CrossMark

ORIGINAL ARTICLE

Streamflow estimation by support vector machine coupled with different methods of time series decomposition in the upper reaches of Yangtze River, China

Shuang Zhu[1] · Jianzhong Zhou[1] · Lei Ye[1] · Changqing Meng[1]

Abstract Machine learning models combined with time series decomposition are widely employed to estimate streamflow, yet the effect of the utilization of different decomposing methods on estimating accuracy is inadequately investigated and compared. In this paper, the main objective is to research the predictability of monthly streamflow using support vector machine model coupled with discrete wavelet transform (DWT) and empirical mode decomposition (EMD). The influence of the noise component of the decomposed time series on the forecast accuracy is also discussed here. Performance is evaluated through an application on Jinsha River, which is located in the upper reaches of Yangtze River in China. Results indicate that both time series decomposition techniques EMD and DWT contribute to improving the accuracy of streamflow prediction, and deeper comparative analysis shows models coupled with DWT have better prediction capabilities than models coupled with EMD. Furthermore, the high frequency component of the original series is indispensable for high-precision streamflow prediction, which is obvious in flood season.

Keywords Discrete wavelet transform · Empirical mode decomposition · Support vector machine · Monthly streamflow forecasting

✉ Jianzhong Zhou
  jz.zhou@hust.edu.cn

  Shuang Zhu
  zhushuanghust08@gmail.com

[1] School of Hydropower and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China

Introduction

Due to the fact that providing accurate and reliable future streamflow information plays an important role in water allocation, flood-control and disaster relief, developing excellent streamflow forecasting methods are significant and has attracted more and more attention of hydrology researchers in this field (Dehghani et al. 2015; Rahman et al. 2014; Wang et al. 2011). For daily or hourly run-off prediction, researchers have developed relatively mature methods such as distributed conceptual models with physical mechanisms. These models use rainfall forecasting as input to simulate the interception, infiltration and formation of run-off and achieve good run-off forecasting results. Compared with short-term run-off forecasting, reliable long-term run-off forecasting is much more difficult to accomplish for the lack of satisfying rainfall forecasting information. To solve this problem, researchers tend to use some data-based artificial intelligence models, which do not seek complex nature of run-off process and do not rely on high-precision rainfall forecasting to approach hydrological processes. Artificial neural network (ANN) and support vector regression (SVR) are very widely used artificial intelligence models. These models have the capacity of representing highly non-linear correlations of input and output, however, they are still found to be inadequate when dealing with non-stationary data (Sehgal et al. 2014).

Hydrological processes are non-stationary and random (Yarar 2014). When simply analyzing and simulating the original run-off series by these models, the details of change are ignored so that forecasting accuracy is reduced. Providing useful decompositions of original non-stationary time series on various resolution levels and extracting the significant information based on the series structure is a feasible way to make improvements in predictive ability of

🖄 Springer

a model. For this reason, in the present investigation, time series decomposition has been combined with the artificial intelligence models for predicting monthly streamflow. Hence, the choices of time series decomposition methods and prediction models are the most important issues in this study.

The decomposition of time series is a statistical method that deconstructs a time series into notional components. Various decomposition techniques based on different principles are recommended for time series prediction. One of the widely used decomposition techniques is discrete wavelet transform (DWT), which is very helpful for non-stationary processes. Labat (2005) made a detailed description of the mechanism of wavelet transform. It decomposes time series into both time and frequency domains and can produce trend, periods, and random component of original series. Since DWT analyzes time series in different scales independently, it is possible to reveal change of run-off on multi-resolution level. Due to these attractive properties, Budu (2013) combined feed-forward neural networks with wavelets preprocessing techniques to predict the daily inflow proved that the wavelets hybrid ANN model performed better compared to artificial neural network. Sahay and Srivastava (2014) proposed a wavelet transform-genetic algorithm-neural network model (WAGANN) for forecasting one-day-ahead monsoon river-flows and WAGANN models are superior to models using original flow-time series (OFTS) for inputs. Tiwari and Adamowski (2014) developed two types of hybrid wavelet-artificial neural network (WANN) to forecast weekly and monthly water demand with limited data availability, the results showed that the proposed methods were able to improve the accuracy and reliability of water demand forecasting by incorporating the capability of wavelet transformation. Another decomposition approach used for dealing with non-linear and non-stationary time series is Empirical mode decomposition (EMD). EMD is a signal processing tool that decomposes a time series into intrinsic modes without fixing any a priori bounds (Huang et al. 1998). EMD method decomposes data from non-stationary and non-linear processes into simple oscillatory functions through the Hilbert transform, and that will yield a meaningful instantaneous frequency. Contrary to 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 and doesn't need any predetermined basis functions (Huang et al. 1998; Lee and Ouarda 2010). Thus, it is empirical, intuitive, direct, and adaptive, with a posterior-idefined basis derived from the data. Decomposing the data through EMD and then building hybrid models to forecast streamflow were explored by Napolitano et al. (2011), Huang et al. (2014) and Wang et al. (2015).

The foregoing research mainly focused on developing all kinds of forecasting models combined with preprocessing decomposed techniques to achieve better performance; it is hard to find comparative investigation of different decomposition methods. Karthikeyan and Kumar (2013) adopted ARMA model based DWT and EMD to predict non-stationary time series and then compared the performance of two kinds of decomposing methods, but the influence of the highest-frequency components of DWT and EMD is not involved in their work. Analysis of high-frequency components is important to find out if they are just noise or if they capture important aspects of the signal dynamics. Therefore, a relatively complete comparative investigation needs to be made to discuss the effect of decomposition methods as well as highest frequency components on streamflow forecasting.

Selection of prediction models is another key step. Artificial intelligence models have achieved success in hydrological applications (Chen et al. 2014; Kasiviswanathan and Sudheer 2013; Sattari et al. 2012). Artificial neural network (ANN) has been widely applied to rainfall–runoff forecast, flood estimation and drought forecast due to its characters of adaptivity, self-organization, and self-learning ability (Hsu et al. 1995; Sattari et al. 2012). In recent decades, a new machine learning method called support vector machine (SVM) has been developed for classification and simulation. The training objective of SVM is to simultaneously minimize both the empirical risk and the model complexity so that local minima and over-fitting problems will be avoided and the training method of SVM is expressed as a convex quadratic programming problem on the high dimensional space which makes the training rapidly solved (Lin et al. 2009). Lin et al. (2009) used SVM to forecast reservoir inflow during typhoon-warning periods and the results indicated that SVM model performed better than back-propagation artificial neural network model. Guo et al. (2011) forecasted monthly streamflow based on improved SVM model and found that the SVM showed better generalization ability and higher prediction accuracy than ANN. Wei et al. (2013) adopted a dynamic particle filter-support vector regression method and obtained good predictions. Inspired by these attractive performances, support vector regression (SVR) is selected as prediction model in this investigation.

