



# Decomposition-ANN Methods for Long-Term Discharge Prediction Based on Fisher's Ordered Clustering with MESA

Fang-Fang Li<sup>1</sup> · Zhi-Yu Wang<sup>2</sup> · Xiao Zhao<sup>1</sup> · En Xie<sup>1</sup> · Jun Qiu<sup>3</sup> 

Received: 8 November 2018 / Accepted: 19 May 2019 /

Published online: 6 June 2019

© Springer Nature B.V. 2019

## Abstract

Precise and reliable long-term streamflow prediction contributes to water resources planning and management. Artificial neural network (ANN) have shown its remarkable ability in forecasting non-linear hydrological processes without involvement of complex, dynamic, hydrological and hydro-climatologic physical process in the water shed. To improve its non-stationary responses, decomposition methods are adopted as pre-processing methods in this study including Empirical Mode Decomposition (EMD), Ensemble Empirical Mode Decomposition (EEMD) and Seasonal-Trend decomposition using Loess (STL). The original time sequence is decomposed to several components, which are then taken as the inputs of the ANN model. EMD and EEMD are data- adaptable methods, and thus the number of Intrinsic Mode Functions (IMFs) might differ for different sequences, leading to the discrepancy of the input number for ANN model in training and predicting. Fisher's ordered clustering is thus used to classify the IMFs into a determined number of classes based on their frequency spectrum resulting from Maximum Entropy Spectral Analysis (MESA). The proposed methodology is applied on four important hydrological stations on the upper stream of the Yellow River and the Yangtze River in China, respectively, to forecast the streamflow of the next whole year with the historical daily data of the past 6 years. The Nash-Sutcliffe efficiencies of the monthly prediction are higher than 0.85 for all of the four cases, and various indicators indicates that the proposed hybrid method of STL-ANN performs better than other compared methods. The highlights of this study lies in that only historical daily streamflow data is used to derive an accurate long-term prediction by data mining based on decomposition technology and mapping relationships between the decomposed components and the original sequence in the future.

**Keywords** Long-term prediction · Ensemble empirical mode decomposition (EEMD) · Artificial neural network (ANN) · Hybrid model · Maximum entropy spectral analysis (MESA)

---

✉ Jun Qiu  
aeroengine@tsinghua.edu.cn

## 1 Introduction

Forecasting streamflow over a long period is of great significance, not only for various evaluations, such as reliability of water supply, risk of drought and flood disasters, security of the eco-environmental maintenance (Wang et al. 2013), but also for the water resources exploitation, such as water transferring and reservoir construction. Although systematic hydrometric stations and advanced monitoring techniques have realized short-term streamflow prediction with high accuracy, the long-term series is still difficult to identify due to the randomness and complexity (Zhu et al. 2016).

The hydrological changes are influenced by various uncertain factors, such as climate, weather, human activities and geographical environment, the mechanism of which haven't been fully understood. Besides, it is a complex process, manifested in: (1) the hydrological system is a highly nonlinear system taking precipitation as the input and the streamflow as the output, and the intermediate processes include evaporation, infiltration, runoff yield and confluences is highly non-linear; (2) it is hard to establish an accurate mathematical model to describe the streamflow process, as such a complex dynamic system involving meteorology, geology, etc., the sub-process cannot be strictly described by mathematical formulae. Limited by various assumptions or conceptual elements or empirical relationships, the simulation is usually an approximation of reality with low accuracy; (3) there exist various uncertainty, randomness, and ambiguity in the streamflow process.

The traditional methods for long-term streamflow forecasting can be categorized into cause analysis and the statistical analysis. The cause analysis methods are based on dynamic physical models of atmospheric circulation (Singhrattna et al. 2012), weather processes (Smiatek et al. 2012) and the physical condition of the underlying surface (Sinha et al. 2014). Many countries and regions have not established complete hydrological monitoring systems, and thus some information needed in the physical model is difficult to obtain. Statistical analysis uses mathematical methods to investigate the changing rules of the hydrological factors, which are then applied on forecasting, mainly including time series analysis (Wang et al. 2015), multivariate regression (Maslova et al. 2016), and similarity prediction (Sun et al. 2014). Although statistical analysis has mature model with simple procedures, its application is restricted by low accuracy without consideration of uncertainties.

In recent years, the advantages of intelligent methods in prediction of nonlinear systems are also taken for long-term streamflow forecasting, mainly including: gray system model (Ma et al. 2013), fuzzy algorithm (Shi et al. 2016), Artificial Neural Network (ANN) (Yu et al. 2014), wavelet analysis (Maheswaran and Khosa 2012), fractal theory (Tao et al. 2011), and mixed pure theory (Nourani et al. 2012).

ANN has strong non-linear mapping ability as a simple and operational model.. Satisfactory results have been achieved from ANN in some short-term streamflow forecasts. Wu et al. (2005) verified the validity of the ANN model for short-term discharge forecast taking historical rainfall and discharge as the inputs. Kisi et al. (2012) compared ANN model with local linear regression (LLR) and dynamic local linear regression (DLLR), and indicated that the ANN models performed better. Vafakhah (2012) used daily streamflow time series to develop ANNs, adaptive neuro-fuzzy inference system (ANFIS), and autoregressive moving average (ARMA) models, respectively, for 1-day ahead streamflow forecasting, and the results showed that the performance of ANNs are superior. Latt and Wittenberg (2014) used historical water level and rainfall data to build stepwise multiple linear regression (SMLR) and ANN models for 1~5 days ahead flood forecasts, and concluded that the performance of ANN

models are superior, particularly in the extreme floods forecasting. Nanda et al. (2016) used daily discharge and average temperature as the inputs of the linear autoregressive moving average with exogenous inputs (ARMAX) and static ANN models for flood forecasting with leading time of 1~3 days, and concluded that the estimated performance of ANN is superior to ARMAX.

In addition to short-term forecasting, developing appropriate ANN models for long-term streamflow forecasting has attracted more and more attention of researchers. Nourani et al. (2012) combined ANN and GP to develop hybrid WGPNN model for long-term forecasting, the accuracy of which is higher than the ANN model. Wang et al. (2015) developed hybrid EEMD-ANN model for long-term forecasting, whose performance is better than single ANN model. Badrzadeh et al. (2016) combined wavelet analysis and backpropagation neural networks (BPNN) model for long-term forecasting. Humphrey et al. (2016) coupled conceptual rainfall-runoff model with Bayesian ANN to forecast monthly streamflow with rainfall, evaporation, antecedent precipitation index (API), and ground water data as the inputs. However, these publications mainly use monthly or annual streamflow sequence as the inputs of ANN models for long-term forecasting, the result of which is often unreliable. The detailed data information contained in the time series with small time intervals haven't been fully excavated, which could be helpful for long-term forecasting.

