Determination of variation uncertainty in runoff time series at multi-temporal scales

Yan Ye, Jinping Zhang, Xunjian Long, Lihua Ma and Yong Ye

ABSTRACT

In order to survey the possible periodic, uncertainty and common features in runoff with multi-temporal scales, the empirical mode decomposition (EMD) method combined with the set pair analysis (SPA) method was applied, with data observed at Zhangjiashan hydrological station. The results showed that the flood season and annual runoff time series consisted of four intrinsic mode function (IMF) components, and the non-flood season time series exhibited three IMF components. Moreover, based on the different coupled set pairs from the time series, the identity, discrepancy, and contrary of different periods at multi-temporal scales were determined by the SPA method. The degree of connection μ between the flood season and annual runoff periods were the highest, with 0.94, 0.77, 0.7 and 0.73, respectively, and the μ between the flood periods and the non-flood periods were the lowest, with 0.66, 0.46, 0.24 and 0.24, respectively. Third, the maximum μ of each SPA appeared in the first mode function. In general, the different extractive periods decomposed by the EMD method reflected the average state of Jinghe River. Results also verified that runoff suffered from seasonal and periodic fluctuations, and fluctuations in the short-term corresponded to the most important variable. Therefore, the conclusions drawn in this study can improve water resources regulation and planning.

Key words | empirical mode decomposition, multi-temporal scales, periodic fluctuation, runoff, set pair analysis

Yan Ye
Xunjian Long (corresponding author)
Lihua Ma
Yong Ye
College of Resources and Environment,
Southwest University,
Chongqing 400715,
China
E-mail: lmcx402@swu.edu.cn

Jinping Zhang
Institute of Water Resources and Environment,
Zhengzhou University,
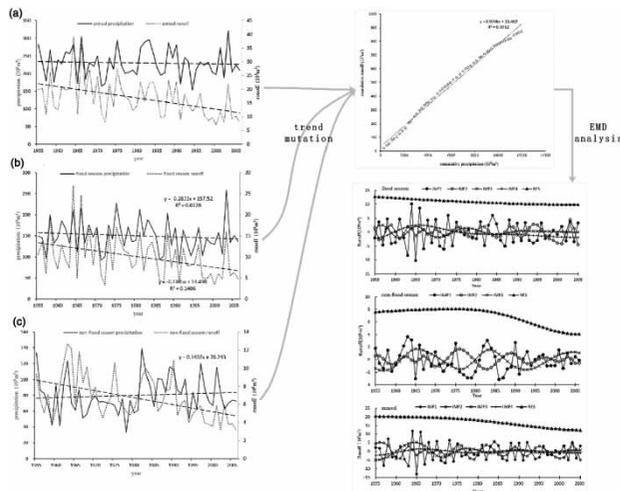Zhengzhou 450001,
China

HIGHLIGHTS

- A coupling model of empirical mode decomposition and set pair analysis is proposed.
- Mutation and variation of runoff under different time scales is provided.
- Trend components and fluctuation features under multi-temporal scales are explored.

GRAPHICAL ABSTRACT

INTRODUCTION

River runoff trends change with climatic conditions. In recent years, research on runoff variability is focusing more on climate factors because runoff has been recognized as an important hydrological phenomenon, also exhibiting long-term and complex variability (Zhao *et al.* 2017). Certainly, it is heavily influenced by several factors, such as precipitation, evaporation, and human activities (Yilinuer *et al.* 2021). A general consensus is that runoff processes tend to be alternating, but random (Jeon *et al.* 2014). In order to better understand the nonlinear, non-stationary frequency components of runoff time series, numerous hydrological models and theoretical methods have been developed and applied (Chau 2017; Birkel & Barahona 2019).

In the past two decades, many models and techniques have been applied to analyze runoff time series. The following two main research directions have been widely discussed: frequency and trends analysis, and prediction or real-time forecasting. These are self-governed and inter-related at the same time (Yu *et al.* 2001). Forecasting is useful to understand the temporal and spatial transformations in a hydrological time series. Various researchers have recommended prediction due to its scientific and rigorous method. These include the artificial neural network model (Nourani 2017), support vector machine

(He *et al.* 2014), and sample entropy (Chou 2014). However, these simulation studies should be based on basic trend analysis or implicated in small sample cases. It could also be argued that models for runoff forecasting are unreliable without frequency and trends calculations. To better understand the underlying patterns and trends, several mathematical methods, such as the Mann–Kendall trend test, regression analysis, and exponential smoothing method (Lloyd & Tommy 2011), as well as the autoregressive integrated moving average (ARIMA) method, ARIMA with exogenous input (ARIMAX), multiple linear regression (Adamowski & Chan 2011; Zhang *et al.* 2011) and wavelet analysis (Nourani *et al.* 2013) have been recommended.

The wavelet analysis method is more commonly used in the analysis of hydrological time series at multi-temporal scales (Chou 2011; Potocki *et al.* 2017; Roushangar *et al.* 2018). Using this method, Chou (2011) applied the wavelet transform to determine the number of resolution levels for a rainfall time series. Roushangar *et al.* (2018) developed a wavelet-based de-noising approach to reduce the effects of noise in a precipitation time series. Although higher resolution was obtained via the wavelet analysis method both in the frequency and temporal domains, many false harmonics still existed as a result of its limitations (Rhif *et al.* 2019).

Moreover, the selection of wavelet base functions significantly impacted the results (Yan 2013).

In order to overcome the shortcomings of runoff time series, such as non-stationary, harmonics or false information during the simulation process, Huang *et al.* (1998, 1999) developed a new signal analysis method called the empirical mode decomposition (EMD). EMD can decompose an original signal into fluctuations or trends at various scales (or frequencies) and produce a stationary data series with different dimensions called the intrinsic mode function (IMF). The IMF can reduce processing complexity and is more effective than the wavelet and almost all the other previous decomposition methods, in particular in the local frequency and time domains (Kim & Oh 2009; Lee & Ouarda 2011; Zhang *et al.* 2017a, 2017b). Thus, EMD is adaptive, empirical, direct, and intuitive. It has been applied to hydrological phenomena with good results (Giulia *et al.* 2011; Karthikeyan & Kumar 2013; Zhang *et al.* 2014).

The complexity of change in runoff mainly lies in its uncertainty. To determine runoff uncertainty, it is necessary to carry out an in-depth study on time series relationships. To resolve difficult and inconvenient calculations by using the mathematical method (Adamowski & Chan 2011; Giulia *et al.* 2011), Zhao & Xuan (1989) first developed a set pair analysis (SPA) method to measure the uncertainty. SPA describes three aspects of an uncertainty relationship between variables: (1) identity, (2) discrepancy, and (3) contrary. It can be applied to deal with the interaction between certainty and uncertainty. As Zhang *et al.* (2019) implicated in water resources, it is suitable for cases with two or more groups from two or more perspectives. Therefore, it has been commonly used in hydrological analyses, such as annual runoff variability (Yu *et al.* 2018); evaluating hydrological forecasting models (Pan *et al.* 2017); analyzing relationships between flood peak and volume (Zhu *et al.* 2007), annual runoff, and classifications (Wang *et al.* 2008). These cases show that SPA is more effective, objective, and reasonable for hydrological uncertainty correlation analysis.