Therefore, in this study, time series decomposition techniques DWT and EMD are separately combined with the model of support vector regression (SVR). The effect of EMD and DWT on streamflow forecasting is compared and the influence of high frequency components on model performance is analyzed. Two kinds of models, i.e., EMD-SVR and DWT-SVR are developed. For the purpose of comparison, in addition to the above-mentioned decomposition-based models, stand-alone SVR model is also

developed. The constructed models are evaluated for forecasting streamflow for one-month-lead-time in the upper reaches of Yangtze River, China.

Methodology

Discrete wavelet transform

Wavelet analysis can reveal localized time and frequency information of non-stationary time series, so it is suitable for streamflow process, which is highly non-linear and includes a lot of random factors. Discrete wavelet transform (DWT) produces a series of approximations (low-pass version) to the original signal and details (high-pass version) at different resolution levels. The principle of DWT is as follows:

Suppose that wavelet function $\psi(t)$ is the mother wavelet satisfying $\int_{-\infty}^{+\infty} \psi(t)\mathrm{d}t = 0$, then successive wavelets $\psi_{a,b}(t)$ can be obtained through compressing and expanding $\psi(t)$ at scale a and location b:

$$\psi_{a,b}(t) = |a|^{-1/2}\psi\left(\frac{t-b}{a}\right) \tag{1}$$

For the discrete time series $f(t)$ with integer time steps, DWT in the dyadic decomposition scheme is defined as:

$$W_\psi f(j,k) = 2^{-j/2}\sum_{t=0}^{N-1} f(t)\bar{\psi}(2^{-j}t - k) \tag{2}$$

where $W_\psi f(j,k)$ is the wavelet coefficient of the discrete wavelet with $a = 2^j$, $b = 2^j k$, $\bar{\psi}(t)$ is the complex conjugate functions of $\psi(t)$.

$W_\psi f(j,k)$ reflects the characteristics of $f(t)$ in frequency and time domain at the same time. When the frequency resolution of wavelet transform is low and the time domain resolution is high, j becomes small. When the time domain resolution is low and the frequency resolution is high, j becomes large (Wang and Ding 2003).

Empirical mode decomposition

Empirical mode decomposition (EMD) produces Hilbert–Huang transform (HHT) by coupling with Hilbert spectral analysis (HSA) (Huang and Wu 2008). HHT acts similar to wavelet analysis, while the differences are that HHT is a posteriori and its theoretical basis is empirical.

The essence of EMD is to analyze characteristic time scales and recognize the intrinsic oscillatory modes empirically and finally, decompose the time series into a sum of different time modes accordingly. Each of these oscillatory modes is represented by an Intrinsic Mode Function (IMF). HHT is consistent with physically meaningful definitions of instantaneous frequency and amplitude

(Huang and Wu 2008), and it obtains the physical meaning related to the full non-linear system through individual components in the linear system rather than a physically linear expansion. In this way, HTT has the ability to acquire the important features of non-stationary and non-linear series (Lee and Ouarda 2011).

EMD is implemented through an adaptively iterative process called "sifting", which is the core of the algorithm. The sifting process serves to not only eliminate riding waves but also smoothen uneven amplitudes. For certain series $x(t)$, the sifting processes of EMD are described by Huang et al. (1998).

After decomposition, the original signal $x(t)$ can be written as a sum of IMFs $C_i(t)$ and a residual $r_n(t)$:

$$X(t) = \sum_{i=1}^{N} C_i(t) + r_n(t) \tag{3}$$

Support vector regression (SVR)

Support vector machine, which is known as classification and then extended to regression, was proposed by Vapnik (1995). Support vector regression (SVR) is the method solving the problem of regression with SVM. Detail description of SVR can be found in many open literature (Lin et al. 2009). The following is a brief description of SVR.

Suppose N known data points for training are $\{(X_i, d_i)\}_i^N$, X_i is m-dimensional input vector and d_i is 1D desired output at sample point i. The aim of SVR is to find a regression function in the form of Eq. (4).

$$y_i = \mathbf{W}\varphi(X_i) + b \quad \forall i = 1, 2, \ldots, N \tag{4}$$

where $\varphi(X)$ is a non-linear mapping, $\mathbf{W}$ is hyperplane, and b is offset.

It is worth mentioning that a penalty function is used in SVR (Guo et al. 2011):

$$\begin{cases} |d_i - y_i| \le \varepsilon, & \text{not allocating } a \text{ penalty} \\ |d_i - y_i| > \varepsilon, & \text{allocating } a \text{ penalty} \end{cases} \tag{5}$$

When the estimated value is within the ε- insensitive tube, the loss value will be zero. Parameters of above regression function can be acquired by minimizing the following objective function:

$$\min\left[\frac{1}{2}\|W\|^2 + C\sum_{i=1}^{N} L^\varepsilon(y_i, d_i)\right] \tag{6}$$

$$L^\varepsilon(y_i, d_i) = \max(0, |y_i - d_i| - \varepsilon) \tag{7}$$

where C represents the regularized constant that weighing the model complexity and the empirical error. A relative importance of the empirical risk will increase when the value of C increases.

Then the method of Lagrange multipliers is introduced to solve the above optimization problem in its dual form:

$$\max\left[\sum_{i=1}^{N}(\alpha_i^+ - \alpha_i^-)y_i - \varepsilon\sum_{i=1}^{N}(\alpha_i^+ + \alpha_i^-)\right.$$
$$\left. - \frac{1}{2}\sum_{i=1,j=1}^{N}(\alpha_i^+ - \alpha_i^-)(\alpha_j^+ - \alpha_j^-)K(X_i, X_j)\right] \tag{8}$$

where α_i^+, α_i^- are the Lagrange multipliers and $K(X_i, X_j)$ is a non-linear kernel function, which can map the lower dimension input into a higher dimension linear space.

Radial basis kernel function is used in this study:

$$K(X_i, X_j) = \exp\left(-\frac{||X_i - X_j||^2}{2\sigma^2}\right), \quad \sigma \in R \tag{9}$$