Streamflow sequence is a time-related waveform, thus time-domain analysis is the fundamental method. To better understand the characteristics of the sequence, its frequency can be used to explore the intrinsic change rule and improve the prediction accuracy. Decomposition of the sequence can deconstruct a sequence into several components with different characteristics. Most of current decomposition methods are on the basis of wavelet analysis. The existing research shows that compared with the original time series, taking the decomposed time series derived from the wavelet analysis as the input of ANN results in higher prediction accuracy (Kasiviswanathan et al. 2016). Nevertheless, there are some considerable disadvantages of wavelet decomposition. For instance, inappropriate mother wavelet function or decomposition level may lead to significant decline in the precision accuracy (Maheswaran and Khosa 2012).

Empirical Mode Decomposition (EMD) (Huang et al. 1998) decomposes nonlinear signal into Intrinsic Mode Functions (IMFs) and one residual component, the time-frequency spectrum of which has physical significance. EMD can be used to decompose non-linear and non-stationary time series. Unlike wavelet transform, EMD works in temporal space directly rather than in the corresponding frequency space, and it is based on the principle of local-scale separation without need of any prior basis functions (Zhu et al. 2016). Decomposing the nonlinear data by EMD and then developing a hybrid model can be applied for streamflow forecasting. Kisi et al. (2014) used coupled EMD-ANN model and individual ANN model to forecast monthly discharge 1 month ahead, and concluded that the accuracy of EMD-ANN model is superior. Huang et al. (2014) coupled EMD with support vector machine (SVM) to forecast monthly streamflow, and the results indicated that EMD-SVM model has a good stability with higher accuracy compared with the individual SVM model. Zhang et al. (2015) combined EMD with ANN and ARMA model, respectively, to hindcast monthly streamflow.

Ensemble empirical mode decomposition (EEMD) is a new noise-assisted data analysis method based on EMD proposed by Wu and Huang (2009) to overcome the mode mixing problem without introducing an intermittent subjective as in EMD, which is employed for time series decomposition in this study, since it can clearly separate scales and requires no a prior subjective selection criterion (Wu and Huang 2009). There have been several studies applying EEMD on streamflow forecasting (Di et al. 2014).

In this study, EMD, EEMD and Seasonal Trends decomposition based on Loess (STL) are integrated with ANN for long-term stream forecasting without need of any basis functions like wavelet analysis. The original daily time series is firstly decomposed into several IMFs by EMD, EEMD, STL, respectively. As data adaptive methods, both EMD and EEMD might results in different number of decomposed components for different time series, which leads to inconsistency of the input number for ANN model in training and testing. Fisher's ordered clustering method is thus used to classify the IMFs into a designated number of clusters based on the Maximum Entropy Spectral Analysis (MESA). These models are used for streamflow forecasting of four cases for a whole year, and the Nash-Sutcliffe efficiencies of the monthly prediction are higher than 0.85 for all of the four cases with only historical daily streamflow data being used.

## 2 Methodology

### 2.1 Empirical Model Decomposition (EMD)

EMD is essentially a stabilization processing for non-linear sequence. Different components with different characteristic scales, called IMF, are decomposed from the original times series. Each IMF needs to meet the following two requirements (Huang et al. 2003): (1) the number of the extreme points and the number of zero points should be equal or differs by 1 at most in the entire data range; (2) the average of the upper and lower envelops should be zero at any point. The detailed EMD procedure can be found in the publication of Huang et al. (2003).

### 2.2 Ensemble Empirical Mode Decomposition (EEMD)

The error from observation and processing introduces uncertainty of the initial pattern of the numerical model. In order to overcome such problems, small random disturbance field can be superimposed on the prime field, and eliminated by taking ensemble average.

Wu and Huang (2009) applied the idea of ensemble prediction to decomposition by adding a certain percentage of the white noise to the original data. The basic principle is that when the additional white noise is evenly distributed throughout the whole time-frequency space, the space is composed of different components with different scales divided by the filter bank. When the signal is coupled with an even-distributed white noise background, the signal areas of different scales are automatically mapped to the appropriate scale corresponding to the background white noise. Finally the added white noise offsets each other by assemble average. Not only the signal information of the original scenario is reserved, but also the mode mixing problem can be overcome to a great extent to ensure the physical uniqueness of the decomposition.

The procedure of EEMD includes: (1) superimpose white noise sequence with the given amplitude on the original sequence; (2) implement EMD on the mixed signal to obtain IMF1; (3) superimpose a noise sequence with the same amplitude on the original sequence which has been stripped with IMF1; and (4) carry out EMD on the new sequence to get IMF2. The above steps are repeated until different IMFs and trend item RES are obtained.

### 2.3 Seasonal-Trend Decomposition Using Loess (STL)

STL uses iterative Loess smoothing to obtain an estimate of the trend and then Loess smoothing again to extract a changing additive seasonal component. A time series  $Y_v$  are decomposed into three components: trend  $T_v$ , seasonal term  $S_v$ , and remainder  $R_v$ , as in Eq. (1).

$$Y_v = T_v + S_v + R_v \tag{1}$$

STL is able to provide confidence intervals for multiple and complex seasonality (Cleveland et al. 1990). However, to balance the computational efficiency, the assumption that the seasonal components and coefficients for flexible predictors do not change quickly is taken in this study.

### 2.4 Fisher’s Ordered Clustering Based on Maximum Entropy Spectral Analysis (MESA)

Frequency domain analysis is generally superior to time domain analysis in analyzing the characteristics of signals, and the Power spectral density (PSD) describes how the power of a signal or time series is distributed over frequency. Compared with classical spectral estimation methods, the parametric modern spectrum analysis methods can be used for short sequence with higher resolution. Maximum Entropy Spectral Analysis (MESA) introduced by Burg (1967) with the basic principle of extrapolating the autocorrelation function under the maximum entropy criterion is adopted in this study to analyze the characteristics of the sub-sequences.

The power spectrum  $S(\omega)$  of an autoregressive (AR) model can be calculated by Eq. (2):

$$S(\omega) = \frac{\Delta t P_m}{\left| 1 + \sum_{i=1}^m a_i^{(m)} \cdot e^{-j\omega m \Delta t} \right|^2} \tag{2}$$

where  $P_m$  is the power of the prediction error. As white noise, the error is a random variable with normal distribution, and satisfies the maximum entropy condition. Equation (2) is thus the computational formula of maximum entropy power spectrum.

