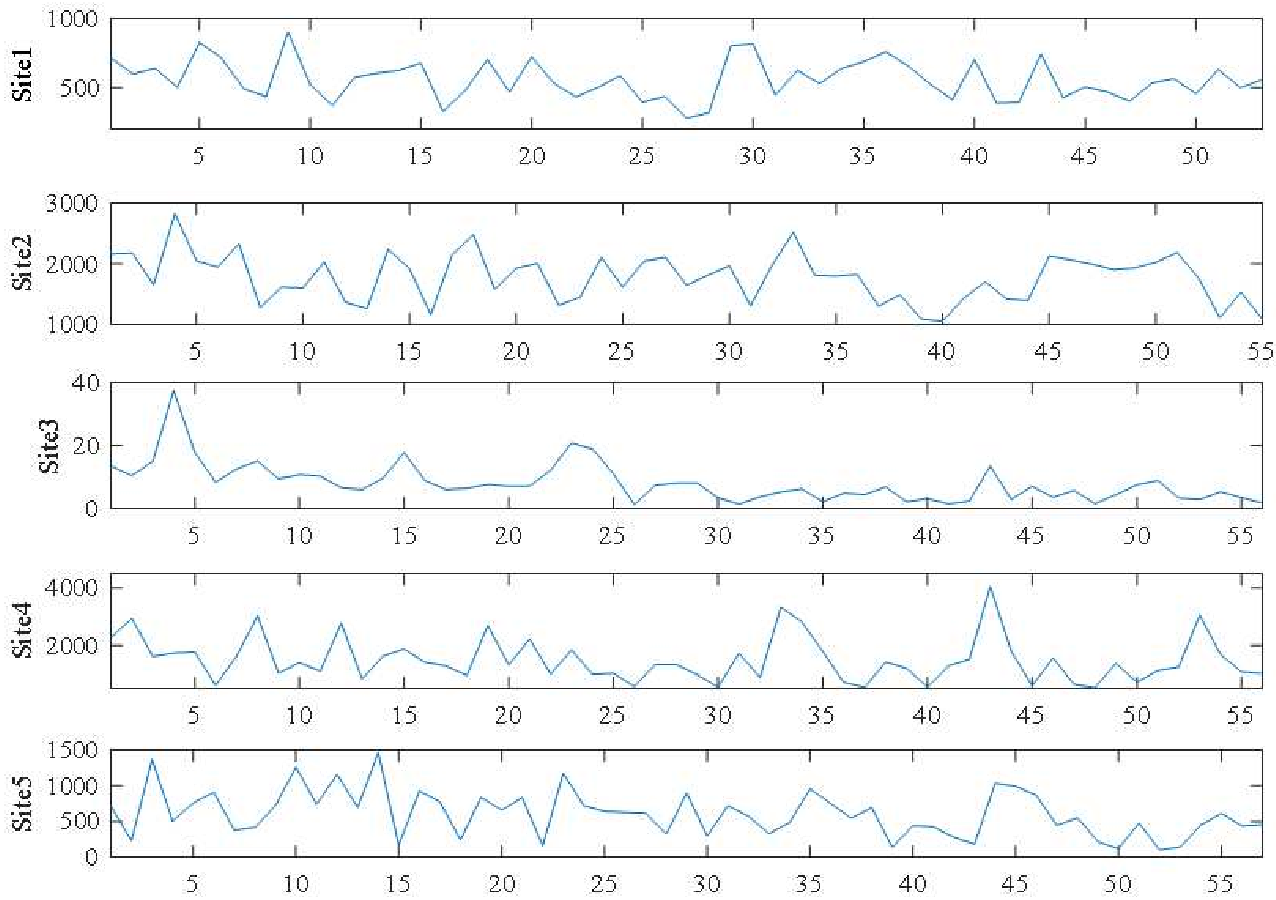

Figure 4

Original runoff series.