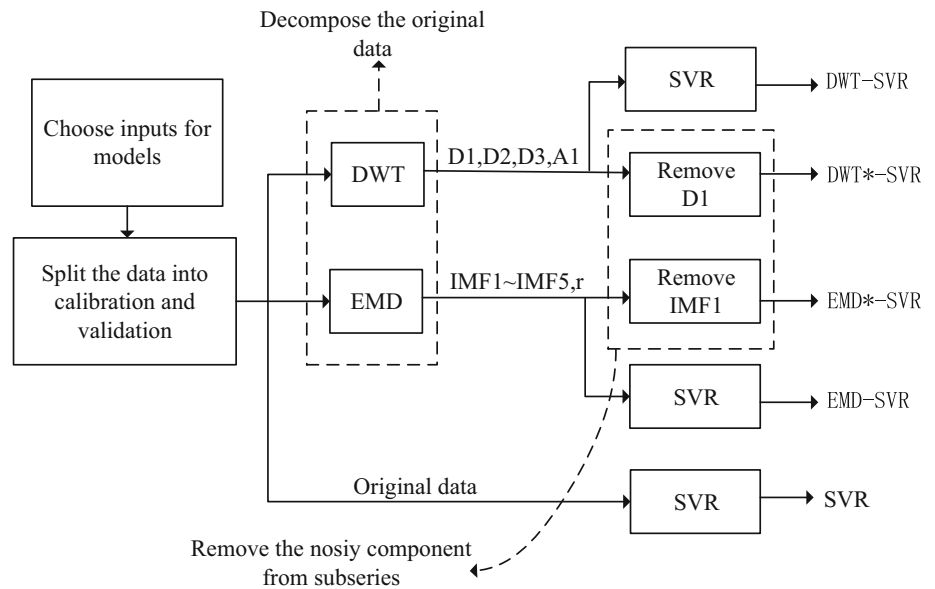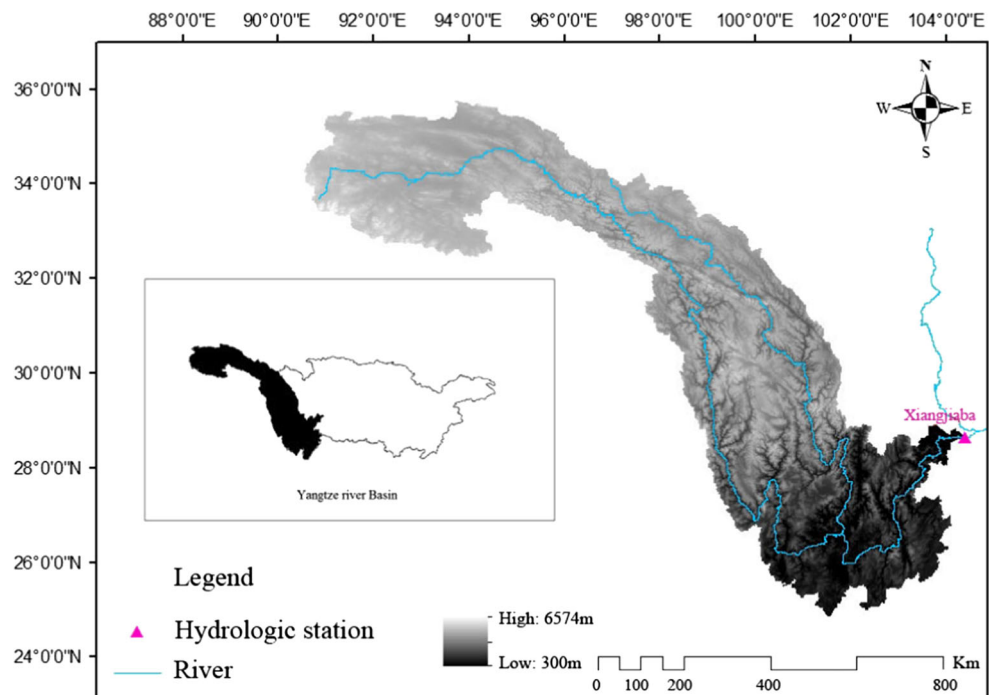Fig. 1 Working structures of various models

Fig. 2 Location of Jinsha River basin and Xiangjiaba hydrologic station

Hybrid decomposing- SVR model

Decomposing-SVR model is a hybrid SVR model coupled with time series decomposition technique. Two types of decomposing technique i.e. DWT and EMD are adopted in this study. Construction of hybrid model is as follows. First, based on the auto-correlation function and the cross-correlation function of observed monthly streamflow and precipitation, appropriate inputs for models are chosen. Then observed streamflow and precipitation time series are decomposed into their sub-time series by DWT and EMD, respectively. For DWT, the sub-time series D1, D2 and D3 represent detail components corresponding to 2-, 4- and 8-months scale or periodicity, and A3 represents approximation component of 8-months scale. For EMD, the sub-time series IMF1 ∼ IMF5 represent different oscillatory modes and r represents a residual after decomposition. Finally, hybrid models DWT-SVR and EMD-SVR are constructed by combining all the decomposed subseries with SVR models. D1 and IMF1 are called the 'noise' components, which are the most fluctuating and uncorrelated in original hydrology time series. To find out the influence of high-frequency components on streamflow forecasting, DWT*-SVR and EMD*-SVR models are established by removing D1 and IMF1 from the decomposed series. The working structures of DWT-SVR, DWT*SVR, EMD-SVR, EMD*-SVR and single SVR are given in Fig. 1.

Application and results

Study area and data division

Jinsha River watershed is taken as the study area in this research. Jinsha River is in the upper reaches of Yangtze River and it originates in the northwest of Yunnan plateau. The total length is 2316 km and the water catchment area is 340,000 km^2. Run-off volume of the river is abundant with an average annual flow of 149.8 billion m^3. It features a big vertical drop of 3300 meters and hydropower resources of Jinsha River are rich, about 1.12 billion kW. 25 hydropower stations are in planning and 4 huge

Fig. 3 Monthly streamflow and rainfall between January 1961 and December 2008
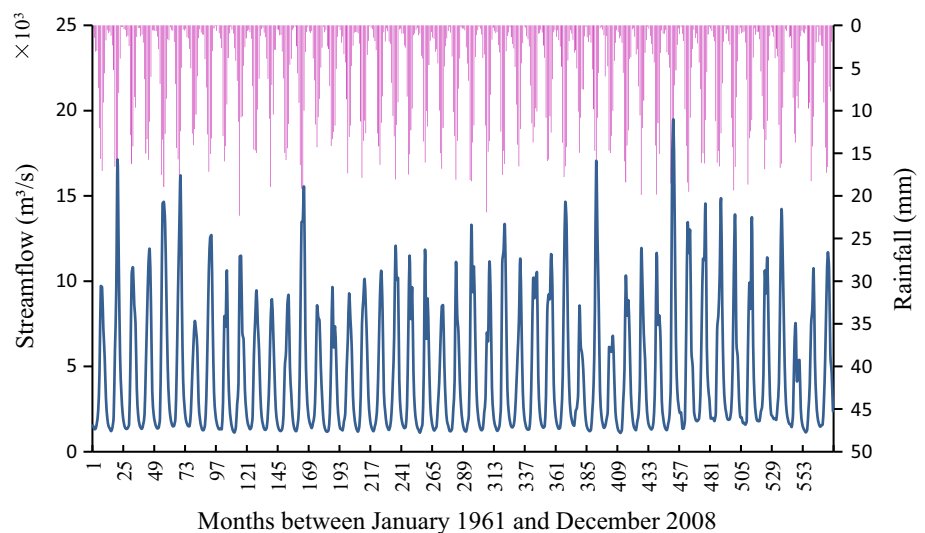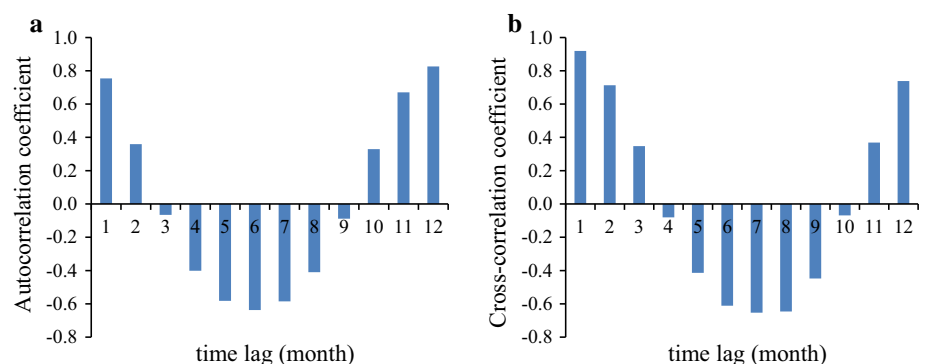