Accordingly, EMD can be used to determine the internal trends of the data series, and SPA can be used to determine the relationships between the data series. Thus, based on the basic rules of the single algorithm of the EMD and SPA, the advantageous point is that they can be combined to accurately determine runoff uncertainty features. Some earlier studies (Yu *et al.* 2018) determined the uncertainty in runoff by decomposing the series with EMD or an improved EMD, and compared the series with SPA. However, these studies did not focus on the different temporal intervals of the data series. Moreover, even the common features were hidden among multiple time series. Furthermore, in order to analyze, predict, and evaluate the phenomena, it is important to explore the uncertainty relationships between different sets in a given pair (Jin *et al.* 2017). In light of the preceding literature review, this study aimed to determine the uncertainties in the flood season, non-flood season, and annual runoff time series at Zhangjiashan hydrological station (ZHS) located at the outflow of Jinghe River in China from 1955 to 2006. The time series were decomposed into multi-temporal scales by EMD method to obtain the fluctuation components for different periods. Then, the SPA method was applied to investigate the uncertainty and the potential interrelationships in the fluctuation components and precipitation with runoff for different periods. Under these time scales, similar periodic and fluctuation trends were observed within each IMF component. On the other hand, the results of multiple SPA patterns showed a high connection degree manifested within short-term relationships. Moreover, via this coupled model, a better way to solve the problem of runoff data scarcity is encountered in hydrological research, engineering design, and water resources management.

DATA SERIES AND METHODS

Data series

Weihe River is one of the tributaries of the Yellow River in China. Jinghe River is the largest tributary of Weihe River, as shown in Figure 1. It originates from Liupanshan Mountain in Ningxia Hui Autonomous Region of China, and is 451 km long with a catchment area of 45,400 km². Jinghe River is known for its violent flooding with a large amount of sediment. The average annual runoff is 1.75 billion m³, and average annual sediment transport is 0.31 billion tons. As such, it experiences major floods and is the source of sediment for the Weihe and Yellow River.
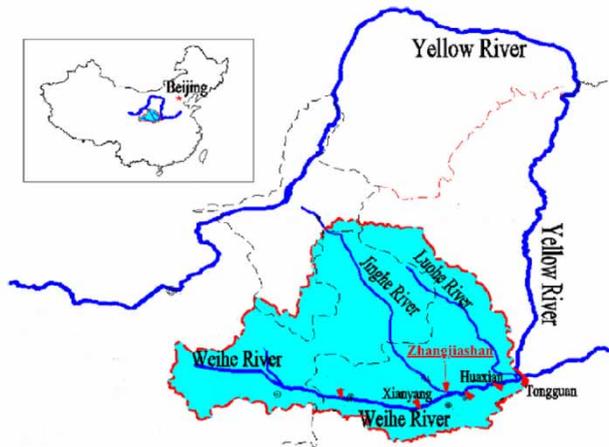
Figure 1 │ Location of ZHS.

ZHS is the main control station for the lower reach of Jinghe River and is located where the river flows into Weihe River. In this study, the homogeneity and reliability of the monthly and annual runoff series from 1955 to 2006 at the ZHS were tested and found to be satisfactory. Figure 2 shows the natural runoff data series at ZHS from 1955 to 2006. In addition to the calculations and the analysis of annual runoff, flood season runoff (July–October) and non-flood season runoff (November–June) were also analyzed in order to identify change features in the intra-annual distributions of runoff. To match the runoff time series, 52 years of precipitation data from 25 meteorological

stations located in the Jinghe River basin were selected for the period 1955–2006.

Methods

Statistical analysis

The analysis was based at ZHS and watershed precipitation stations for annual runoff and area averaged precipitation. In statistically significant cases, the average surface rainfall in the study catchment was calculated by using the Theil–Sen estimator (Blöschl *et al.* 2017). Trend analysis between river runoff and precipitation was performed for multiple series and the stationarity of the relationship between annual runoff and areal precipitation totals was analyzed by double mass curves (Searcy & Hardison 1960).

Empirical mode decomposition

Based on local characteristics and the temporal domain of the data, Huang *et al.* (1999) developed the EMD method, which could decompose any nonlinear and non-stationary signal into a set of IMFs, amplitude, and frequency modulated signals on separate time scales. In the EMD decomposition progress, no prior knowledge is required. According to the characteristics of the input signal, IMF is
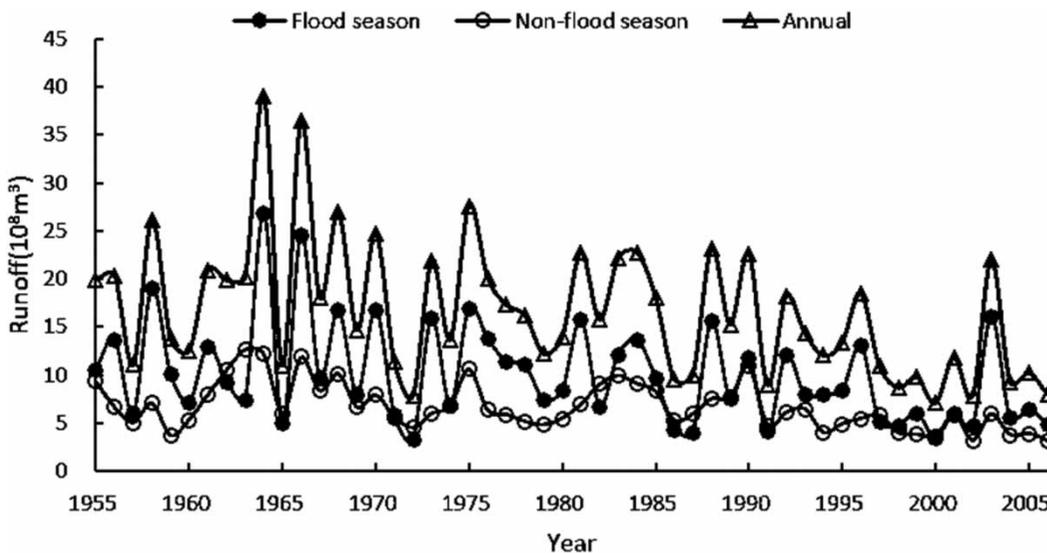

Figure 2 │ Natural runoff series at ZHS.

obtained through multiple adaptive screening decomposition and each IMF needs to be screened out several times which depends on the local mean value of the upper and lower bounds conditions. During calculation, each IMF satisfies the following two basic conditions in order to obtain a meaningful instantaneous frequency: (1) In the entire data range, the number of extreme points and the number of zero crossings must be either equal or their difference should not be greater than one; (2) At any point, the mean of the upper envelope formed by all local maxima and the lower envelope formed by all local minima is zero.