The key of MESA is to determine the parameter of the AR model  $a_m$  and the prediction error power  $P_m$ . In this study, the burg algorithm is adopted, using the sum of forward and backward errors as the predictive error, to minimize which, the  $a_m^{(m)}$  is determined, as shown in Eq. (3).

$$a_m^{(m)} = \frac{2 \sum_{t=m}^N e_t^{+(m)} e_t^{-(m)}}{\sum_{t=m}^N \left[ \left( e_t^{+(m)} \right)^2 + \left( e_t^{-(m)} \right)^2 \right]} \tag{3}$$

where  $e_t^{+(m)}$  and  $e_t^{-(m)}$  are the  $m$ -th order forward and backward prediction errors calculated by Eqs. (4) and (5), respectively.

$$e_t^{+(m)} = \sum_{j=0}^m a_j^{m-1} x_{t-j} = \sum_{j=0}^m a_{m-j}^{m-1} x_{t-m+j} \quad , \quad t = m, \dots, N \tag{4}$$

$$e_t^{-(m)} = \sum_{j=0}^m a_j^{m-1} x_{t-m+j} = \sum_{j=0}^m a_{m-j}^{m-1} x_{t-j} \quad , \quad t = m, \dots, N \tag{5}$$

Given  $a_m^{(m)}$ , the power-density spectrum can be obtained by Eqs. (6) and (7):

$$P_m = P_{m-1} \left( 1 - |a_m^{(m)}|^2 \right) \tag{6}$$

$$P_0 = \frac{1}{N} \sum_{j=1}^N x_j^2 \quad , \quad e_t^{+(0)} = e_t^{-(0)} = x_j \tag{7}$$

Plot  $P_m$  of the IMFs with different frequency, and if there exists a peak, the corresponding frequency is taken as the dominant frequency. MESA is suitable for the process with unknown basic distribution. The disadvantages of the classical Spectral Analysis such as subjective assumption of missing data are overcome by MESA.

With the given frequency, the IMFs are clustered by Fisher’s ordered clustering method, which divides  $N$ -sample sequence  $\{x_j\}$  into  $k$  classes and searches an optimal clustering scheme to make the sum of the dispersion minimum (Fisher 1958). There are two features of this method: one is that these  $N$  samples are ordered, and the other is that the continuity of the sample order is maintained during the clustering without any jump. The optimal clustering criterion is to get the minimum variance within the same class, and maximum variance among different classes.

### 2.5 Rolling EEMD-ANN and EMD-ANN Hybrid Methods

ANN is a network system simulating artificial intelligence with strong ability of nonlinear computation and learning. A neuron receives the weighted accumulation of all inputs connected, and the weighted sum is compared with a threshold. If the sum is greater than the

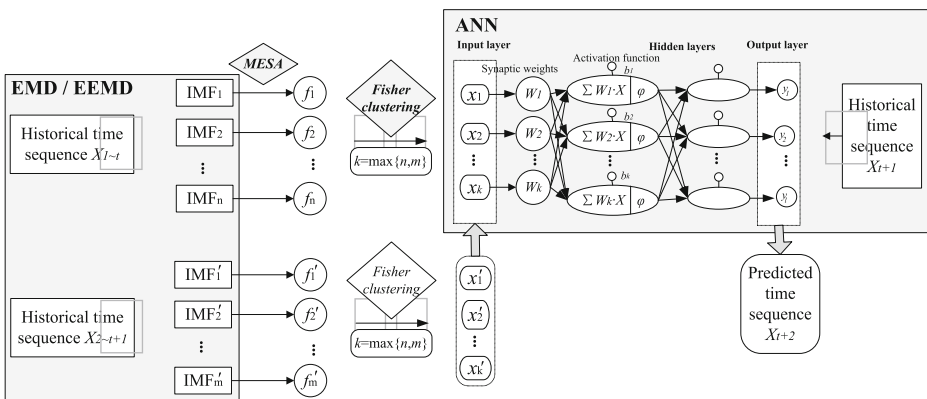
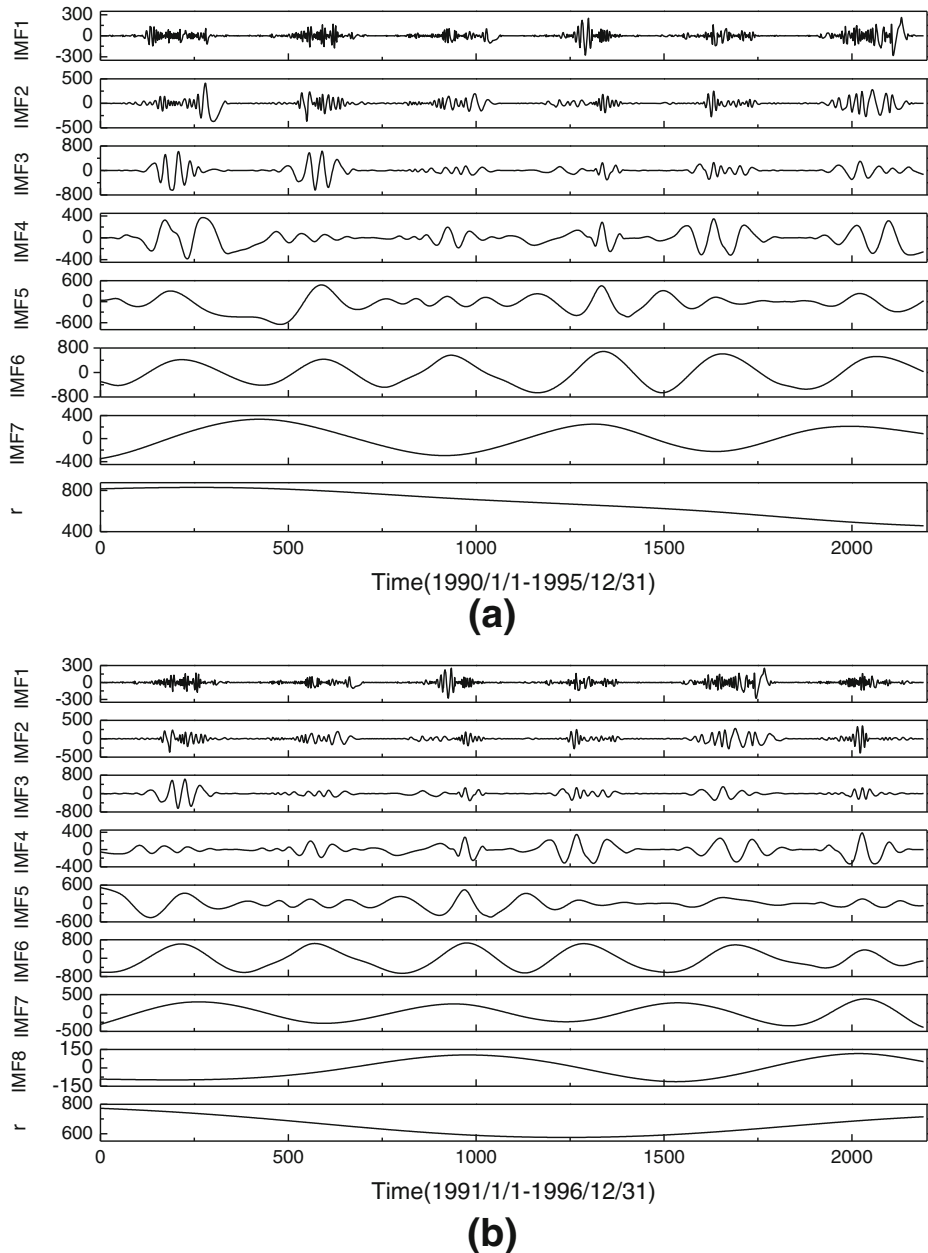


Fig. 1 Framework of the proposed rolling EEMD-ANN hybrid method

threshold, the artificial neuron is activated, and the signal is transmitted to the higher order neuron associated with it.