Fig. 4 a Auto-correlation coefficients for streamflow; **b** Cross-correlation coefficient for streamflow and rainfall

hydropower plants are already under construction. Monthly streamflow forecasting of Xiangjiaba, which is located in the basin outlet of Jinsha River, is researched in this study. Location of Jinsha River basin and Xiangjiaba station is shown in Fig. 2.

Data consist of run-off records from Xiangjiaba hydrometric station, and rainfall records from 32 meteorological stations in Jinsha River basin. To reduce the number of the model inputs and the complexity of the model structure, the rainfall used is the mean of the above 32 rainfall records. 48 years (January 1961 to December 2008) of the streamflow and rainfall data is plotted in Fig. 3. Data from the year of 1961–1997 (444 months) are used for model training and remaining data from 1998 to 2008 (132 months) are used for model validation.

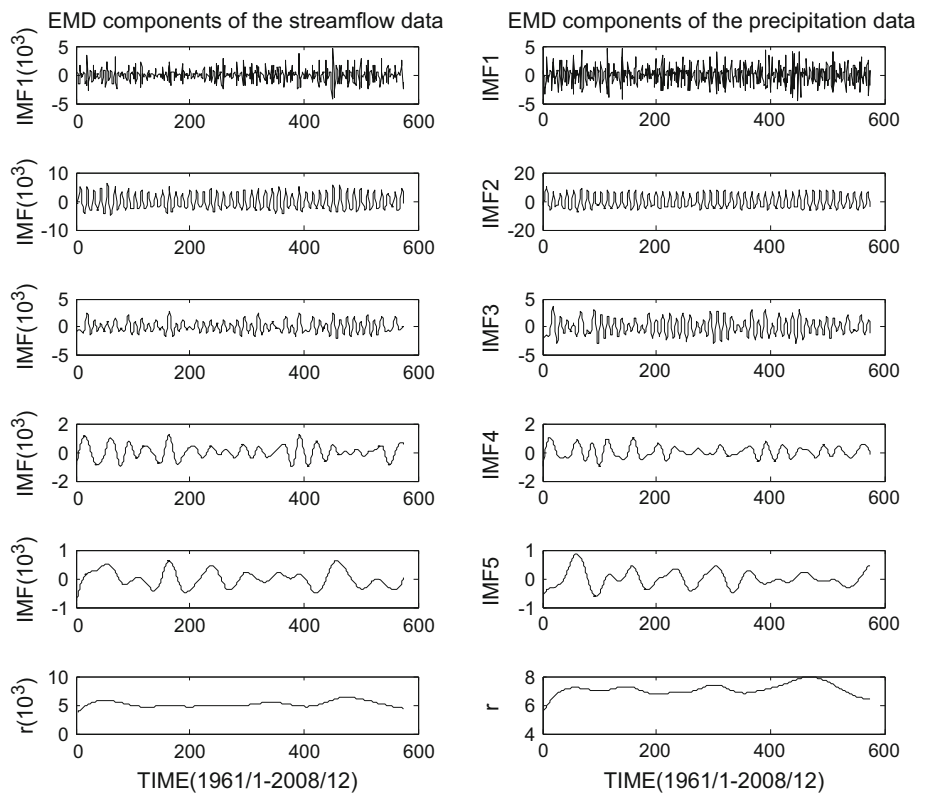

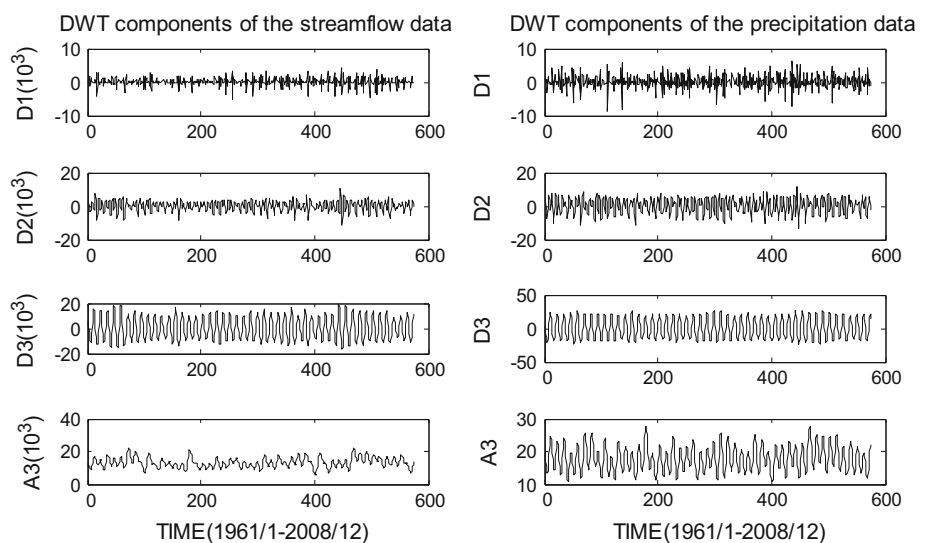**Fig. 5** Sub-time series components decomposed by EMD

Fig. 6 Sub-time series components decomposed by DWT

Selection of model inputs

Selection of model inputs is an important phase in model calibration and it extremely affects the precision of the model. Generally, for the sake of simplicity, streamflow is the only input factor in many researches (Dehghani et al. 2015; Wu and Chau 2010; Yilmaz and Muttil 2013). In this paper, the influence of precipitation on the streamflow forecasting will be revealed by adding precipitation to the model. A result about the auto-correlation function and the cross-correlation function of streamflow and precipitation is shown in Fig. 4. Based on the analysis, the most correlated variables Q_{t-1}, Q_{t-11}, Q_{t-12} and P_{t-1}, P_{t-2}, P_{t-12} are selected as input variables of models, where Q_{t-1}, Q_{t-11}, Q_{t-12} stand for streamflow at 1, 11, 12 months ago, respectively; P_{t-1}, P_{t-2}, P_{t-12} stand for precipitation at 1, 11, 12 months ago, respectively.