Usually, any eligible time series $x(t)$ can be decomposed into several IMFs and one final residual r_n as follows:

$$x(t) = \sum_{i=1}^{n} c_i + r_n \tag{1}$$

where c_i is the ith IMF, r_n is the final residue term, and n is the number of IMFs.

The key function of the EMD is to extract the IMF from a given time series $x(t)$. The basic steps to successively extract the IMFs are as follows (Yang *et al.* 2010; Huang *et al.* 2014):

(1) Extract all local extreme values including the local minima and maxima of $x(t)$.
(2) Interpolate the identified local maxima and minima points with a cubic spline, and construct the upper and lower envelopes e_{up} and e_{low}.
(3) Calculate the mean $m(t)$ of the two envelopes, as follows:

$$m(t) = (e_{up} + e_{low})/2 \tag{2}$$

(4) Construct the new time series:

$$h(t) = x(t) - m(t) \tag{3}$$

Usually, $h(t)$ is not an expected IMF component. Therefore, this shifting process needs to be repeated k times until the first IMF component $c_1(t) = h_{1k}(t)$ from the data $x(t)$ is achieved. The residual $r_1(t) = x(t) - c_1(t)$ gets further decomposed until it satisfies the limiting standard deviation (SD).

The criterion of the SD is defined as follows:

$$SD = \sum_{k=1}^{T} (|h_{1(k-1)}(t) - h_{1k}(t)|^2 / h_{1(k-1)}(t)^2) \tag{4}$$

where T is the length of the time series and SD lies between 0.2 and 0.3.

(5) Repeat Steps 1–4 until all IMFs are extracted:

$$r_2(t) = r_1(t) - c_2(t) \tag{5}$$

$$r_n(t) = r_{(n-1)}(t) - c_n(t) \tag{6}$$

Set pair analysis

The SPA is a method that can systematically analyze certainty and uncertainty, as well as the certainty–uncertainty interaction between two sets in a set pair for a given problem. In general, the analysis is performed by establishing the connection between the two sets under consideration (Searcy & Hardison 1960). The basic concepts of the SPA include set pairs and their degree of connection, in which the common features (identical), contrary features (contrary), and features that are neither common nor contrary in the two sets are identified.

For set pair $H(A, B)$, which consists of two interrelated sets A and B, the properties of these two sets include identity degree, discrepancy degree, and contrary degree. They are defined as follows:

$$\mu = S/N + (F/N)i + (P/N)j \tag{7}$$

where μ is the connection degree of the set pair $H(A, B)$; S is the number of identity characteristics; F is the number of discrepancy characteristics; P is the number of contrary characteristics; N is the total number of characteristics of the set pair, and $N = S + F + P$; i and j are the coefficients of the discrepancy degree F/N and the contrary P/N, respectively; i is an uncertain value under different conditions between -1 and 1, i.e. $i \in [-1, 1]$, sometimes it can be only regarded as a discrepancy marker; and j is the contradictory coefficient $j = -1$, which is determined based on each specific situation.

Given the ratios $a = S/N$(*identity degree*), $b = F/N$(*discrepancy degree*), and $c = P/N$(*contrary degree*), Equation (5) can be rewritten as follows:

$$\mu = a + bi + cj \tag{8}$$

Clearly, $a, b, c > 0$, and $a + b + c = 1$.

RESULTS

Statistical characteristics of annual precipitation and runoff

The annual runoff series for the study catchment in the study period ranged from 7.05×10^8 to $39.07 \times 10^8 \, \text{m}^3$, and precipitation varied from 153.39×10^8 to $321.92 \times 10^8 \, \text{m}^3$ (Figure 3(a)). During the study period, the precipitation time series decreased substantially except during the non-flood season, but runoff decreased in all the time series (Figure 3). The interannual variability trend was mainly driven by the flood season.

Considering the different variability trend between flood season and non-flood season, the double mass curve was applied to test the break site. The straight line pattern of the double mass curve for precipitation and runoff at ZHS showed the stationarity of the relationship between precipitation and runoff during the study period (Figure 4). Moreover, the coefficient of the curve between precipitation and runoff is 0.99. However, the curve also indicated a change in the runoff/precipitation ratio during the years 1966 and 1999 (Figure 4).
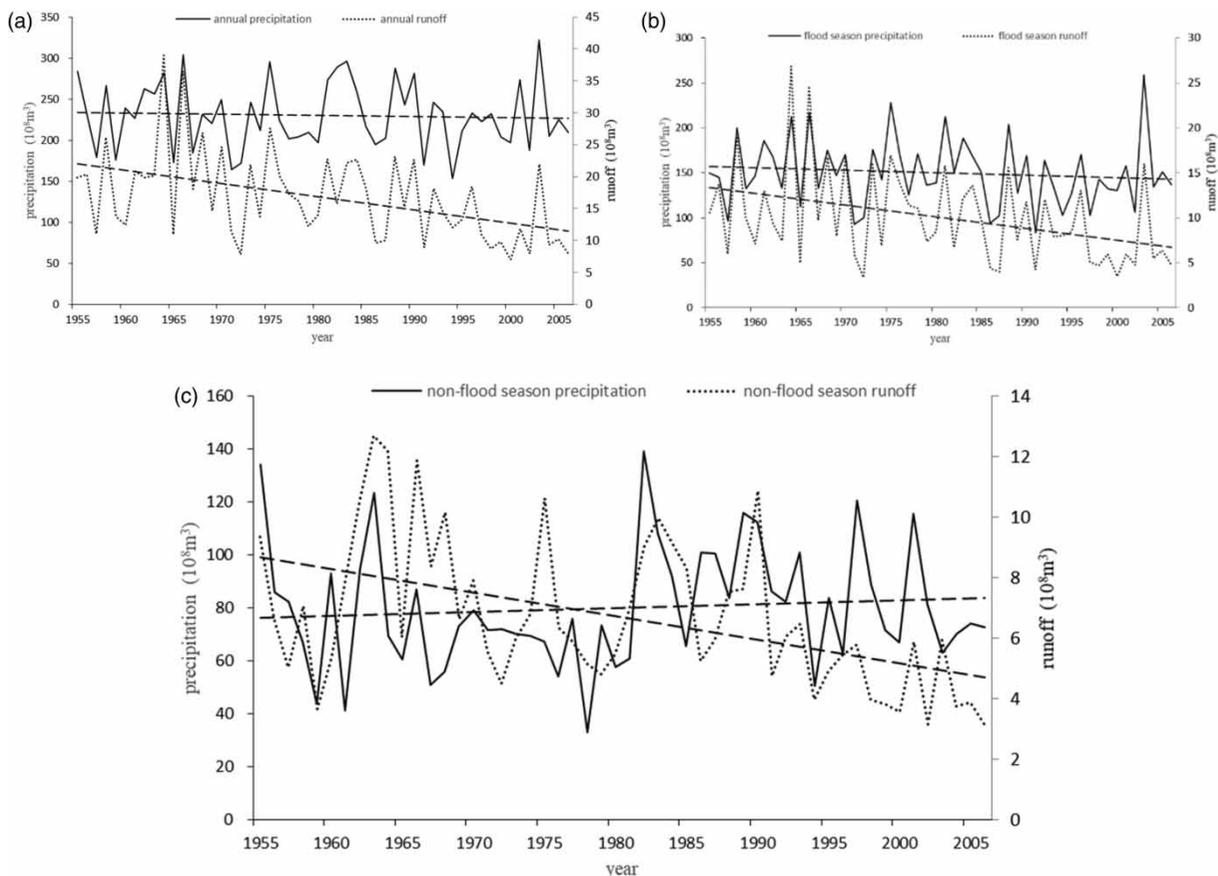