BP (Back-Propagation)-ANN is a multi-layer feed-forward network trained by transmitting error back, so that the weights and thresholds in the network are constantly adjusted to make the output of the network satisfying. It has been proved mathematically that the three-layer



**Fig. 2** Decomposition results of the historical data at the Yalong river by EMD: **a** from the year 1990 to 1995; **b** from the year 1991 to 1996

neural network with sigmoid nonlinear transfer function is able to arbitrarily approximate any continuous function.

Figure 1 shows the schematic diagram of the proposed decomposition-ANN hybrid method. The original sequence is firstly decomposed into  $n$  components by various methods, such as EMD, EEMD, and STL in this study.  $W$  and  $b$  are the weight matrix and the threshold vector of the neurons in the hidden layer, respectively. The sigmoid function is applied as the transfer function  $\varphi$ , where the hidden layer uses  $\text{tansig}(\cdot)$  and the output layer uses  $\text{logsig}(\cdot)$ . The original signal is used to adjust the weight and threshold, so as to find out the intrinsic relationship between the input and output. The BP algorithm essentially transforms a set of sample input and output problems into a nonlinear optimization problem, and uses its gradient to adjust the network parameters. If the output is not satisfying, the backward propagation begins, and the error signal returns along the original path. The error is minimized by repeatedly modifying the weights and thresholds of neurons in each layer.

The numbers of input nodes for both the training and the testing data need to be the same for ANN, while EMD/EEMD is a data adaptive method, i.e., different data set produces different numbers of IMF components. It is important to classify the IMFs into a settled number of categories.

Streamflow sequence contains multiple information with complex variation characteristics, such as high nonlinearity, periodicity in multiple time scales, and variability. STL uses LOESS to decompose a time series into three sub-series, and EMD/EEMD decompose the original signal into different IMF components with different frequency or trends. Hence, the frequency of the sub-series is selected as the classification criterion, estimated by MESA first, and then classified by on Fisher's ordered clustering to unify the number of the decomposed components, which are then taken as the input of the ANN.

When unifying the input neurons numbers of the training and the testing data for the ANN model, the classification number  $k$  of the Fisher's ordered clustering method is defined in Eq. (8) to reserve the information from the IMFs as much as possible:

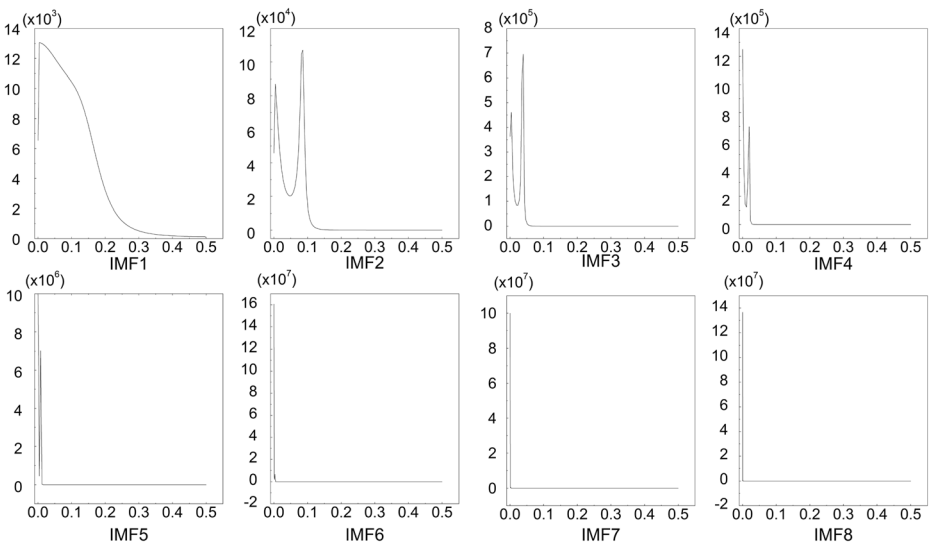
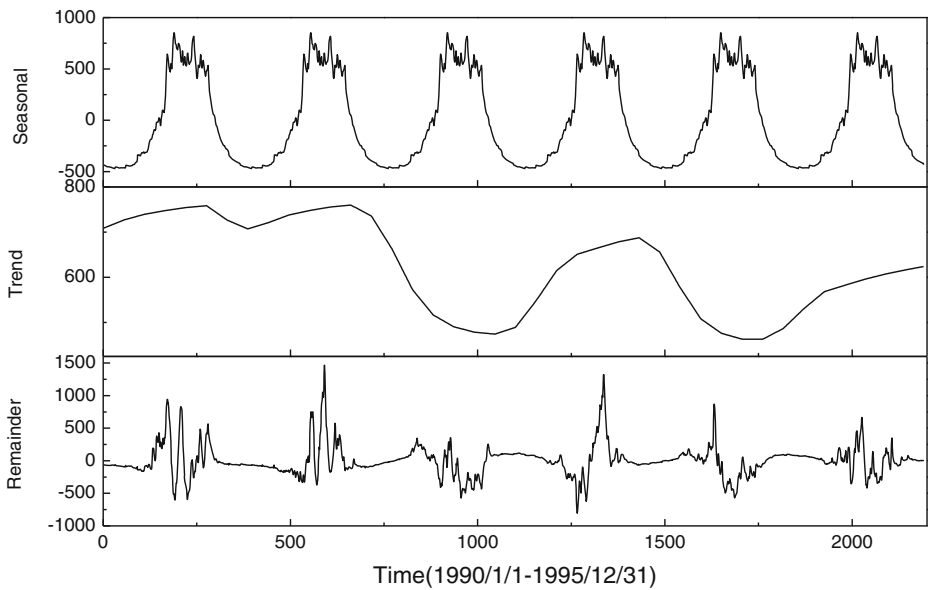
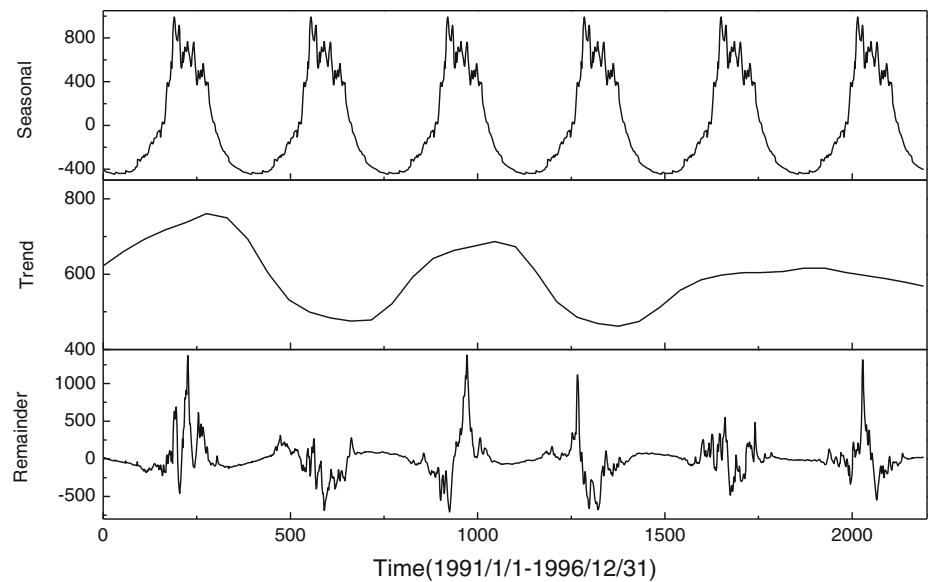