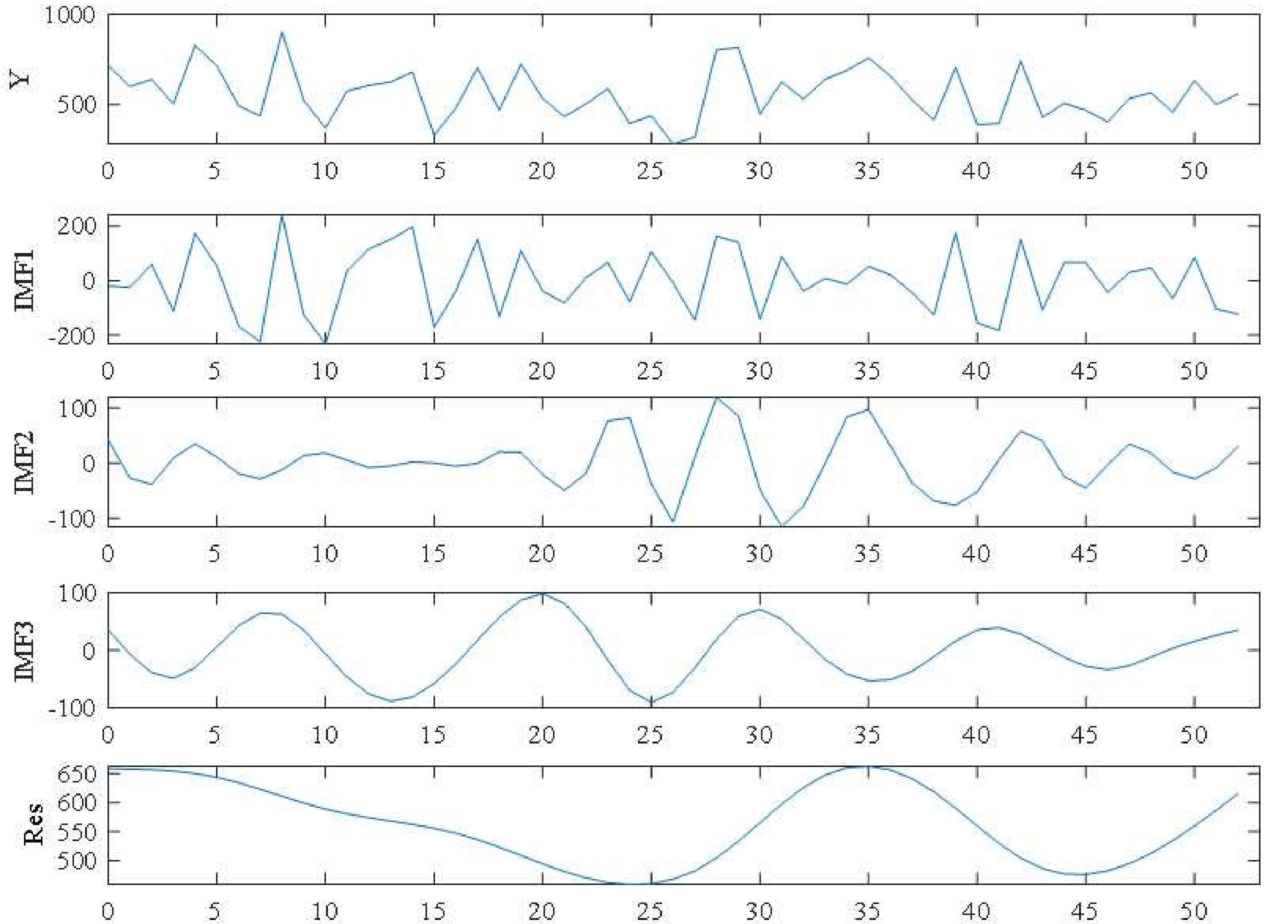

Figure 5

Decomposition at Site 1 by ESMD.