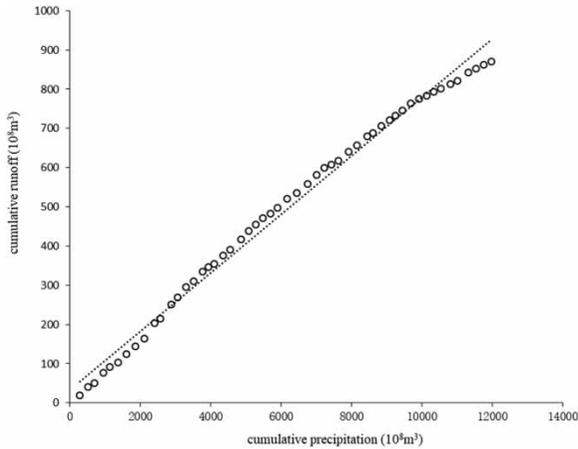**Figure 3** | Variability trend of precipitation and runoff.

Figure 4 │ Double mass curve between precipitation and runoff.

Changes in the multi-temporal scale of natural runoff

Using the EMD method, multi-temporal time series for the flood season, non-flood season, and annual runoff at ZHS from 1955 to 2006 were analyzed with an SD of 0.25 (Jin *et al.* 2017). As calculated by using Equations (1)–(3), each IMF component of the time series indicated feasible fluctuation periods. In general, the number of components of the basic model that depends on the response of the original signal to the SD and calculation criteria was not fixed. Further, Figures 5–7 show the decomposed runoff time series for the flood season, non-flood season, and annual by the EMD method.

The flood season time series (Figure 5) indicates that quasi-periodic fluctuations of 3–5 years were present in the IMF1 components time series except for some intervals. The largest runoff fluctuation occurred in the early 1960s. During the other periods, runoff was considered stable with a small range of variability, and slightly decreased amplitude was observed during the late 1990s. For the IMF2 components periodic fluctuations, the time series only had one interval with different fluctuations. Quasi-periodic fluctuations of 4–7 years were also observed. The IMF2 time series could be separated into two periods based on the 1970s. Before the 1970s, the fluctuation presented smaller amplitude and shorter periods than those after the 1970s. In particular, the amplitude increased with time and the largest fluctuation occurred from the year 2001. In the IMF3 components time series, quasi-periodic fluctuations of 8–12 years were present. The period wavelength increased while the amplitude of runoff fluctuation decreased gradually year after year, and the largest fluctuation was in the 1960s. In the IMF4 components of flood season time series, quasi-periodic fluctuations of 24 years were present between the 1950s and the mid-1970s. Further, there were no obvious quasi-periodic fluctuations. Figure 5 demonstrates that the residual components of the flood season were smoother and decreased over the years. In general, the quasi-periodic fluctuations presented an increasing trend from IMF1 to IMF4, until only a smooth decreasing trend of residual components was obtained. However, amplitude variation of IMF1, IMF3, and IMF4 showed a decreasing tendency, while IMF2 increased. Consequently, the flood season time series exhibited four types of time scale, 2–5, 4–7, 12, and 24 years.
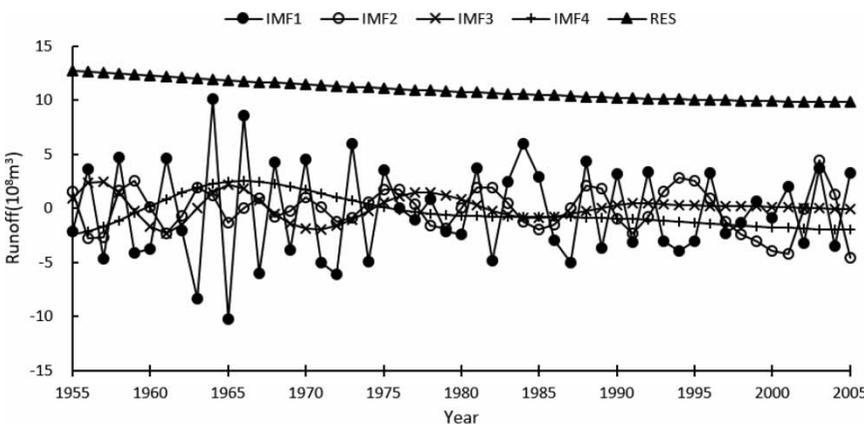

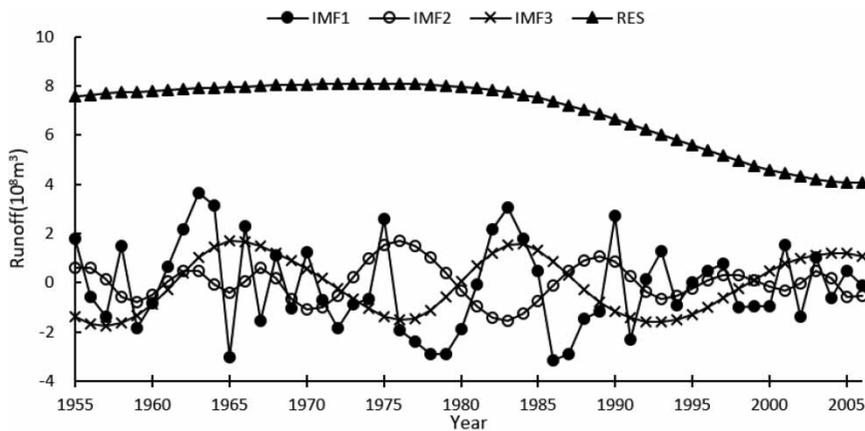**Figure 5** │ IMF components time series for flood season runoff at ZHS.