Fig. 3 PSD of IMFs decomposed by EMD with testing data at the Yalongjiang River





(a)



(b)

Fig. 4 Sub-time series components of original dataset decomposed by STL at the Yalongjiang River

$$k = \text{Min}\{n, m\} \tag{8}$$

where  $n$  and  $m$  are the number of IMFs derived from the trained and tested data of the ANN, respectively, as shown in Fig. 1.

### 3 Case Study

#### 3.1 Study Cases

The hybrid methods proposed in this study are applied for the streamflow forecasting at four hydrological stations including Tangnaihai and Lanzhou station located on the Yellow River, and Shigu and Yalongjiang station located on the Yangtze River in China. The daily streamflow data of 8 years are adopted in this study. The first 6 years is used to train the BP-ANN model, while the 2nd~7th year is used for test. The established model is used to predict the daily streamflow of the 8th year, which is compared with the historical data.

The Yellow River ranks as the sixth longest river in the world with the length of 5464 km. Tangnaihai station is the boundary of the source region of the Yellow River, where the annual runoff is about  $20 \times 10^9 \text{ m}^3$ . As the major water source area, the upper reaches of the Yellow River provides abundant water. Recent decades, with the rapid exploitation and development of the Yellow River Basin, the long-term streamflow forecast becomes more and more important for the water allocation and management. Lanzhou Station is also on upstream of the Yellow River. The daily streamflow time series of 8 years (from 2010 to 2017) are used for the two stations.

The Shigu station is located on the Jinshajiang River which is the upper reaches of the Yangtze River in China. The Yangtze River is the third longest river in the world. The daily

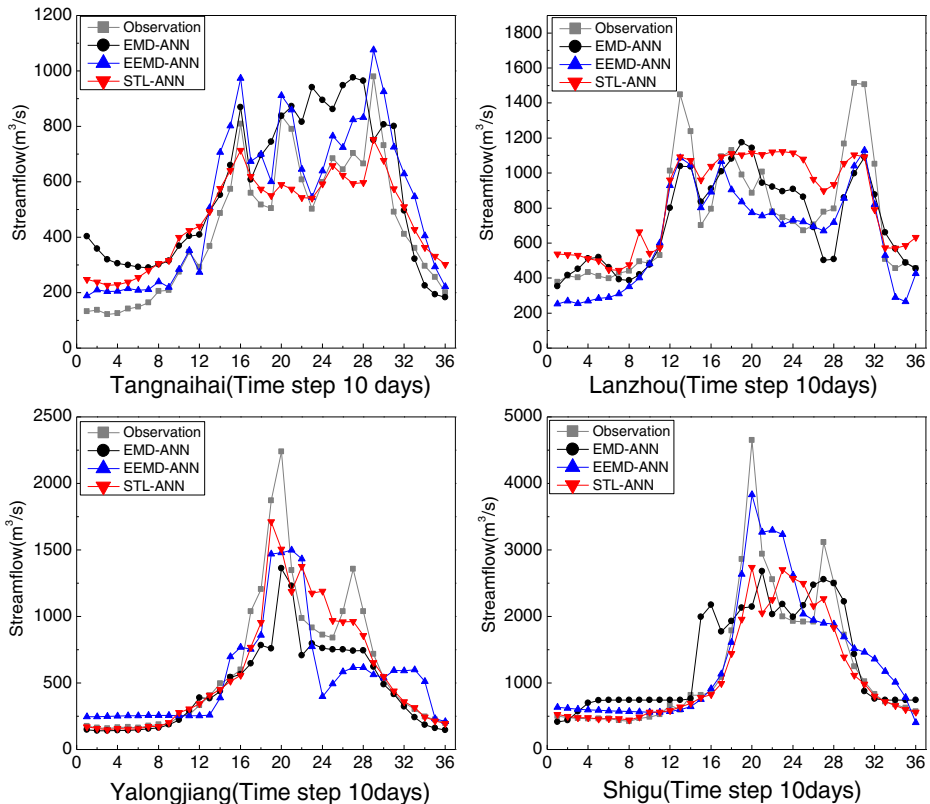


Fig. 5 Ten-day prediction results by the proposed hybrid methods for study cases

streamflow time series of 8 years (from 1990 to 1997) are utilized. The Yalongjiang station is located on Yalongjiang River which is the largest tributary of the Jinshajiang River, and also one of the rivers with the most abundant hydropower resources in China with the annual flow of about  $1860 \text{ m}^3/\text{s}$ . The daily streamflow data from 1990 to 1997 are chosen for study.

### 3.2 Decomposition from EMD/EEMD

Figure 2 shows the decomposition results of the historical streamflow of the Yalongjiang River by EMD. Figure 2a uses the training data from the year 1990 to 1995, deriving 7 IMFs and a residual; while Fig. 2b uses the testing data from the year 1991 to 1996, and results in 8 IMFs and a residual. Such discordance between IMFs which decomposed by EMD method proves the necessity of Fisher's clustering based on MESA, which helps to unify the number of inputs for the trained and tested data into the ANN model.

### 3.3 Classification of IMFs by Fisher's Ordered Clustering

For the Yalongjiang River, the PSD for each IMF of testing data can be plot as shown in Fig. 3 by using MESA, and the dominant frequency can be identified. For other study cases, the number of IMFs for the training and testing data into the ANN model are equal.