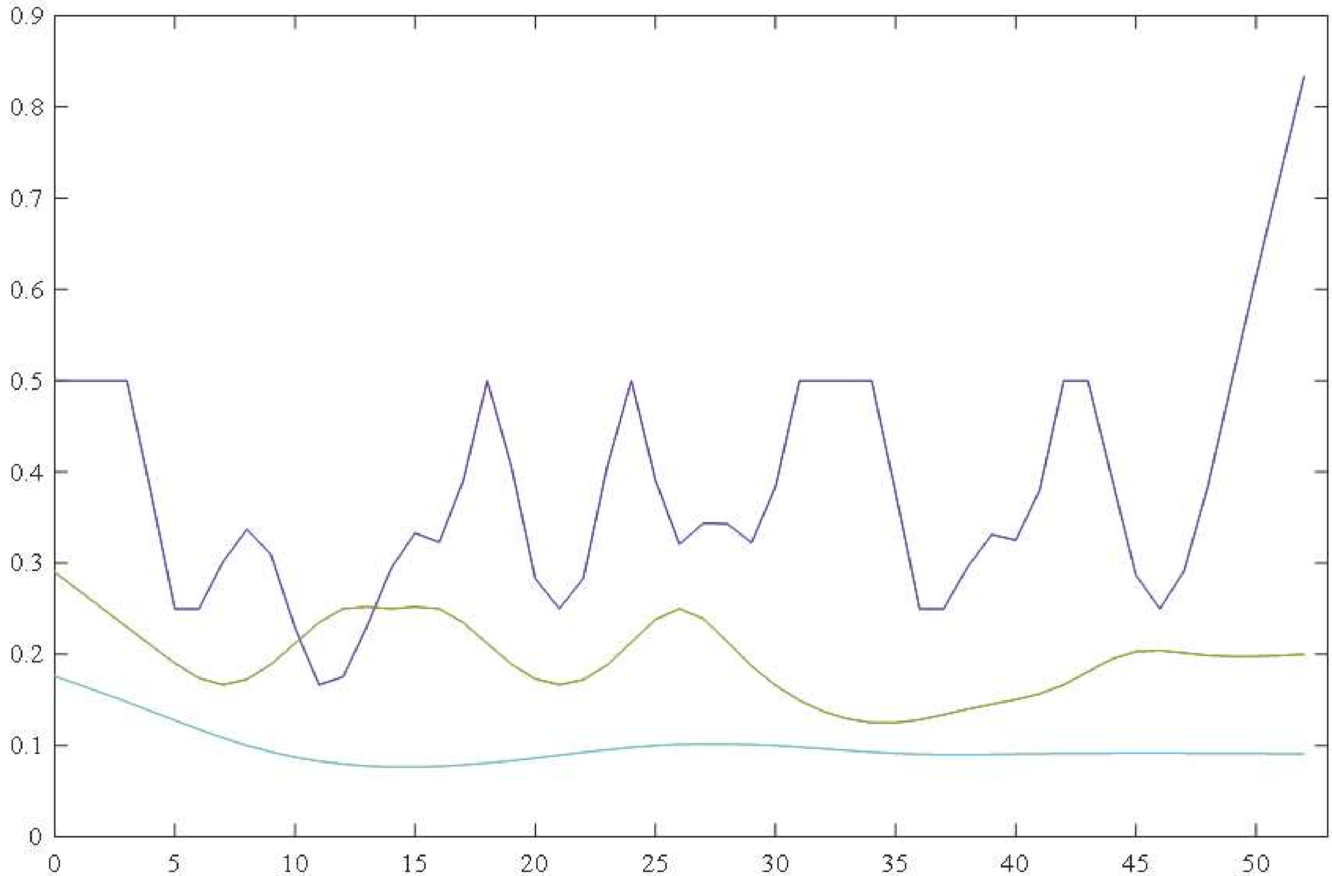

Figure 6

Frequency distribution of each IMF at Site 1.

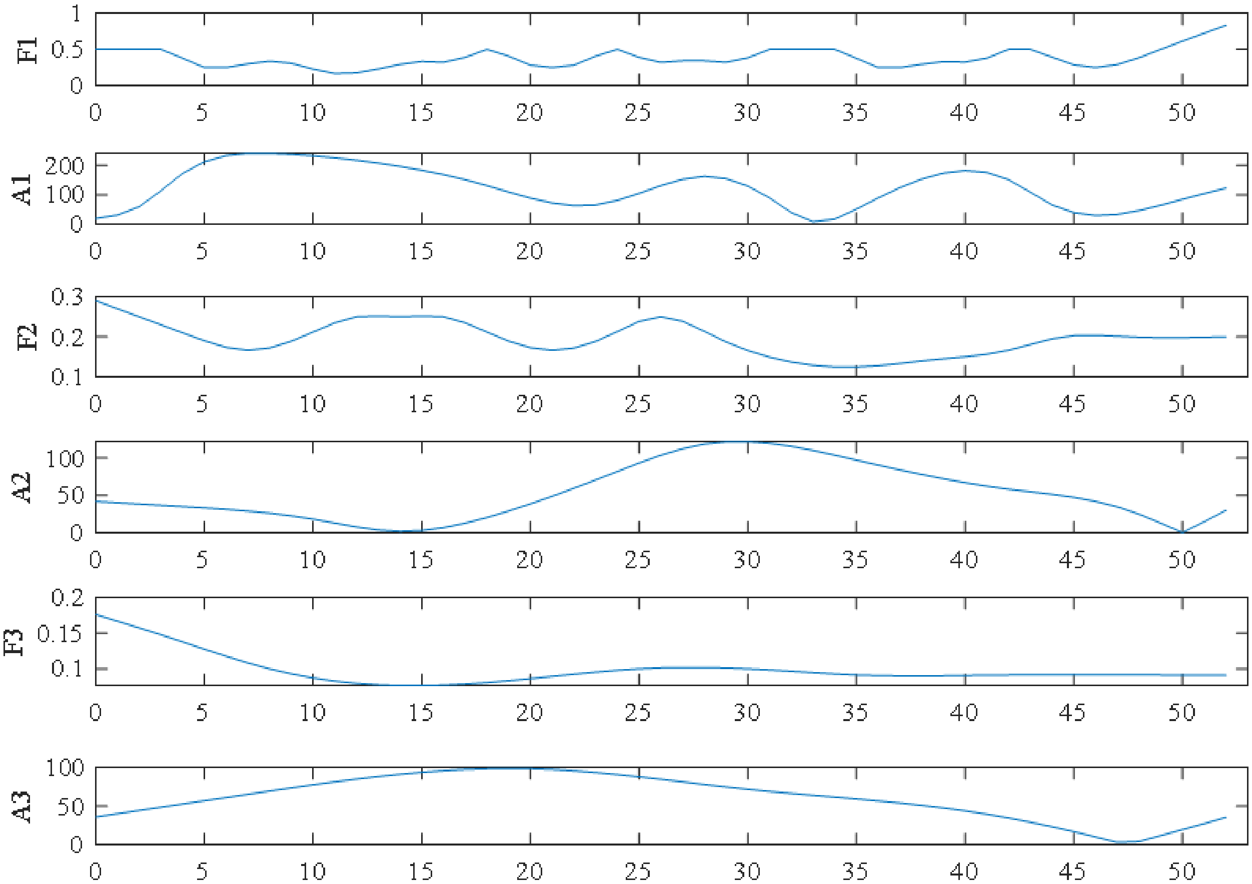

Figure 7

Amplitude of each IMF at Site 1.