Figure 6 | IMF components time series for non-flood season runoff at ZHS.
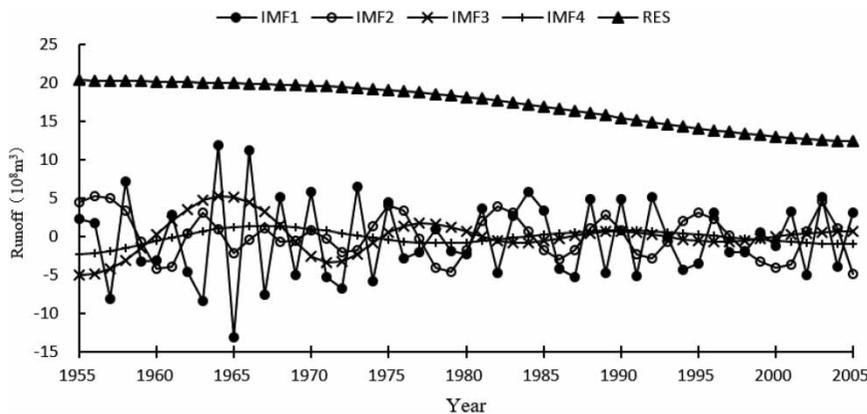

Figure 7 | IMF components time series for annual runoff time series at ZHS.

Figure 6 illustrates that quasi-periodic fluctuations of 3–7 years are present in the IMF1 components time series. The largest runoff fluctuation occurred in the early 1960s, and the second largest occurred from the mid-1970s to the early 1990s. During the other periods, runoff was stable.

The IMF2 components periodic fluctuations indicated that the time series only had one interval with different fluctuations. Quasi-periodic fluctuations of 5–10 years were observed, except for 1973–1986. The IMF2 time series was separated into two types: (1) greater amplitude larger time scale during 1973–1986, and (2) less amplitude shorter time scale during the other time with 5–7 years. However, the amplitudes of the fluctuations in the non-flood season time series decreased each year. Further, the largest fluctuation occurred in the mid-1970s.

For the IMF3 components time series, quasi-periodic fluctuations of 12–22 years were observed in the non-flood season time series. Moreover, along with time, the amplitudes

of non-flood season runoff were relatively small. The period wavelength was longer than that in the flood season time series, and the longest period occurred from 1972 to 1994. After the 1990s, the periodic fluctuation gradually shortened. The amplitude of runoff fluctuation was stable before the 1980s, and gradually decreased after the 1980s.

The residual components from the non-flood season showed faint varying trends. They seemed to be steady from prior to the mid-1970s, and then decreased.

In order to compare the difference between annual scale and seasonal scale, Figure 7 shows the decomposed results for annual runoff from 1955 to 2006 using the EMD method, which included four IMF components and one residual. It can be seen that:

(1) Quasi-periodic fluctuations of 3–5 years were present in the IMF1 component time series. The largest fluctuation

occurred in the mid-1960s. From the mid-1960s to the early 21st century, the runoff fluctuations were stable.

(2) The different amplitudes periods were very close to the IMF1. Except in the mid-1960s, quasi-periodic fluctuations of 4–9 years were present in the IMF2 component time series. The longest period wavelength was from the 1950s to the 1960s. The fluctuations around the 1950s and after the mid-1970s showed the largest amplitudes, and they remained relatively stable.

(3) Quasi-periodic fluctuations of 12–15 years were observed in the IMF3 component time series. In the entire time series the amplitude decreased consequently, while the fluctuation periods increased. Moreover, the largest fluctuation occurred around the mid-1960s.

(4) In the IMF4 component time series, there were only two complete waves. Quasi-periodic fluctuations of 23–27 years were present. One occurred from the 1950s to the 1970s when the quasi-periodic fluctuation of 23 years was present. The other was from the 1970s to the 1990s, when the quasi-periodic fluctuation of 27 years was present. Otherwise, in the entire time series, the fluctuations decreased gradually.

(5) In the residual component time series, the overall trend varied and decreased from 1955 to 2006. Especially, the decline was steeper after the 1990s.

The common point of these IMF functions was that different variation occurred in the mid-1960s and the late 1990s, and quasi-periodic fluctuation increased from the start year. In the meantime, all extracted factors exhibited shortened amplitude with time period except IMF3.

In general, the IMF extracted from different time series showed similar variation characteristics. Fluctuations during the mid-1960s and late-1990s showed obviously different characteristics from other periods. Furthermore, the quasi-periodic characteristic of fluctuation increased from the start year, while the majority of fluctuation amplitude decreased along with time.

SPA on runoff at multi-temporal scales

The runoff time series was decomposed to the IMF1–IMF4 components by the EMD method, and each component showed its own fluctuation periods. Therefore, the SPA assumed that the fluctuation period of the IMF1 component was short, that of the IMF2 component period was medium, and those of the IMF3 and IMF4 periods were long. The residual was not analyzed because it had no periodic fluctuation. In order to describe the relationship between the annual, flood season, and non-flood season time series, the IMFs were placed in different set pairs, as presented in Table 1. On the other hand, due to the relationship between precipitation and runoff, the SPA set different temporal scales between them (Table 1).

(1) The flood season runoff series for Jinghe River $Y_{flood} = (y_1, y_2, y_3, \ldots, y_{70})$ was denoted as Set *A*. Its four IMF components $X_m = (x_{m1}, x_{m2}, x_{m3}, \ldots, x_{m70})$ were specified as Set A_m, where m = 1, 2, 3, and 4. The non-flood season runoff series for Jinghe River $Y_{nonflood} = (y_1, y_2, y_3, \ldots, y_{70})$ was denoted as Set *B*. Its three IMF components $X_m = (x_{m1}, x_{m2}, x_{m3}, \ldots, x_{m70})$ were specified as Set B_m, where m = 1, 2, and 3. Note, the non-flood season runoff was only decomposed into three IMF components, but the IMF3 component involved the medium-long and long periods. The four

Table 1 | SPA on runoff with different temporal scales

Components/runoff	H (Am, Bm)			
	m = 1	m = 2	m = 3	m = 4
Flood and non-flood runoff	IMF1,IMF1	IMF2,IMF2	IMF3,IMF3	IMF4,IMF3
Flood and annual runoff	IMF1,IMF1	IMF2,IMF2	IMF3,IMF3	IMF4,IMF4
Non-flood and annual runoff	IMF1,IMF1	IMF2,IMF2	IMF3,IMF3	IMF3,IMF4
Annual ruoff and IMF	A,IMF1	A,IMF2	A,IMF3	A,IMF4
Annual precipitation and runoff	P,R	P_{flood}, R_{flood}	$P_{non-flood}, R_{non-flood}$	

Runoff at different temporal scales employed for each set (A_m, B_m) in Table 1 is denoted as:.

set pairs were denoted as $(A_1, B_1), (A_2, B_2), (A_3, B_3)$, and (A_4, B_3), which were made up of the four decomposed flood season runoff IMF components, and three decomposed non-flood season runoff IMF components.