Decomposition of the original data

The original data are, respectively, decomposed by DWT and EMD. EMD doesn't need any predetermined basis functions, therefore five IMFs and one residual are generated directly from the initial series, which are displayed in Fig. 5. As for DWT decomposing procedures, subseries are obtained using fast Mallat algorithm on three decomposition levels, which is recommended by Aussem and Murtagh (1998) and Nourani et al. (2009). Selecting an appropriate mother wavelet is critical to obtain a better wavelet hybrid model. Maheswaran and Khosa (2012) pointed out that wavelets with wider support and higher vanishing moments are more suitable for irregular and oscillating hydrology series. Hence, an irregular mother wavelet, the Daubechies wavelet with five vanishing moments (db5), which has very high number of vanishing moments for a given support width, is selected. Three decomposition levels series of approximations (A) and

details (D) through the high-pass and low-pass filter coefficients of the chosen db5 are displayed in Fig. 6, and each subseries may represent a special level of the temporal characteristics of the original time series. It can be seen from Figs. 5 and 6 that the high frequency decomposed components IMF1 and D1 are the most non-linear and disorder.

Model development

SVR, EMD-SVR, DWT-SVR, EMD*-SVR, DWT*-SVR discussed in "Methodology" section are evaluated for forecasting one-month-ahead streamflow in the Jinsha River of China. Based on different input combinations, eight models SVR-Q, SVR-QP, DWT-SVR-Q, DWT-SVR-QP, EMD-SVR-Q, EMD-SVR-QP, DWT*-SVR-QP and EMD*-SVR-QP are developed eventually in this research. SVR-Q has original streamflow with leading time of 1, 11, 12 months as inputs, DWT-SVR-Q has D1, D2, D3, A3 series (Fig. 5) produced by DWT decomposition of original streamflow as inputs. DWT-SVR-QP has D1, D2, D3, A3 series (Fig. 5) produced by DWT decomposition of original streamflow and precipitation as inputs. DWT*-SVR-QP has D2, D3, A3 series produced by DWT decomposition of original streamflow and precipitation as inputs. For EMD-SVR-Q, EMD-SVR-QP and EMD*-SVR-QP, the difference to above is that inputs are taken from subseries produced by EMD rather than DWT. The desired output of all models is one-month-ahead streamflow. All models and the corresponding inputs are provided in Table 1. In Table 1, Q_{t-1} and P_{t-1} denote original streamflow and precipitation at time t-1 respectively; Q_{t-1} (IMF1–IMF5, r) denote EMD components IMF1, IMF2, IMF3, IMF4, IMF5 and residue, and Q_{t-1} (D1–D3, A3) denote DWT components D1, D2, D3 and A3.

Based on the input–output pairs in training period, models will be separately optimized. When the

Table 1 Models and their corresponding inputs

Model	Input variables
SVR-Q	Q_{t-1}, Q_{t-11}, Q_{t-12}
SVR-QP	Q_{t-1}, Q_{t-11}, Q_{t-12}, P_{t-1}, P_{t-2}, P_{t-12}
EMD-SVR-Q	Q_{t-1} (IMF1–IMF5, r), Q_{t-11} (IMF1–IMF5, r), Q_{t-12} (IMF1–IMF5, r)
EMD-SVR-QP	Q_{t-1} (IMF1–IMF5, r), Q_{t-11} (IMF1–IMF5, r), Q_{t-12} (IMF1–IMF5, r), P_{t-1} (IMF1–IMF5, r), P_{t-2} (IMF1–IMF5, r), P_{t-12} (IMF1–IMF5, r)
EMD*-SVR-QP	Q_{t-1} (IMF2–IMF5, r), Q_{t-11}(IMF2–IMF5, r), Q_{t-12} (IMF2–IMF5, r), P_{t-1} (IMF2–IMF5, r), P_{t-2} (IMF2–IMF5, r), P_{t-12} (IMF2–IMF5, r)
DWT-SVR-Q	Q_{t-1} (D1–D3, A3), Q_{t-11} (D1–D3, A3), Q_{t-12} (D1–D3, A3)
DWT-SVR-QP	Q_{t-1} (D1–D3, A3), Q_{t-11} (D1–D3, A3), Q_{t-12} (D1–D3, A3), P_{t-1} (D1–D3, A3), P_{t-2} (D1–D3, A3), P_{t-12} (D1–D3, A3)
DWT*-SVR-QP	Q_{t-1} (D2, D3, A3), Q_{t-11} (D2, D3, A3), Q_{t-12} (D2, D3, A3), P_{t-1} (D2, D3, A3), P_{t-2} (D2, D3, A3), P_{t-12} (D2,D3, A3)

optimization phase is done, all models will be applied to forecast monthly streamflow using the data from the historical series in validation period. The data are decomposed in advance by EMD or DWT based on different needs. For single SVR models, original series are directly used as inputs, for DWT-SVR hybrid models, original series will be decomposed by DWT and for EMD-SVR hybrid models, original series will be decomposed by EMD.

To improve the efficiency of training, model inputs and the desired output are conveniently standardized and scaled to the range [0, 1]. Model parameters are optimized by cross-validation (CV). CV is a standard technique for training SVR model. Typical fivefold cross-validation is adopted here.
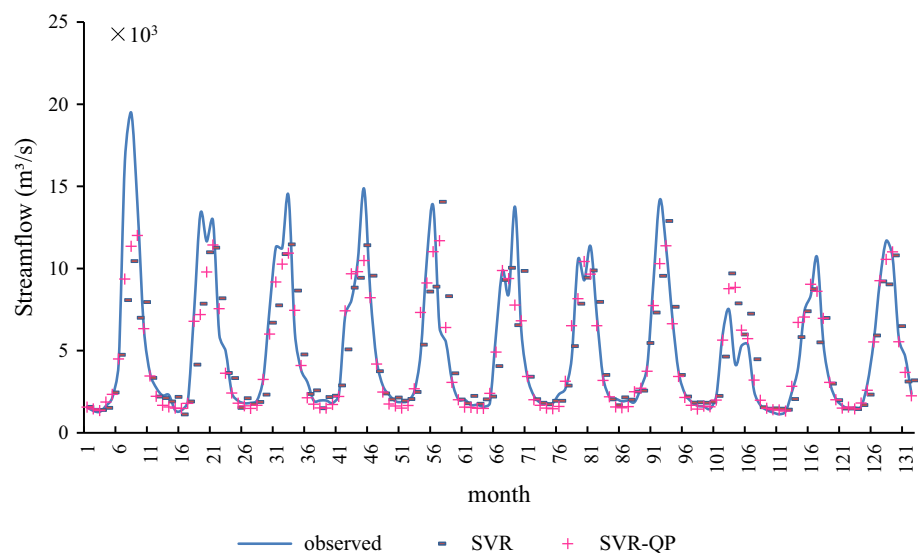
Performance evaluation

The evaluation indicators of root mean square errors (RMSE), mean absolute errors (MAE), mean absolute percentage error (MAPE), correlation coefficient (R) and run time (RT) are employed to evaluate the accuracy and time cost of models. RMSE assesses the goodness of the fit related to high streamflow values whereas MAE measures a more balanced perspective of the fitness at moderate streamflows. R shows the degree which two variables are linearly related to. RT stands for time cost of model. RMSE, MAE and MAPE are defined as:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (Y_i^{obs} - Y_i^{est})^2} \qquad (10)$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left| Y_i^{obs} - Y_i^{est} \right| \qquad (11)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{Y_i^{obs} - Y_i^{est}}{Y_i^{obs}} \right| \times 100\% \qquad (12)$$

where N denotes the number of datasets, Y_i^{obs} represents the observed monthly streamflow. Y_i^{est} represents the estimated monthly streamflow.