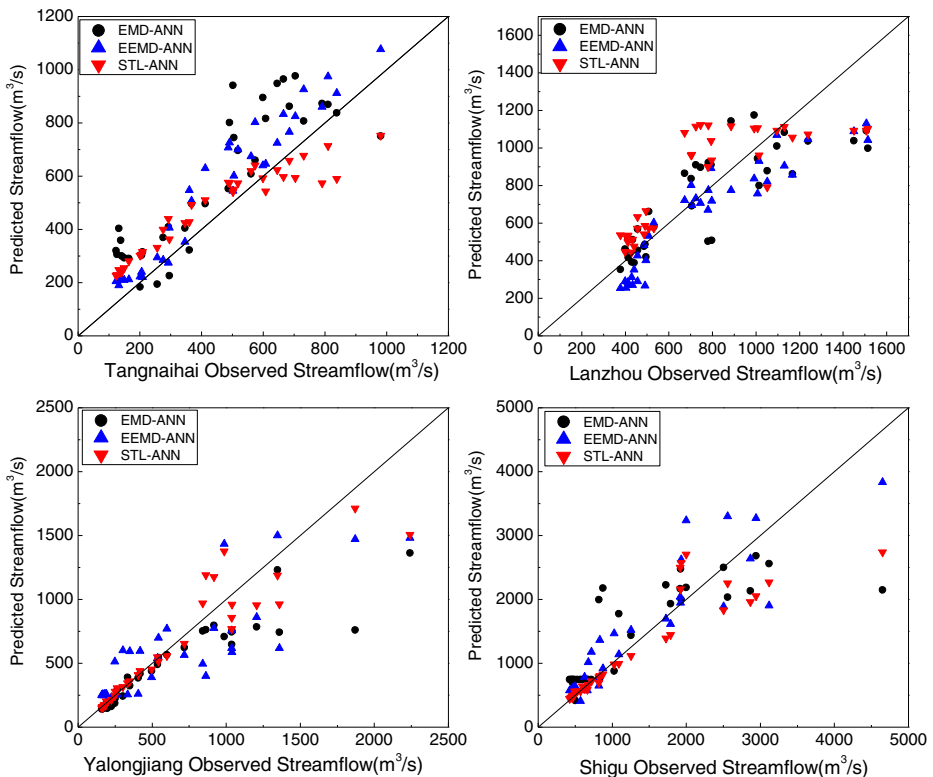


Fig. 6 Scatter plots of the 10-day predicted and observed data for study cases

The IMFs of testing data shown in Fig. 2 are classified in to 8 categories for the EMD by Fisher’s ordered clustering method which equate the number of IMFs for training data: IMF1, IMF2, IMF3, IMF4, IMF5, IMF6, IMF7 + IMF8, and r.

### 3.4 Data Processing for STL

The decomposed result by STL is shown in Fig. 4. Unlike the data adaptive method such as EMD and EEMD, STL results in three terms for any time series: seasonal term, trend term, and residual term. The training dataset from the year 1990 to 1995 are used to decompose by STL, as shown in Fig. 4a, and the testing dataset from the year 1991 to 1996 are used to decompose by STL, as shown in Fig. 4b.

## 4 Results and Discussion

The long-term streamflow forecasting results for four study cases by the proposed STL-ANN, EEMD-ANN, and EMD-ANN methods based on Fisher’s ordered clustering with MESA are compared in Figs. 5, 6, 7 and 8. Figures 5 and 6 show the 10-day prediction results, the monthly means of which is presented in Figs. 7 and 8.