(2) The flood season runoff series for Jinghe River $Y_{flood} = (y_1, y_2, y_3, \ldots, y_{70})$ was denoted as Set A. Its four IMF components $X_m = (x_{m1}, x_{m2}, x_{m3}, \ldots, x_{m70})$ were specified as Set $A_m (m = 1, 2, 3, 4)$. The annual runoff series for Jinghe River $Y_{annual} = (y_1, y_2, y_3, \ldots, y_{70})$ was denoted as Set B. Its four IMF components $X_m = (x_{m1}, x_{m2}, x_{m3}, \ldots, x_{m70})$ were specified as Set $B_m (m = 1, 2, 3, 4)$. The four set pairs were denoted as $(A_1, B_1), (A_2, B_2), (A_3, B_3)$ and (A_4, B_4), which were made up of the four decomposed flood season runoff IMF components, and the four decomposed annual runoff IMF components.

(3) The non-flood season runoff series for Jinghe River $Y_{flood} = (y_1, y_2, y_3, \ldots, y_{70})$ was denoted as Set A. Its three IMF components $X_m = (x_{m1}, x_{m2}, x_{m3}, \ldots, x_{m70})$ were specified as Set $A_m (m = 1, 2, 3)$. The annual runoff series for Jinghe River $Y_{annual} = (y_1, y_2, y_3, \ldots, y_{70})$ was denoted as Set B. Its four IMF components $X_m = (x_{m1}, x_{m2}, x_{m3}, \ldots, x_{m70})$ were specified as Set $B_m (m = 1, 2, 3, 4)$. The four set pairs were denoted as (A_1, B_1), (A_2, B_2), (A_3, B_3), and (A_3, B_4), which were made up of the three decomposed non-flood season runoff IMF components, and the four decomposed annual runoff IMF components.

(4) The natural annual runoff series for Jinghe River $Y_{annual} = (y_1, y_2, y_3, \ldots, y_{70})$ was denoted as Set A. Its four IMF components $X_m = (x_{m1}, x_{m2}, x_{m3}, \ldots, x_{m70})$ were specified as Set $B_m (m = 1, 2, 3, 4)$. The four set pairs were denoted as (A, B_1), (A, B_2), (A, B_3), and (A, B_4), which were made up of the four decomposed IMF components, and the natural annual runoff series.

(5) The precipitation series including annual, flood season, and non-flood season for Jinghe basin $P_m = (y_1, y_2, y_3, \ldots, y_{52})$ was denoted as Set $A_m (m = 1, 2, 3)$. Natural runoff including annual, flood season and non-flood season $R_m = (x_{m1}, x_{m2}, x_{m3}, \ldots, x_{m52})$, were specified as Set $B_m (m = 1, 2, 3)$. The three set pairs were denoted as (A_1, B_1), (A_2, B_2), and (A_3, B_3) which were made up of the annual, flood season, and non-flood season series.

For these SPA groups, runoff from the annual, flood season, and non-flood season time series were decomposed into IMFs. Further, each IMF element was classified into three parts: (I) poor, (II) normal, and (III) rich. By using the mean SD method, the IMF components were classified into the three parts and their corresponding classification ranges were $(-\infty, \text{Ex} - 0.5\text{d})$, $(\text{Ex} - 0.5\text{d}, \text{Ex} + 0.5\text{d})$, and $(\text{Ex} + 0.5\text{d}, +\infty)$, where Ex is the mean value of the set in a set pair and d is the SD. Table 2 presents the classification results for the annual, flood season, and non-flood season time series, respectively.

Based on the classification results presented in Table 2 and Equations (5) and (6), for $i = 0.5$, j was specified as -1, the identity degree a, the discrepancy degree b, the contrary degree c, and the connection degree μ at multi-temporal scales were then obtained, which are listed in Table 3.

Table 2 │ Classification results for multiple time series

Time series	Classification	Y	X₁	X₂	X₃	X₄
Annual runoff	Poor (I)	$(-\infty, 14.62]$	$(-\infty, -3.09]$	$(-\infty, -1.09]$	$(-\infty, -0.74]$	$(-\infty, -0.42]$
	Normal (II)	$(14.62, 20.45)$	$(-3.09, 1.76)$	$(-1.09, 1.03)$	$(-0.74, 0.80)$	$(-0.42, 0.62)$
	rich (III)	$[20.45, +\infty)$	$[1.76, +\infty)$	$[1.03, +\infty)$	$[0.80, +\infty)$	$[0.62, +\infty)$
Flood season runoff	poor (I)	$(-\infty, 8.50]$	$(-\infty, -2.52]$	$(-\infty, -0.87]$	$(-\infty, -0.48]$	$(-\infty, -0.89]$
	Normal (II)	$(8.50, 13.03)$	$(-2.52, 1.51)$	$(-0.87, 0.58)$	$(-0.48, 0.65)$	$(-0.89, 0.50)$
	Rich (III)	$[13.03, +\infty)$	$[1.51, +\infty)$	$[0.58, +\infty)$	$[0.65, +\infty)$	$[0.50, +\infty)$
Non-flood season runoff	Poor (I)	$(-\infty, 5.86]$	$(-\infty, -0.97]$	$(-\infty, -0.25]$	$(-\infty, -0.47]$	
	Normal (II)	$(5.86, 7.68)$	$(-0.97, 0.45)$	$(-0.25, 0.43)$	$(-0.47, 0.38)$	
	Rich (III)	$[7.68, +\infty)$	$[0.45, +\infty)$	$[0.43, +\infty)$	$[0.38, +\infty)$	
Annual precipitation	Poor (I)	$(-\infty, 493.69]$	$(-\infty, 313.02]$	$(-\infty, 164.16]$		
	Normal (II)	$(493.69, 571.03)$	$(313.02, 381.23)$	$(164.16, 206.31)$		
	Rich (III)	$[571.03, +\infty)$	$[381.23, +\infty)$	$[206.31, +\infty)$		

Further, the meaning of the set pair index in Table 3 is the same as in Table 1.

For the set pairs in the sequence from (A1, B1) to (A4, B3), the fluctuation period increased, and SPA results for the flood season and non-flood season runoff (Table 3) showed that:

(1) The identity degree a of the flood season and non-flood season runoff decreased. For the identity degree, the maximum was 0.53 on the fluctuation with a short period, while the minimum was 0.21 on the fluctuation with a long period. The identity degree of the flood season and non-flood season runoff were mainly controlled by short periods.