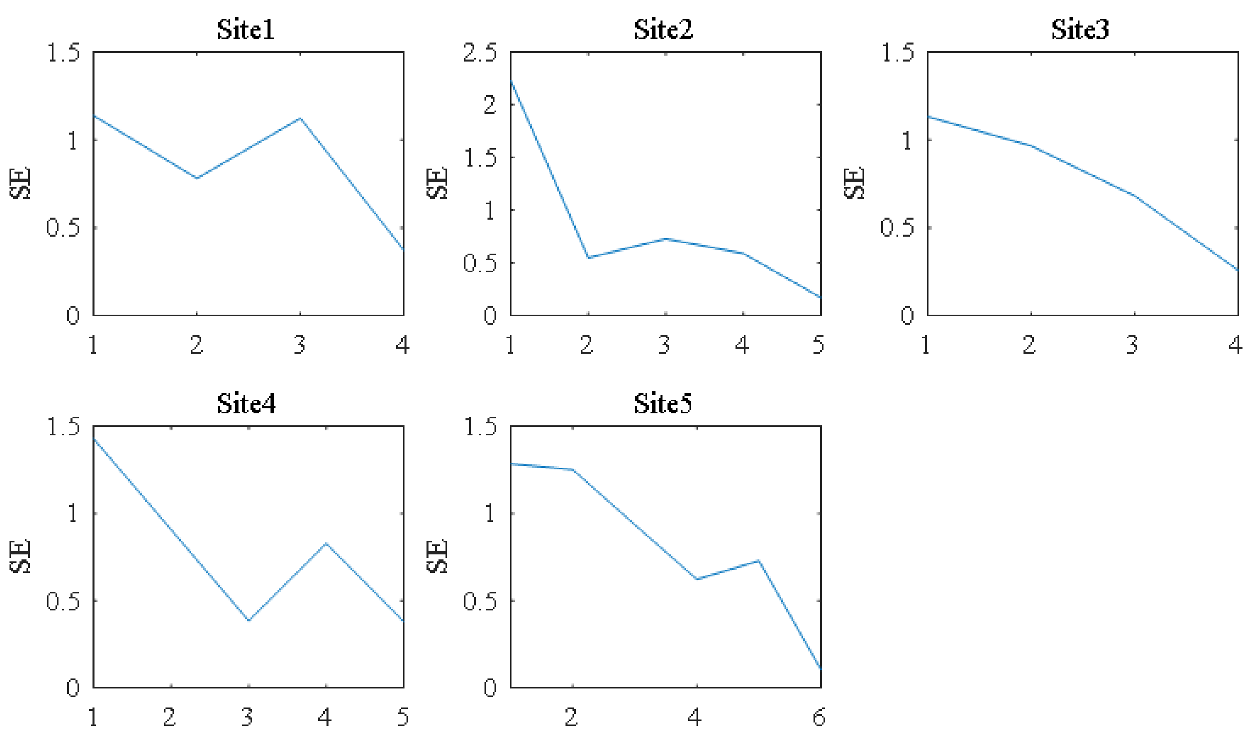

Figure 8

SE of each sequence decomposed by ESMD.