Results analysis

Results of SVR models, SVR models combined with EMD and SVR models combined with DWT in validation period (January 1998 to December 2008) are shown in Figs. 7, 8 and 9, respectively. RMSE, MAE, MAPE, R and run time (RT) in the calibration and validation period are, respectively, given in Tables 2, 3 and 4. In view of indices RMSE, MAE, MAPE and R, Table 2 indicates that SVR-QP model, in which precipitation and streamflow as inputs, has a better accuracy than the SVR-Q model that only adopts streamflow as inputs. Table 3 shows that the performance of EMD-SVR-QP is superior to EMD-SVR-Q and proves again that the use of precipitation can improve modeling precision. Then EMD-SVR-QP and EMD*-SVR-QP are compared, it is found that removing the high frequency component IMF1 is inappropriate for a better accuracy. Results of DWT-SVR models are given in Table 4 and the same conclusions can be made for DWT-SVR models. In view of RT indices, it can be concluded from Tables 2, 3 and 4 that introducing precipitation data into model inputs will greatly influence the time cost in model calibration. Generally, more inputs lead to more cost of run time. However, once models have been calibrated, run time in model validation or forecast will be almost same.

Fig. 7 Monthly streamflow estimations of single SVR models in the validation period

Fig. 8 Monthly streamflow estimations of SVR models combined with EMD in the validation period
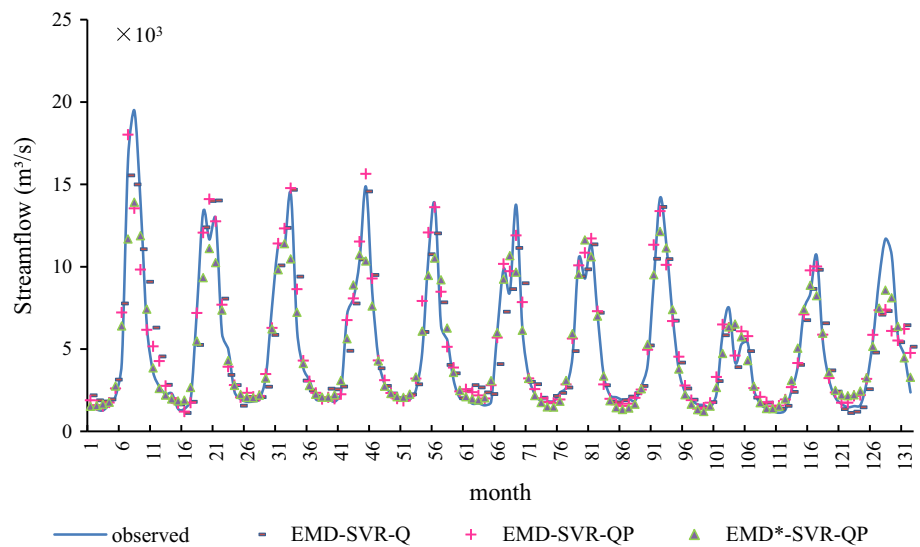

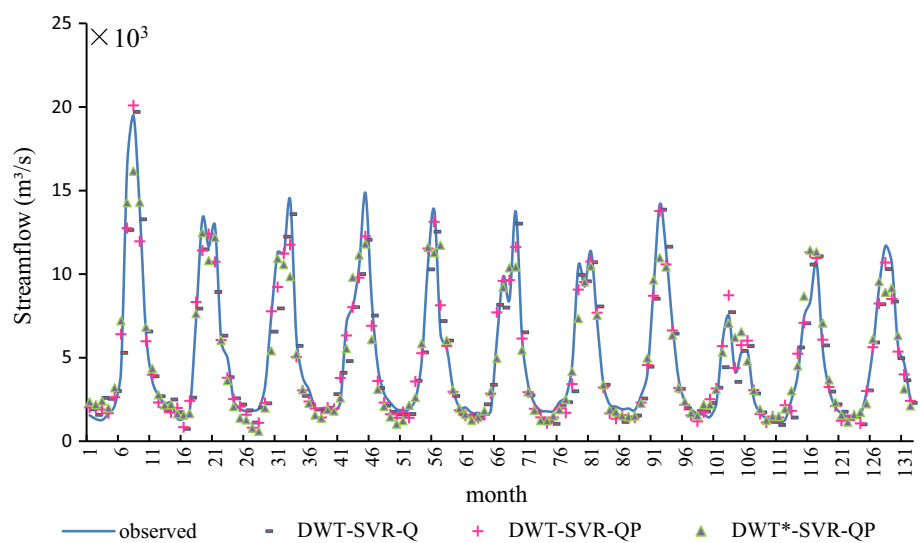**Fig. 9** Monthly streamflow estimations of SVR models combined with DWT in the validation period

Table 2 RMSE, MAE, MAPE, R and RT of the single SVR models

Model	Calibration					Validation				
	RMSE	MAE	MAPE (%)	R	RT(s)	RMSE	MAE	MAPE (%)	R	RT(s)
SVR-Q	1305	706	13	0.88	23.00	2114	1185	20	0.74	0.02
SVR-QP	1257	708	13	0.88	20.80	1693	908	16	0.84	0.02

Table 3 RMSE, MAE, MAPE, R and RT of the EMD-SVR based models

Model	Calibration					Validation				
	RMSE	MAE	MAPE (%)	R	RT(s)	RMSE	MAE	MAPE (%)	R	RT(s)
EMD-SVR-Q	512	324	8	0.98	32.06	1278	840	19	0.9	0.02
EMD-SVR-QP	384	244	7	0.99	42.27	1213	748	17	0.91	0.02
EMD*-SVR-QP	895	468	9	0.94	42.09	1373	863	18	0.91	0.02

Table 4 RMSE, MAE, MAPE, *R* and RT of the DWT-SVR based models

Model	Calibration					Validation				
	RMSE	MAE	MAPE (%)	R	RT(*s*)	RMSE	MAE	MAPE (%)	*R*	RT(*s*)
DWT-SVR-Q	349	224	7	0.99	24.67	1015	715	18	0.94	0.02
DWT-SVR-QP	365	233	7	0.99	33.96	951	689	15	0.95	0.03
DWT*-SVR-QP	875	491	10	0.94	32.72	1275	874	19	0.91	0.02