(2) The discrepancy degree b of the flood season and non-flood season runoff first increased, then decreased, and increased again. For the discrepancy degree, the maximum was 0.54 on the fluctuation with a long period, while the minimum was 0.27, on the fluctuation with a medium-long period.

(3) The contrary degree c of the flood season and non-flood season runoff first increased and then decreased. For the contrary degree, the maximum was 0.31 on the fluctuation with a medium-long period, while the minimum was 0.07 on the fluctuation with a short period.

(4) The connection degree μ of the flood season and non-flood season runoff decreased. For the connection degree, the maximum was 0.66 on the fluctuation with a short period; however, the minimum was 0.24 on the fluctuation with a long period. The results for the connection degree were consistent with those for the identity degree.

SPA results for flood season and annual runoff (Table 3) showed that:

(1) The identity degree of the flood season and annual runoff first decreased, then increased, and decreased again. For the identity degree, the maximum was 0.89 on the fluctuation with a short period, while the minimum was 0.63 on the fluctuation with a medium-long period. Thus, the identity degree of the flood season and annual runoff was mainly on the fluctuation with a short period.

(2) The discrepancy degree of the flood season and annual runoff first increased, then decreased, and increased again. For the discrepancy degree, the maximum was 0.34 on the fluctuation with a medium period, while the minimum was 0.11 on the fluctuation with a short period.

(3) The contrary degree of the flood season and annual runoff first increased and then decreased. For the contrary degree, the maximum was 0.10 on the fluctuation with a medium-long period, while the minimum was 0 on the fluctuation with a short period.

(4) The connection degree of the flood season and non-flood season runoff decreased from 0.94 to 0.70. The results for the connection degree were consistent with those for the identity degree.

SPA results for non-flood season and annual runoff (Table 3) showed that:

(1) The identity degree of the non-flood season and annual runoff first decreased, then increased, and for the fluctuations with long periods decreased again. For the identity degree, the maximum was 0.57 on the fluctuation with a short period; however, the minimum was 0.43 on the fluctuation with a long period. The identity degree of the non-flood season and annual runoff was mainly on the fluctuation with a short period.

(2) The discrepancy degree of the non-flood season and annual runoff first increased, then decreased, and increased again. For the contrary degree, the maximum was 0.41 on the fluctuation with a medium period, while the minimum was 0.30 on the fluctuation with a medium-long period.

(3) The contrary degree of the non-flood season and annual runoff increased throughout. For the contrary degree, the maximum was 0.17 on the fluctuation with a long period; however, the minimum was 0.09 on the fluctuation with a short period.

(4) The connection degree of the non-flood season and annual runoff ranged from 0.66 to 0.46. The results for the connection degree were consistent with those for the identity degree.

SPA results for annual runoff at multi-temporal scales (Table 3) showed that:

(1) The identity degree of the annual runoff at multi-temporal scales first decreased, and then increased. For the identity degree, the maximum was 0.59 on the fluctuation with a short period, while the minimum was 0.30 on the fluctuation with a medium-long period. The identity degree of the annual runoff at multi-temporal scales was mainly on the fluctuation with a short period.

(2) The discrepancy degree of the annual runoff at multi-temporal scales first decreased, then increased, and decreased again. For the discrepancy degree, the maximum was 0.59 on the fluctuation with a medium-long period; nonetheless, the minimum was 0.37 on the fluctuation with a medium period.

(3) The contrary degree of the annual runoff at multi-temporal scales decreased throughout. For the contrary degree, the maximum was 0.30 on the fluctuation with a short period, while the minimum was 0.11 on the fluctuation with a medium-long period.

(4) The connection degree of the annual runoff at multi-temporal scales decreased from 0.75 to 0.48. The results for the connection degree were consistent with those for the identity degree.

SPA results for precipitation and runoff at multi-temporal scales (Table 3) showed that:

For different temporal sets, Table 3 summarizes that for the identity degree, the maximum was 0.69 on the fluctuation with the annual period, while the minimum was 0.37 on the fluctuation with the non-flood period. The identity degree results for annual runoff and flood season temporal scales were relatively similar. For the discrepancy degree, the maximum was 0.48 with the non-flood period, while the minimum was 0.29 in annual sets. The maximum of the contrary degree was 0.15 with the non-flood period. The results for the connection degree were consistent with those for the identity degree.

The characteristics of properties for multi-temporal scales could be concluded from data summarized in Table 3. For the identity degree *a*, the maximum with high similarity was presented in the short period. Unfortunately, the maximum of discrepancy degree *b* and contrary degree *c* were present in different periods. Based on the fuzziness and randomness coefficients, the connection degree within all set pairs was decreased from short period to long period. This variation could be ascribed to the fact that the short period was the best response scale between precipitation and runoff, as well as different time scales for runoff.

Table 3 | SPA for multi-temporal scales

Set pairs		a	b	c	μ
Flood season and non-flood season runoff	(A1,B1)	0.53	0.40	0.07	0.66
	(A2,B2)	0.36	0.50	0.14	0.46
	(A3,B3)	0.41	0.27	0.31	0.24
	(A4,B3)	0.21	0.54	0.24	0.24
Flood season and annual runoff	(A1,B1)	0.89	0.11	0	0.94
	(A2,B2)	0.63	0.34	0.03	0.77
	(A3,B3)	0.7	0.2	0.1	0.7
	(A4,B4)	0.63	0.31	0.06	0.73
Non-flood season and annual runoff	(A1,B1)	0.57	0.34	0.09	0.66
	(A2,B2)	0.49	0.41	0.1	0.59
	(A3,B3)	0.54	0.3	0.16	0.54
	(A3,B4)	0.43	0.4	0.17	0.46
Annual runoff	(A,B1)	0.59	0.39	0.3	0.75
	(A,B2)	0.5	0.37	0.13	0.56
	(A,B3)	0.3	0.59	0.11	0.48
	(A,B4)	0.39	0.47	0.14	0.48
Precipitation and runoff	(A1,B1)	0.69	0.29	0.02	0.82
	(A2,B2)	0.62	0.38	0	0.81
	(A3,B3)	0.37	0.48	0.15	0.45

DISCUSSION

From 1955 to 2006, the original time series of annual runoff and precipitation did not change significantly, as presented in Figure 3. However, fluctuation trends between runoff and precipitation in the non-flood season were different. The most obvious point in the non-flood series occurred in the year 1975 (Figure 3(c)). Precipitation in the flood season of 1975 was 3.38 times of that in the non-flood season, possibly caused by the hysteresis between runoff and precipitation. Nevertheless, the relationship between accumulate precipitation and runoff verified that the real feasible mutated points were not the trend crossing year, but the years 1966 and 1999.