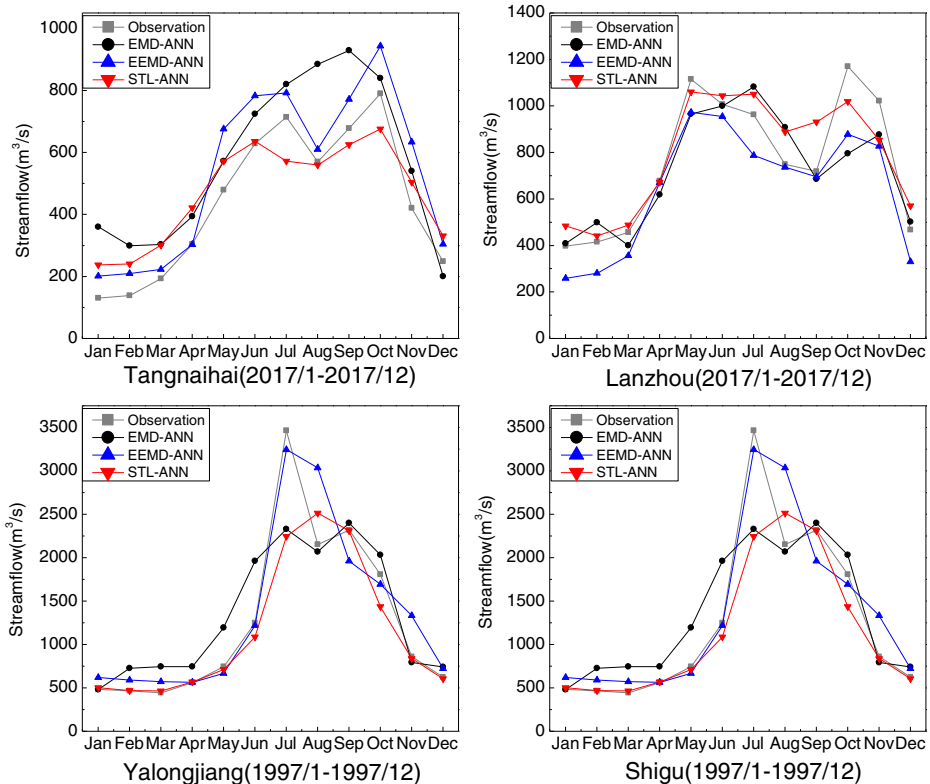


Fig. 7 Monthly prediction results by the proposed hybrid methods for study cases

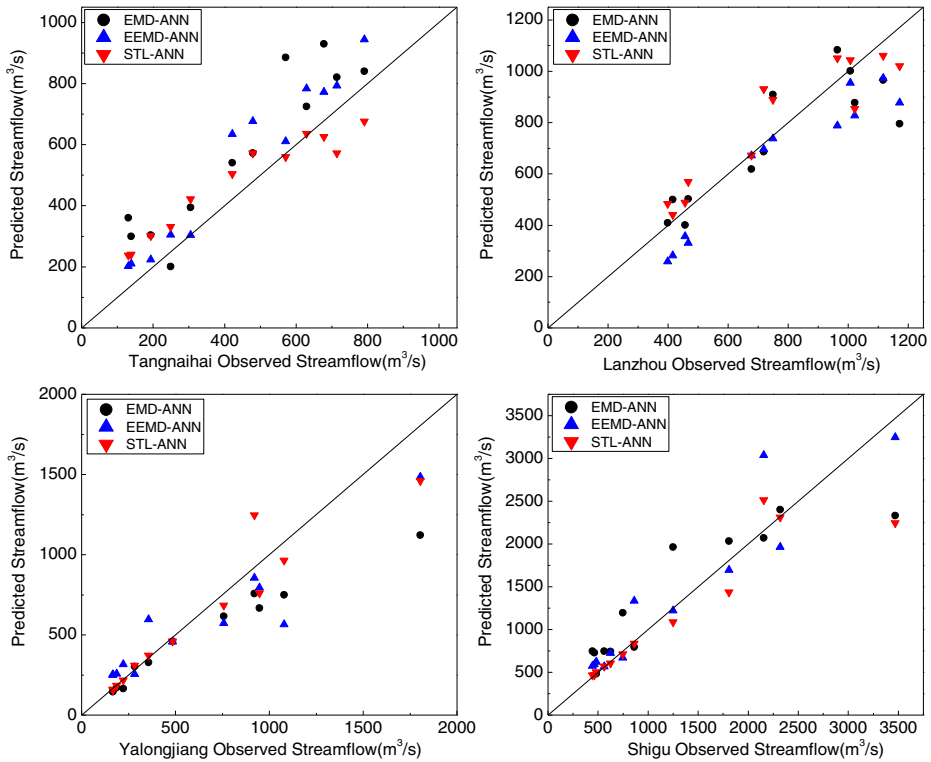


Fig. 8 Scatter plots of the monthly predicted and observed data for study cases

It can be seen that the prediction by STL-ANN method are generally closer to the observations than the EMD-ANN and EEMD-ANN hybrid methods from scatter plots. STL-ANN hybrid method can result in the prediction evenly distributed on both sides of the observation line.

The Root Mean Square Error (*RMSE*), Mean Absolute Error (*MAE*), as well as Nash-Sutcliffe Efficiency (*NSE*) are used as evaluation indices in this study, as defined from Eqs. (9) to (11), respectively. *RMSE* can be used to evaluate the fitting

Table 1 Evaluations of the prediction by different hybrid methods

Cases	EMD-ANN			EEMD-ANN			STL-ANN		
	RMSE	MAE	NSE	RMSE	MAE	NSE	RMSE	MAE	NSE
Ten-day prediction									
Tangnaihai	183.78	153.98	0.406	118.47	97.33	0.753	107.86	94.88	0.795
Lanzhou	190.73	144.27	0.675	175.63	136.99	0.724	157.02	120.25	0.780
Shigu	588.44	358.72	0.652	422.56	285.80	0.821	475.20	258.24	0.773
Yalongjiang	290.77	152.83	0.676	282.49	208.55	0.694	188.38	103.80	0.864
Monthly prediction									
Tangnaihai	160.31	138.85	0.763	116.21	96.62	0.875	93.82	84.85	0.919
Lanzhou	142.77	117.49	0.941	141.46	102.69	0.942	110.87	91.95	0.964
Shigu	435.36	301.83	0.827	324.54	221.19	0.904	386.35	187.90	0.864
Yalongjiang	242.89	149.19	0.786	207.30	155.43	0.844	152.84	94.35	0.915

degree of the predicted value and the high flow observed data; while *MAE* is adopted to evaluate the fitting degree of the predicted value and the observed data with middle/low flow. *NSE* represents the performance of hybrid methods.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i^o - x_i^p)^2} \quad (9)$$

$$MAE = \frac{1}{N} |x_i^o - x_i^p| \quad (10)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (x_i^o - x_i^p)^2}{\sum_{i=1}^n (x_i^o - \bar{x}^o)^2} \quad (11)$$

where  $N$  denotes the number of datasets;  $x_i^o$  and  $x_i^p$  represent the observation and the prediction, respectively.

The evaluation results are shown in Table 1.

Both the graphic comparison and the statistics of prediction accuracy indicate that the STL-ANN method performs the best among the three predictive methods in this study. The NSE of the STL-ANN hybrid model for monthly prediction are higher than 0.85, and the prediction result is satisfying.

The EMD method might have mode mixing problem because of signal intermittency which is defined as that a single IMF contains signals of widely disparate scales, or a signal of a similar scale resides in different IMF components. Such problem would result in the inability of representing instinct different time scale characteristics of original time series for those IMFs (Wu and Huang 2009). To resolve the mode mixing problem, EEMD method was developed. EEMD defines the true IMF component as the mean of an ensemble of trials, and each trial consists of the original signal and a white noise with finite amplitude. The additional white noise exists through the whole time-frequency space uniformly with the constituting components of different scales. EEMD can eliminate the mode mixing problem to a great extent and preserve physical uniqueness of the decomposition (Wu and Huang 2009). With the help of the EEMD method, each component represents instinct change rule for different time scale of the original data, therefore the accuracy of EEMD-ANN method is higher than EMD-ANN.

## 5 Conclusion

As data-driven methods, ANN does not involve complex, dynamic, hydrological and hydro-climatologic physical process in the water shed, and have shown promise in modeling and forecasting non-linear hydrological processes. However, it has some drawbacks with non-stationary responses. To handle such instances of “seasonality”, the input data pre-processing is required. In this paper, a hybrid method mainly coupling EMD/ EEMD/ STL and ANN model is proposed. The difficulty of applying EMD/EEMD/STL methods on ANN input pre-processing lies in their self- adaptability, i.e., different number of IMFs will result from

different original time series, which leads to the disagreement of the input data between the training and predicting process. The Fisher's ordered clustering is thus adopted to classify the IMFs into a determined number of classes according to the frequency spectrum of each IMF from MESA. Three statistical performance evaluation measures (*MAE*, *RMSE*, and *NSE*) are adopted to evaluate various methods. The statistical results indicates that the proposed STL-ANN method can perform superiorly to the other hybrid methods, which is conducive to long-term hydrological prediction.

**Acknowledgements** This research was supported by the Integration Program of the Major Research Plan of the National Natural Science Foundation of China (91847302), the National Natural Science Foundation of China (51879137), and National Key R&D Program of China (2017YFC0403600, 2017YFC0403602).