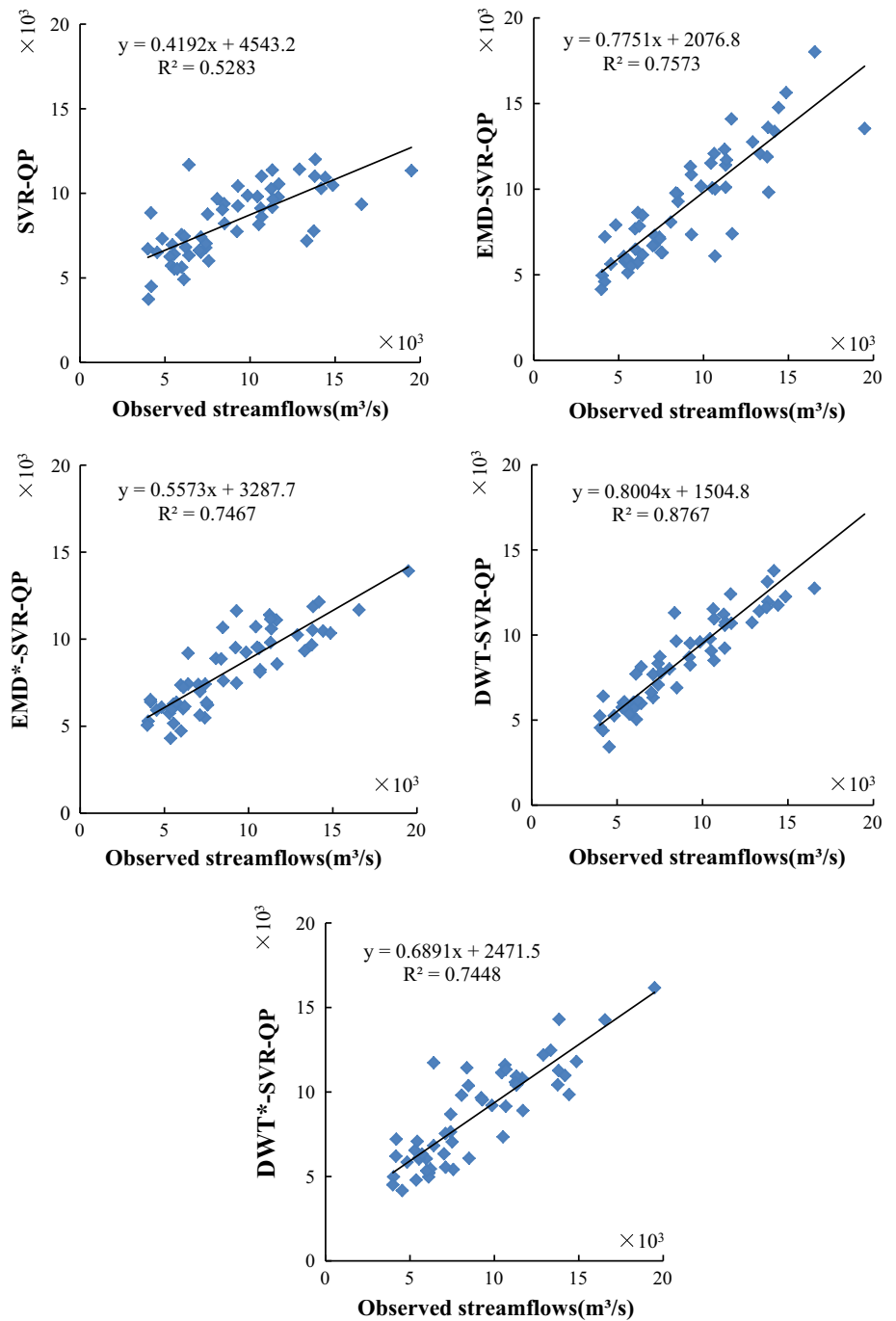**Fig. 10** Monthly streamflow estimations in the flood season

Table 5 Qualified rate of forecasting in the flood season (May–October)

Model	SVR-QP	EMD-SVR-QP	EMD*-SVR-QP	DWT-SVR-QP	DWT*-SVR-QP
Qualified rate (%)	67	73	56	85	65

Then a comprehensive analysis needs to be made to reveal the effect of different decomposing techniques on model accuracy based on Table 2, 3 and 4. Compare SVR-QP, EMD-SVR-QP and DWT-SVR-QP, all of which are optimum models in each group, it can be acquired that the models combined with decomposition techniques perform much better than single SVR model and DWT-SVR-QP improves the accuracy of prediction more highly than EMD-SVR-QP in both calibration and validation periods. Therefore, using time series decomposing techniques contributes to improving performance of forecasting, and decomposing technique DWT is more suitable than EMD for monthly streamflow modeling.

Furthermore, due to the reason that streamflow in flood season severely fluctuates, high-precision forecasting in that period is very challenging. An evaluation of the flood season (May–October) streamflow forecasting is discussed here. Results of SVR-QP, EMD-SVR-QP and EMD*-SVR-QP, DWT-SVR-QP and DWT*-SVR-QP in the validation period are illustrated in Fig. 10. It can be recognized that estimations of EMD-SVR-QP and DWT-SVR-QP approximate the observed streamflows better than single SVR model and hybrid models removing high frequency components. 20 % is considered as a reasonable and acceptable relative error in this study. Percentage of acceptable predictions, which is called qualified rate, is shown in Table 5. The results indicate that DWT-SVR-QP raises the qualified rate from 67 % of SVR-QP to 85 %. It can be seen from the table that DWT-SVR-QP achieves the greatest forecasting ability for flood season streamflow. Removing high-frequency components from original sub-series leads to an obvious reduction in forecast performance. The reason may be that streamflows in the flood season are extremely fluctuant and the high-frequency components contain the important information which is indispensable to estimate fluctuations.

Conclusions

Accuracy of monthly streamflow forecasting models coupled with different decomposition techniques was investigated in this paper. SVR, EMD-SVR, and DWT-SVR based models were, respectively, obtained by single support vector regression, SVR, combined with empirical mode decomposition, and SVR combined with discrete wavelet transform. To research the influence of precipitation on the forecast accuracy, all models were implemented with two kinds of inputs depending on whether antecedent precipitation was included. All models were applied to Xiangjiaba to perform one-month-ahead streamflow forecasting. Results can be summarized as follows:

1. All of the models that add antecedent precipitation in inputs exhibit a significant improvement, so a more excellent model can be built when precipitation information is taken into account.
2. Removing the high-frequency component from the subseries fails to improve the forecasting ability, especially for the forecasting in flood season. This is because sub-time series of first or second decomposition levels show major information on extreme flows. Therefore high-frequency component is indispensable to monthly streamflow forecasting.
3. Comparison results of SVR-QP, EMD-SVR-QP and DWT-SVR-QP models show that SVR-QP produces the worst performance, EMD and DWT both could significantly increase accuracy of monthly streamflow forecasting. Meanwhile, it can be acquired that DWT outperforms EMD in terms of the evaluation indices of RMSE, MAE, MAPE, R and RT.
4. For the flood season forecasting, SVR models combining EMD and DWT both raise the qualified rate based on the results from May to October. DWT-SVR-QP is superior to EMD-SVR-QP with a highest qualified rate of 85 %.