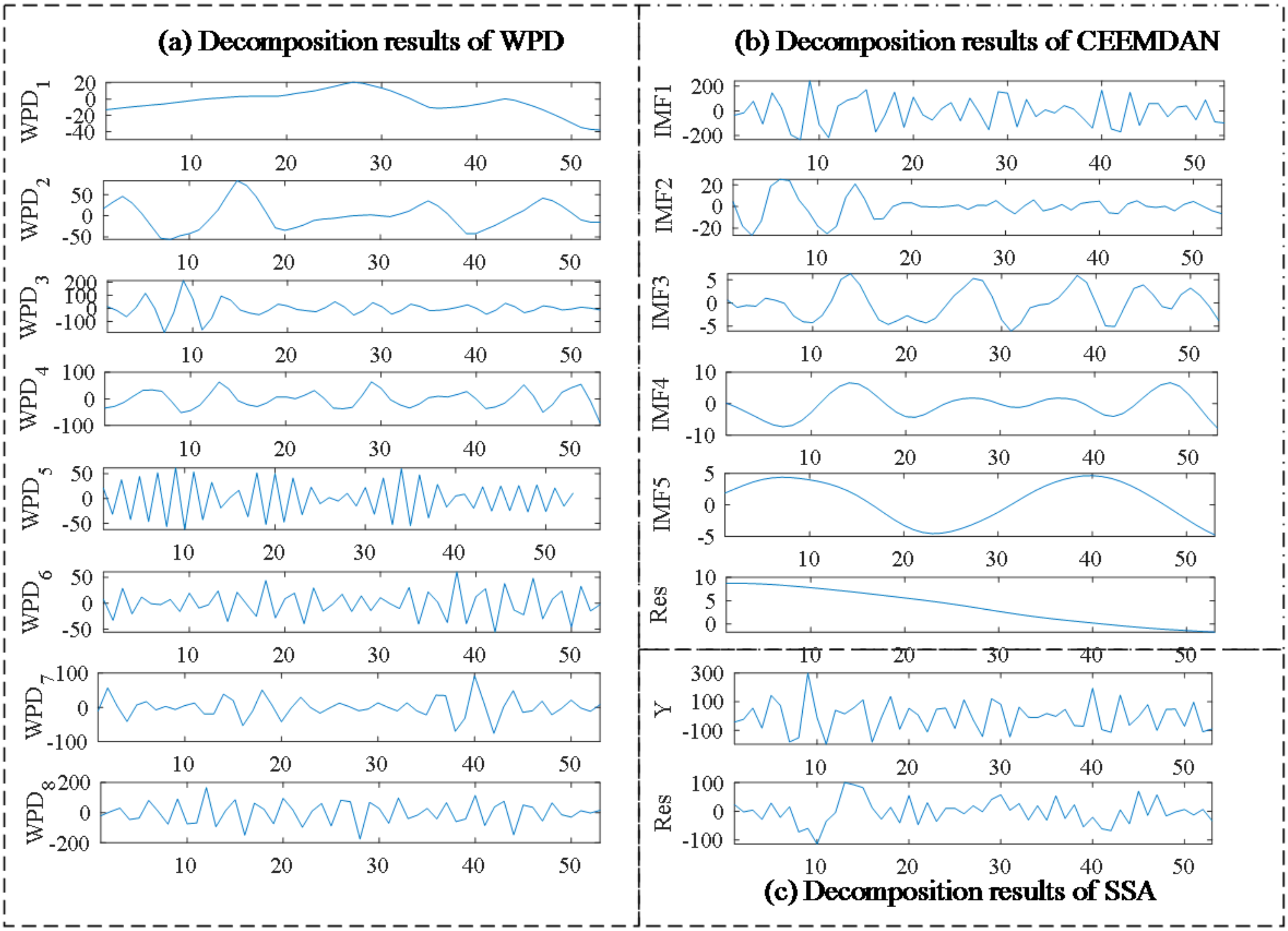

Figure 9

Decomposition results of IMF1 at Site 1.

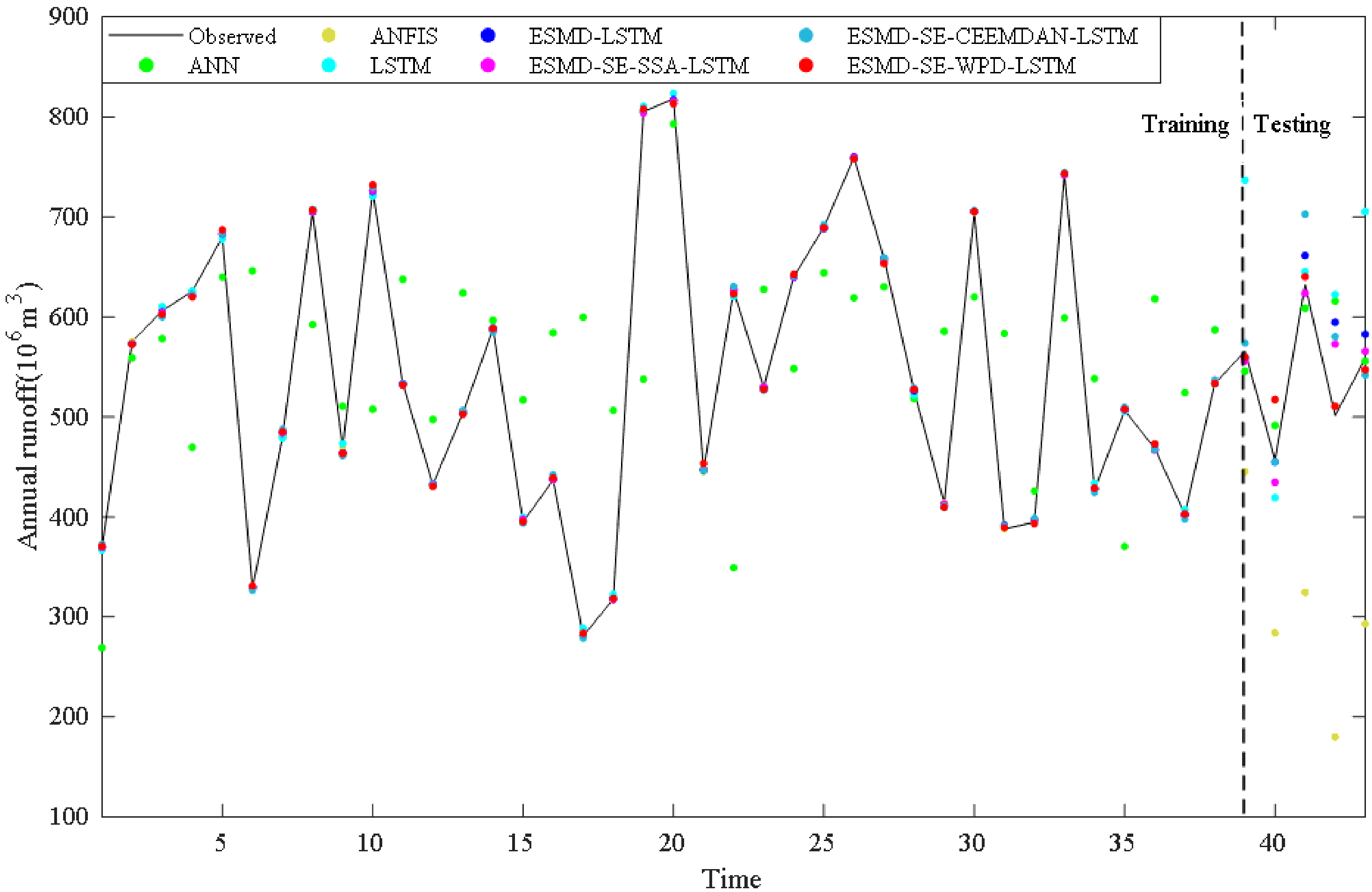

Figure 10

Forecasting results at Site 1.