## Compliance with Ethical Standards

**Conflict of Interest** None.

## References

- Badrzadeh H, Sarukkalige R, Jayawardena AW (2016) Improving ANN-based short-term and long-term seasonal river flow forecasting with signal processing techniques. *River Res Appl* 32:245–256
- Burg JP (1967) Maximum entropy spectral analysis. In: 37<sup>th</sup> annual international meeting. Society of Exploration Geophysics, Oklahoma
- Cleveland RB, Cleveland WS, McRae JE, Terpenning I (1990) STL: a seasonal-trend decomposition procedure based on loess. *J Off Stat* 6(1):3–73
- Di CL, Yang XH, Wang XC (2014) A four-stage hybrid model for hydrological time series forecasting. *PLoS One* 9:8
- Fisher DW (1958) On grouping for maximum homogeneity. *J Am Stat Assoc* 53(284):789–798
- Holt CC (2004) Forecasting seasonals and trends by exponentially weighted moving averages. *Int J Forecast* 20: 5–10
- Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, Liu HH (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. In: *Proceedings of the Royal Society of London a: mathematical, physical and engineering sciences*, the Royal Society
- Huang NE, Wu ML, Qu W, Long SR, Shen SS (2003) Applications of Hilbert-Huang transform to non-stationary financial time series analysis. *Appl Stoch Model Bus Ind* 19(3):245–268
- Huang SZ, Chang JX, Huang Q, Chen YT (2014) Monthly streamflow prediction using modified EMD-based support vector machine. *Journal of Hydrology* 511:764–775
- Humphrey GB, Gibbs MS, Dandy GC, Maier HR (2016) A hybrid approach to monthly streamflow forecasting: integrating hydrological model outputs into a Bayesian artificial neural network. *J Hydrol* 540:623–640
- Hyndman RJ, Koehler AB, Snyder RD, Grose S (2002) A state space framework for automatic forecasting using exponential smoothing methods. *Int J Forecast* 18:439–454
- Kasiviswanathan KS, He J, Sudheer KP, Tay JH (2016) Potential application of wavelet neural network ensemble to forecast streamflow for flood management. *J Hydrol* 536:161–173
- Kisi O, Nia AM, Gosheh MG, Tajabadi MRJ, Ahmadi A (2012) Intermittent streamflow forecasting by using several data driven techniques. *Water Resour Manag* 26:457–474
- Kisi O, Latifoglu L, Latifoglu F (2014) Investigation of empirical mode decomposition in forecasting of hydrological time series. *Water Resour Manag* 28:4045–4057
- Latt ZZ, Wittenberg H (2014) Improving flood forecasting in a developing country: a comparative study of stepwise multiple linear regression and artificial neural network. *Water Resour Manag* 28:2109–2128
- Ma ZK, Li ZJ, Zhang M, Fan ZW (2013) Bayesian statistic forecasting model for middle-term and long-term runoff of a hydropower station. *J Hydrol Eng* 18(11):1458–1463
- Maheswaran R, Khosa R (2012) Comparative study of different wavelets for hydrologic forecasting. *Comput & Geosci* 46:284–295

- Maslova I, Toclavilca AM, McKee M (2016) Adjusting wavelet-based multiresolution analysis boundary conditions for long-term streamflow forecasting. *Hydrol Process* 30(1):57–74
- Nanda T, Sahoo B, Beria H, Chatterjee C (2016) A wavelet-based non-linear autoregressive with exogenous inputs (WNARX) dynamic neural network model for real-time flood forecasting using satellite-based rainfall products. *Journal of Hydrology* 539: 57–73
- Nourani V, Komasi M, Alami M (2012) Hybrid wavelet- genetic programming approach to optimize ANN modeling of rainfall- runoff process. *J Hydrol Eng* 17(6):724–741
- Shi B, Hu CH, Yu XH, Hu XX (2016) New fuzzy neural network–Markov model and application in mid- to long-term runoff forecast. *Hydrol Sci J* 61(6):1157–1169
- Singhtrattna N, Babel MS, Perret SR (2012) Hydroclimate variability and long-lead forecasting of rainfall over Thailand by large-scale atmospheric variables. *Hydrol Sci J* 57(1):26–41
- Sinha T, Sankarasubramanian A, Mazrooei A (2014) Decomposition of sources of errors in monthly to seasonal streamflow forecasts in a rainfall- runoff regime. *J Hydrometeorol* 15:2470–2483
- Smiatek G, Kunstmann H, Werhahn J (2012) Implementation and performance analysis of a high resolution coupled numerical weather and river runoff prediction model system for an alpine catchment. *Environ Model Softw* 38:231–243
- Sudheer C, Maheswaran R, Panigrahi BK, Mathur S (2014) A hybrid SVM-PSO model for forecasting monthly streamflow. *Neural Comput & Applic* 24:1381–1389
- Sun AY, Wang DB, Xu XL (2014) Monthly streamflow forecasting using Gaussian process regression. *J Hydrol* 511:72–81
- Tao J, Chen X-H, Wang L, Xie Y-W (2011) Study on fractal characteristics of runoff time series in the Beijiang River. *Acta Scientiarum Naturalium Universitatis Sunyatseni* 50:148–152
- Vafakhah M (2012) Application of artificial neural networks and adaptive neuro-fuzzy inference system models to short-term streamflow forecasting. *Can J Civ Eng* 39(4):402–414
- Wang W-C, Xu D-M, Chau K-W, Chen S (2013) Improved annual rainfall-runoff forecasting using PSO–SVM model based on EEMD. *J Hydroinf* 15(4):1377–1390
- Wang WC, Chau KW, Qiu L, Chen YB (2015). Improving forecasting accuracy of medium and long-term runoff using artificial neural network based on EEMD decomposition. *Environmental Research* 139:46–54
- Wu Z, Huang NE (2009) Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Adv Adapt Data Anal* 1(01):1–41
- Wu JS, Han J, Annambhotla S, Bryant S (2005) Artificial neural networks for forecasting watershed runoff and stream flows. *J Hydrol Eng* 10:216–222
- Yu JJ, Qin XS, Larsen O, Chua LHC (2014) Comparison between response surface models and artificial neural networks in hydrologic forecasting. *J Hydrol Eng* 19(3):473–481
- Zhang XL, Peng Y, Zhang C, Wang BD (2015) Are hybrid models integrated with data preprocessing techniques suitable for monthly streamflow forecasting? Some experiment evidences. *J Hydrol* 530:137–152
- Zhu S, Zhou JZ, Ye L, Meng CQ (2016) Streamflow estimation by support vector machine coupled with different methods of time series decomposition in the upper reaches of Yangtze River, China. *Environ Earth Sci* 75(6): 531

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

## Affiliations

Fang-Fang Li<sup>1</sup> · Zhi-Yu Wang<sup>2</sup> · Xiao Zhao<sup>1</sup> · En Xie<sup>1</sup> · Jun Qiu<sup>3</sup>

<sup>1</sup> College of Water Resources & Civil Engineering, China Agricultural University, Beijing 100083, China

<sup>2</sup> Shandong Water Conservancy Vocational College, Rizhao 276826, China

<sup>3</sup> State Key Laboratory of Hydroscience & Engineering, Tsinghua University, Beijing 100084, China