In conclusion, results in this paper indicate that models coupled with decomposition techniques perform better than the single models and decomposition technique DWT provides a superior alternative to EMD in monthly streamflow forecasting. Among all the developed models, DWT-SVR-QP which combining discrete wavelet transform and support vector regression, has the best performance and is recommended as an alternative to Xiangjiaba monthly streamflow forecasting.

References

Aussem ACJ, Murtagh F (1998) Wavelet-based feature extraction and decomposition strategies for financial forecasting. J Comput Intell Finance 6:7

Budu K (2013) Comparison of wavelet-based ANN and regression models for reservoir inflow forecasting. J Hydrol Eng 19:1385–1400. doi:10.1061/(ASCE)HE.1943-5584.0000892

Chen L, Singh V, Guo S, Zhou J, Ye L (2014) Copula entropy coupled with artificial neural network for rainfall–runoff simulation. Stoch Environ Res Risk Assess 28:1755–1767. doi:10.1007/s00477-013-0838-3

Dehghani M, Saghafian B, Rivaz F, Khodadadi A (2015) Monthly stream flow forecasting via dynamic spatio-temporal models. Stoch Environ Res Risk Assess 29:861–874. doi:10.1007/s00477-014-0967-3

Guo J, Zhou JZ, Qin H, Zou Q, Li QQ (2011) Monthly streamflow forecasting based on improved support vector machine model. Expert Syst Appl 38:13073–13081. doi:10.1016/j.eswa.2011.04.114

Hsu K-L, Gupta HV, Sorooshian S (1995) Artificial neural network modeling of the rainfall-runoff process. Water Resour Res 31:2517–2530. doi:10.1029/95WR01955

Huang NE, Wu Z (2008) A review on Hilbert-Huang transform: method and its applications to geophysical studies. Rev Geophys 46:RG2006. doi:10.1029/2007RG000228

Huang NE et al (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc Lond Ser A Math Phys Eng Sci 454:903–995. doi:10.1098/rspa.1998.0193

Huang SZ, Chang JX, Huang Q, Chen YT (2014) Monthly streamflow prediction using modified EMD-based support vector machine. J Hydrol 511:764–775. doi:10.1016/j.jhydrol.2014.01.062

Karthikeyan L, Kumar DN (2013) Predictability of nonstationary time series using wavelet and EMD based ARMA models. J Hydrol 502:103–119. doi:10.1016/j.jhydrol.2013.08.030

Kasiviswanathan KS, Sudheer KP (2013) Quantification of the predictive uncertainty of artificial neural network based river flow forecast models. Stoch Environ Res Risk Assess 27:137–146. doi:10.1007/s00477-012-0600-2

Labat D (2005) Recent advances in wavelet analyses: part 1. Rev Concepts J Hydrol 314:275–288. doi:10.1016/j.jhydrol.2005.04.003

Lee T, Ouarda TBMJ (2010) Long-term prediction of precipitation and hydrologic extremes with nonstationary oscillation processes. J Geophys Res Atmos. doi:10.1029/2009jd012801

Lee T, Ouarda TBMJ (2011) Prediction of climate nonstationary oscillation processes with empirical mode decomposition. J Geophys Res Atmos 116:D06107. doi:10.1029/2010JD015142

Lin GF, Chen GR, Huang PY, Chou YC (2009) Support vector machine-based models for hourly reservoir inflow forecasting during typhoon-warning periods. J Hydrol 372:17–29. doi:10.1016/j.jhydrol.2009.03.032

Maheswaran R, Khosa R (2012) Comparative study of different wavelets for hydrologic forecasting. Comput Geosci Uk 46:284–295. doi:10.1016/j.cageo.2011.12.015

Napolitano G, Serinaldi F, See L (2011) Impact of EMD decomposition and random initialisation of weights in ANN hindcasting of daily stream flow series: an empirical examination. J Hydrol 406:199–214. doi:10.1016/j.jhydrol.2011.06.015

Nourani V, Komasi M, Mano A (2009) A multivariate ANN-wavelet approach for rainfall-runoff modeling. Water Resour Manage 23:2877–2894. doi:10.1007/s11269-009-9414-5

Rahman K, Etienne C, Gago-Silva A, Maringanti C, Beniston M, Lehmann A (2014) Streamflow response to regional climate model output in the mountainous watershed: a case study from the Swiss Alps Environ. Earth Sci 72:4357–4369. doi:10.1007/s12665-014-3336-0

Sahay RR, Srivastava A (2014) Predicting monsoon floods in rivers embedding wavelet transform. Genet Algorithm Neural Netw Water Resour Manage 28:301–317. doi:10.1007/s11269-013-0446-5

Sattari MT, Apaydin H, Ozturk F (2012) Flow estimations for the Sohu Stream using artificial neural networks. Environ Earth Sci 66:2031–2045. doi:10.1007/s12665-011-1428-7

Sehgal V, Sahay R, Chatterjee C (2014) Effect of utilization of discrete wavelet components on flood forecasting performance of wavelet based ANFIS models. Water Resour Manage 28:1733–1749. doi:10.1007/s11269-014-0584-4

Tiwari M, Adamowski J (2014) Medium-term urban water demand forecasting with limited data using an ensemble wavelet-bootstrap machine-learning approach. J Water Resour Plan Manage 141:04014053. doi:10.1061/(ASCE)WR.1943-5452.0000454

Vapnik VN (1995) The nature of statistical learning theory. Springer, New York

Wang W, Ding J (2003) Wavelet network model and its application to the prediction of hydrology. Science 1:67–71

Wang EL, Zhang YQ, Luo JM, Chiew FHS, Wang QJ (2011) Monthly and seasonal streamflow forecasts using rainfall-runoff modeling and historical weather data. Water Resour Res. doi:10.1029/2010wr009922

Wang W, Chau K, Qiu L, Chen Y (2015) Improving forecasting accuracy of medium and long-term runoff using artificial neural network based on EEMD decomposition. Environ Res 139:46–54. doi:10.1016/j.envres.2015.02.002

Wei Z, Tao T, ZhuoShu D, Zio E (2013) A dynamic particle filter-support vector regression method for reliability prediction. Reliab Eng Syst Saf 119:109–116. doi:10.1016/j.ress.2013.05.021

Wu CL, Chau KW (2010) Data-driven models for monthly streamflow time series prediction. Eng Appl Artif Intell 23:1350–1367. doi:10.1016/j.engappai.2010.04.003

Yarar A (2014) A hybrid wavelet and neuro-fuzzy model for forecasting the monthly streamflow data. Water Resour Manage 28:553–565. doi:10.1007/s11269-013-0502-1

Yilmaz A, Muttil N (2013) Runoff estimation by machine learning methods and application to the euphrates basin in Turkey. J Hydrol Eng 19:1015–1025. doi:10.1061/(ASCE)HE.1943-5584.0000869