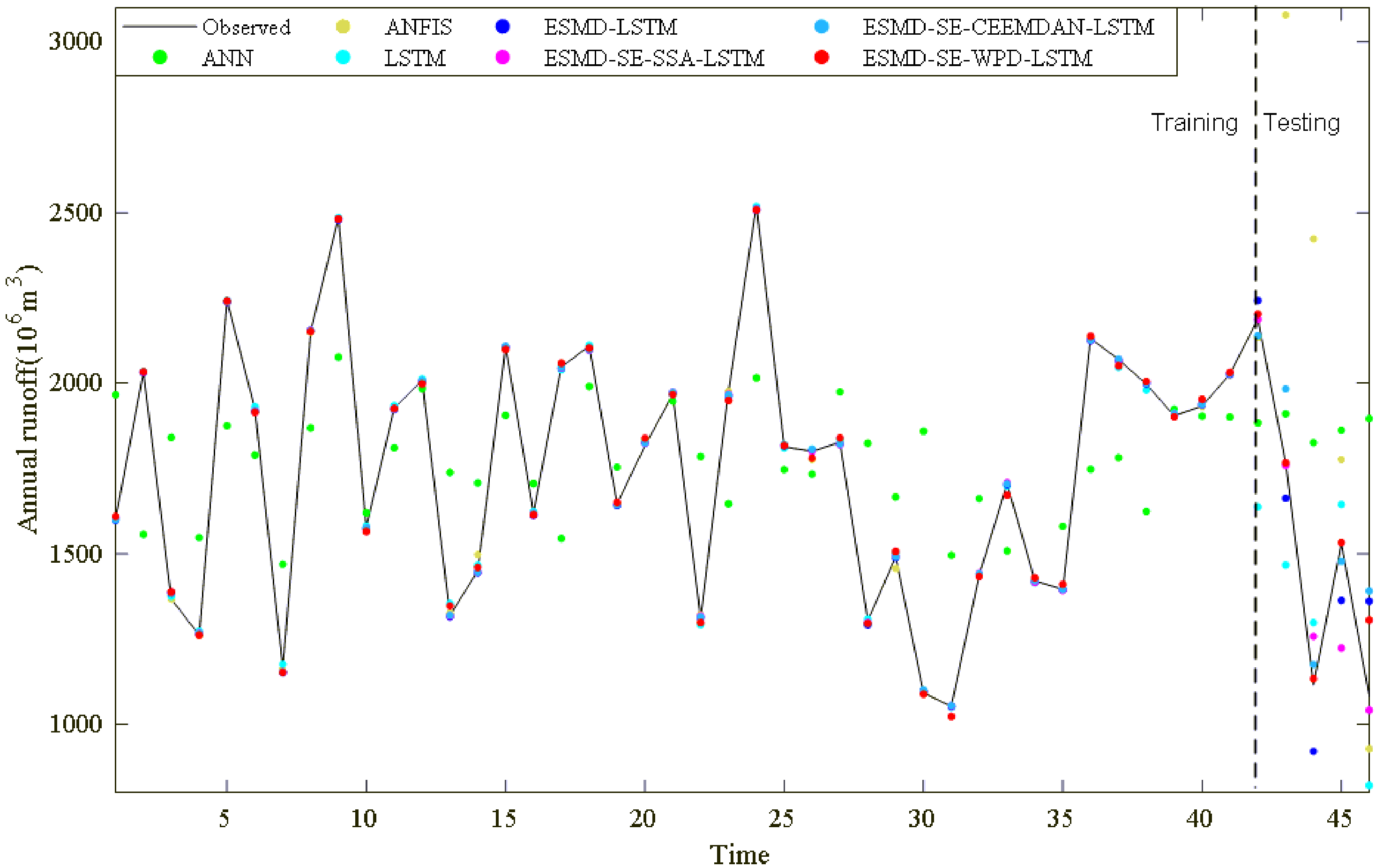

Figure 11

Forecasting results at Site 2.

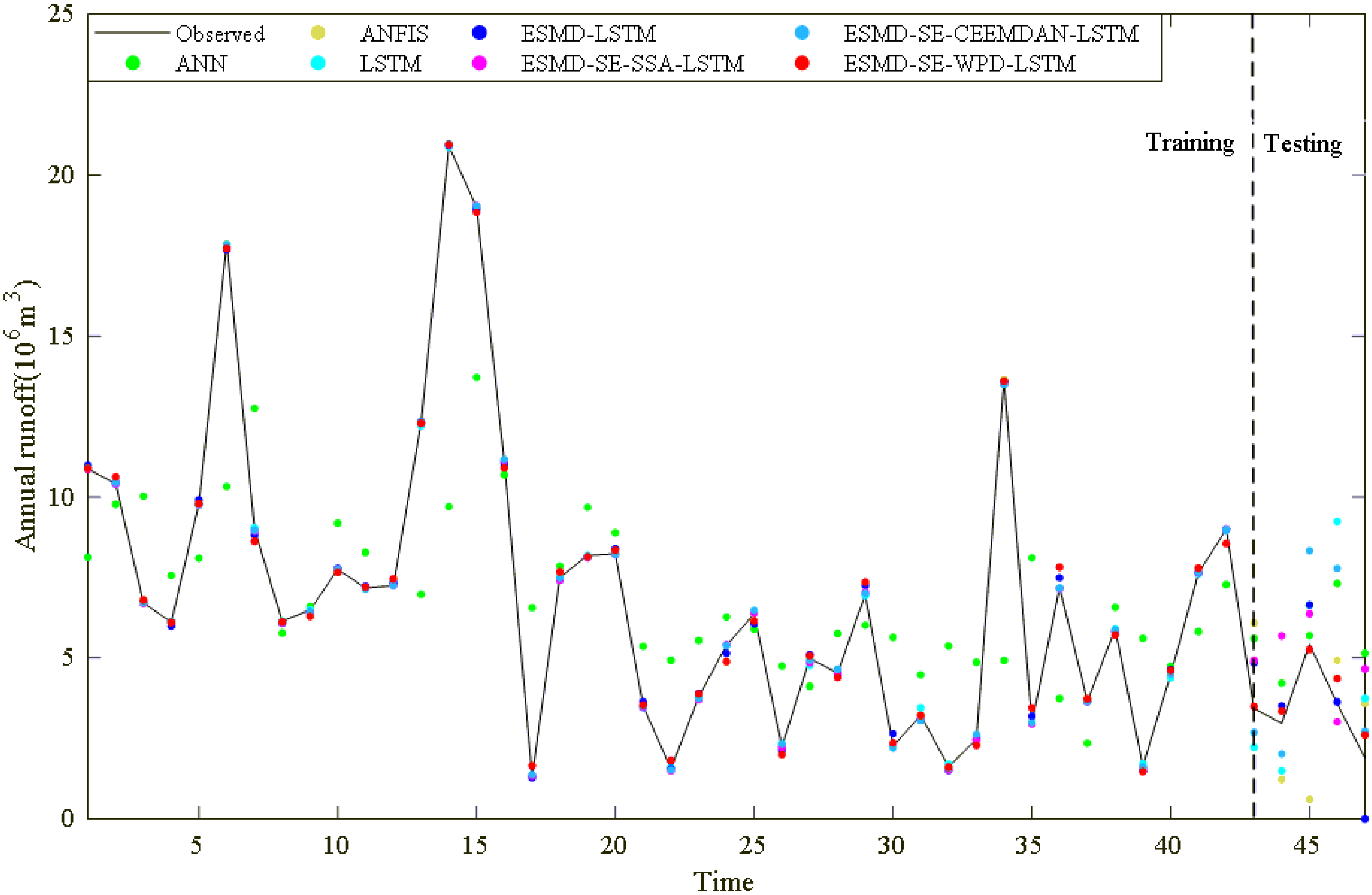

Figure 12

Forecasting results at Site 3.

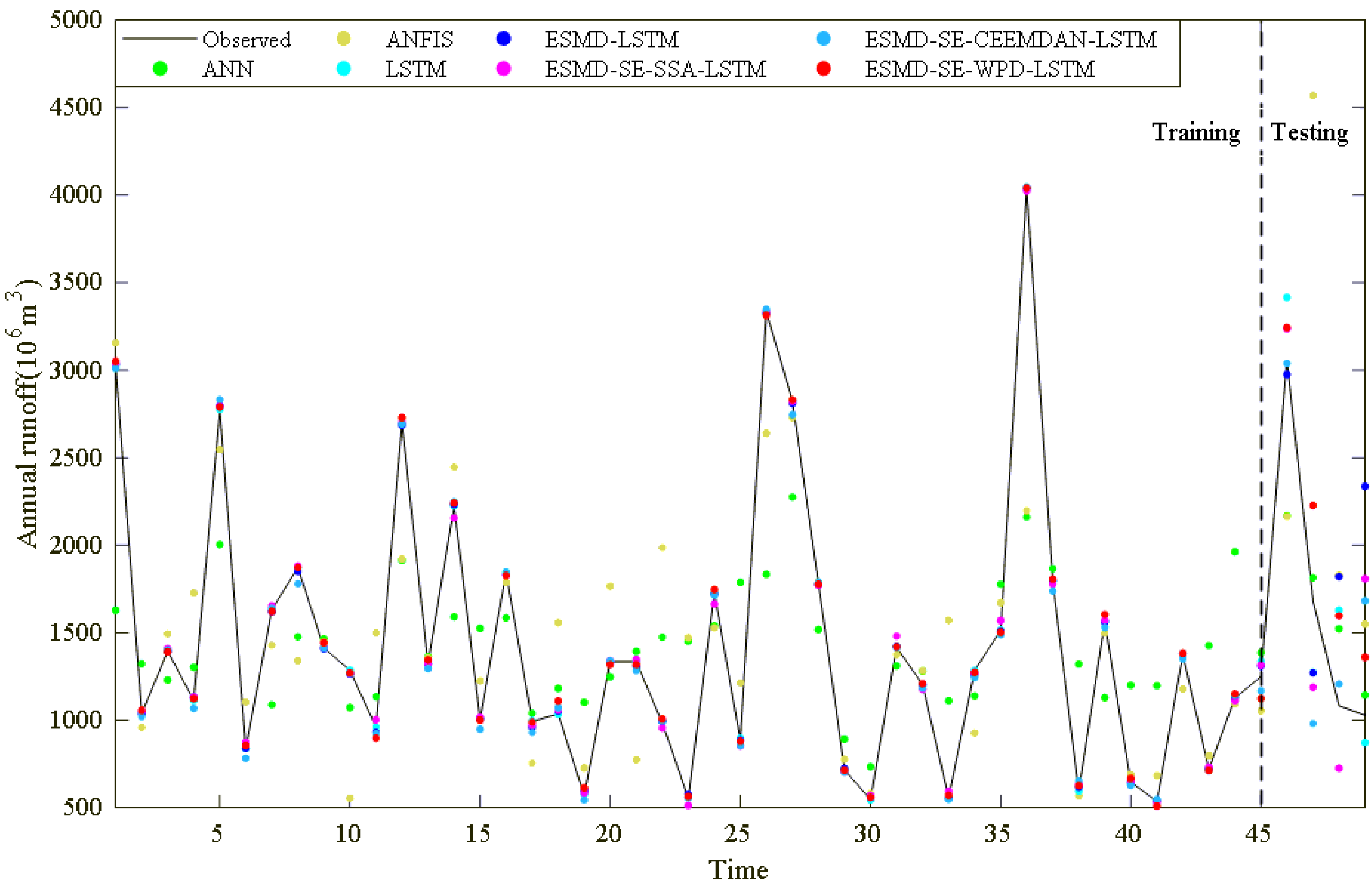

Figure 13

Forecasting results at Site 4.

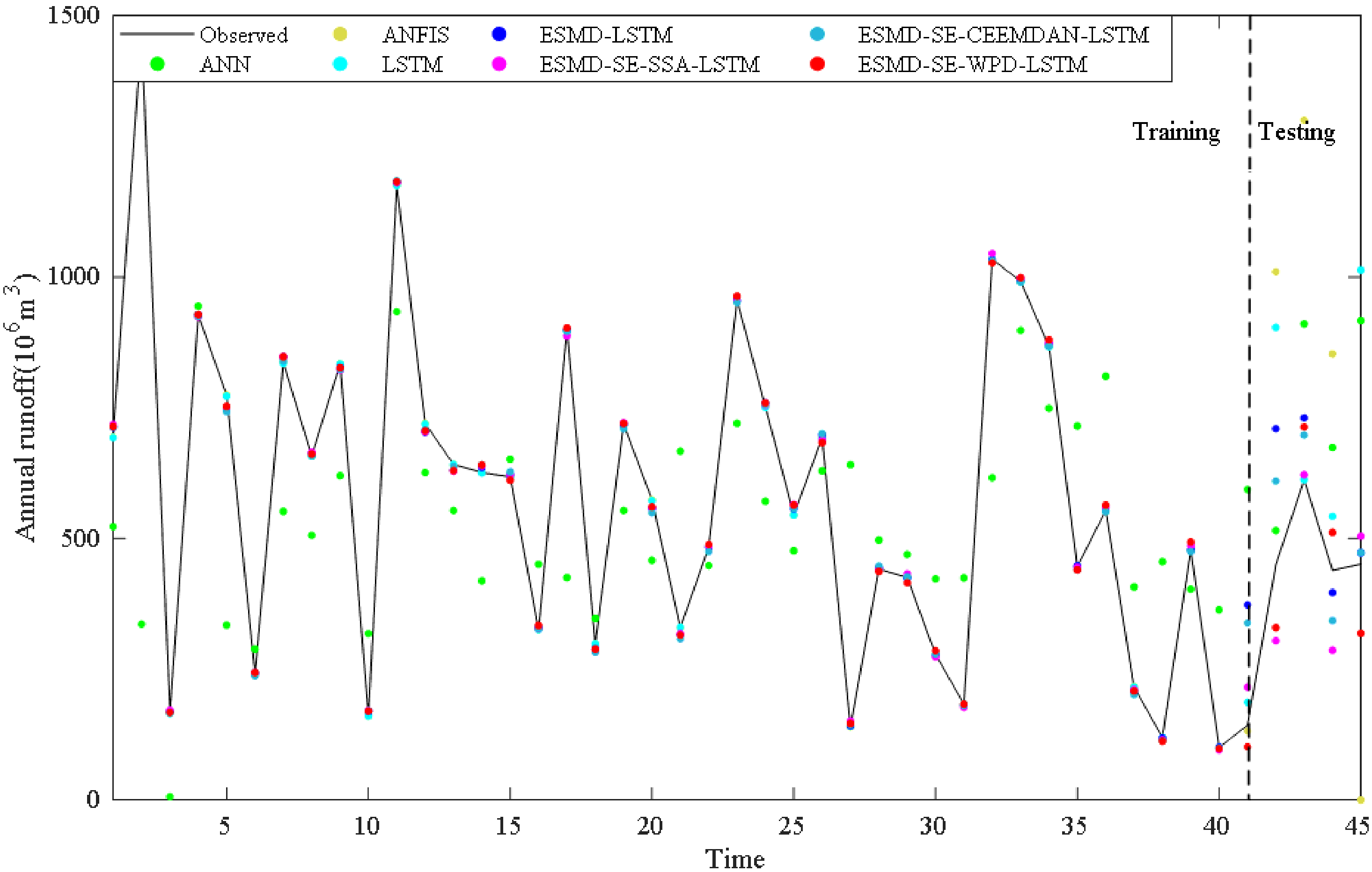

Figure 14

Forecasting results at Site 5.