However, only the single line chart could not manifest the elucidation of the internal mechanism. Therefore, considering the self-adaption, non-linear, and non-stationary characteristics of runoff time series, EMD was used to

decompose runoff at multi-temporal scales. The results showed that all the IMF1 and IMF2 components exhibited similar fluctuation periods. The IMF3 and IMF4 components also showed similar fluctuation periods for the flood season and annual runoff; however, the IMF3 components periodic fluctuation period of the non-flood season was between the flood season and annual runoff.

Comparative analysis of the amplitude results for different time series indicates that the IMF1 amplitudes for non-flood season runoff were greater than those for the flood season runoff. For the IMF2 components periodic fluctuations, both the flood and non-flood season time series showed only one interval with different fluctuations. Moreover, the corresponding time node occurred consistently with the mutation point of the accumulation curve. However, the largest fluctuations in the flood season and non-flood season time series occurred during the early 21st century and the mid-1960s, respectively. On the other hand, from the perspective of oscillation, all IMF exhibits the characteristic of extended cycle. Moreover, clearly, the significant variation position of the periodic boundary of annual runoff and flood season runoff just coincides with the abrupt transition point of the cumulative rainfall curve. The break period of non-flood season presents 3–5 years delay. This shows that rainfall in Jing River acts on runoff obviously and immediately during flood season and annual runoff. However, it generates a delayed response to non-flood season. Furthermore, the results of SPA presented in Table 3 also prove the conclusion. The maximum value of identity degree and connection degree occurred in the short period.

Even though uncertainty in hydrological phenomena has been widely accepted and discussed in earlier studies (Karthikeyan & Kumar 2013; Zhang *et al.* 2014), in this study, the EMD method was used to verify that a longer time series improves the response of inherent runoff characteristics. Based on the historical documents, the reconstruction of the hydro-climatic spatial patterns showed that a strong El Niño event was likely to induce precipitation anomalies over various regions, in particular, in the Asia-Pacific region, including the study area (Zhang *et al.* 2017a, 2017b, 2018). Further, the results of many earlier studies from observation or reanalysis data indicated that the impact of El Niño on precipitation anomalies in China

was mainly characterized by less rainfall in northern China (Hao *et al.* 2019). Comparative analysis of these El Niño variable periods and the quasi-periodic fluctuations showed that the existence of the quasi-periodic fluctuations of 3.5 and 4–8 years was an El Niño phenomenon (Zhang *et al.* 2016, 2017a, 2017b). El Niño variability was implicated as a variation single of precipitation (Zhang *et al.* 2018), and the double mass curve showed high correlation with a correlation coefficient of 0.99 between runoff and precipitation. Runoff fluctuation periods for IMF1 and IMF2 components were very similar to the cycle periods of the El Niño phenomenon's effects. This showed that there could have been a close relationship between El Niño and runoff in Jinghe River. Moreover, the IMF3 and IMF4 components may be related to the air–sea intersection and the long period of solar activity. Zhang *et al.* (2018) simulated the El Niño interannual variability and discussed impacts and interactions, in north China. By comparing the statistical ENSO mature phase year with runoff fluctuations for the Jinghe River, they showed relatively synperiodic but lagged effects. It was commonly concluded that when El Niño occurred, runoff increased (Zhang *et al.* 2016; Hao *et al.* 2019).

Despite the fact that the runoff was influenced by several factors, features within different temporal time scales need to be further classified by the coupling method of SPA with EMD factors. Therefore, based on the connection degree, SPA was applied to identify the classification based on different runoff volumes for each scale. Table 3 presents the existence of several common features. The maximum values of a and μ for all the sets were in the IMF1 components. On the other hand, most maximum values of b were in the IMF2 components except for the set of annual runoff at multi-temporal scales. The most significant difference of c was in the IMF3 components when the flood season was one of the set components (Table 3); however, the other two sets (Table 3) revealed different components. Comparing the indicator c with the quasi-periodic fluctuations of the IMF3 components demonstrated that the quasi-periodic fluctuation of IMF3 in the flood season was different from the other two temporal scales. As precipitation is one of the important natural determinants of change in river runoff, the SPA between precipitation and runoff with multiple time series showed quite similar results to IMF1. However, it can be considered consistent with the

coupled method application. From the basic definition of SPA, it can be concluded that the short period embodies the best connection degree between each set pair. With the increase of time scale, the connection degree between set pair indicators decreases, indicating that the mutual connection between them weakens and the influence of the other factors increases (Zhang *et al.* 2016).

CONCLUSIONS

The runoff series sounds to be a chaotic series, and the traditional methods could not perform the real logical characteristics. To investigate the fluctuation uncertainty and response relationship to runoff, the EMD-SPA method was determined and the results showed the following benefits:

(1) During 1955–2006, runoff and rainfall in Jinghe River catchment showed decreasing trends. Moreover, accumulative curve of rainfall reflected two noteworthy periods in the late 1960s and 1990s. These variation periods were also observed and investigated by the EMD-SPA method.

(2) By employing the EMD method, time series for the flood season, non-flood season, and annual runoff were decomposed into four or three components. The results showed that quasi-periodic fluctuations of all the time series were lengthening and the oscillation amplitude was decreasing. Amplitude mutation periods of annual runoff and flood season runoff occurred corresponding to the time of rainfall. However, non-flood season appeared a little later corresponding to the abrupt change of rainfall by 2–3 years.

(3) The SPA for runoff at multi-temporal scales and runoff with rainfall at multi-temporal scales determined the relationships between the IMF components based on the parameters of identity, discrepancy, and contrary. It showed that identity degrees of short period were the main component. Furthermore, it also indicated that the highest connection degree of the multiple series was present in the short period, and then the value decreased with the increase in the quasi-periodic feature. These results showed that runoff changes over the short term could basically reflect the average state

of the Jinghe River. Therefore, enhancing short-term forecasting accuracy can improve the regulation and planning of water resources.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

REFERENCES

Adamowski, J. & Chan, H. F. 2011 A wavelet neural network conjunction model for groundwater level forecasting. *Journal of Hydrology* **407**, 28–40. doi:10.1016/j.jhydrol.2011.06.013.

Birkel, C. & Barahona, A. C. 2019 Rainfall-runoff modeling: a brief overview. *Reference Module in Earth Systems and Environmental Sciences*. doi:10.1016/B978-0-12-409548-9.11595-7.

Blöschl, G., Hall, J., Parajka, J., Perdigão, R. A. P., Merz, B., Arheimer, B., Aronica, G. T., Bilibashi, A., Bonacci, O., Borga, M., Čanjevac, I., Castellarin, A., Chirico, G. B., Claps, P., Fiala, K., Frolova, N., Gorbachova, L., Gül, A., Hannaford, J., Harrigan, S., Kireeva, M., Kiss, A., Kjeldsen, T. R., Kohnová, S., Koskela, J. J., Ledvinka, O., Macdonald, N., Mavrova-Guirguinova, M., Mediero, L., Merz, R., Molnar, P., Montanari, A., Murphy, C., Osuch, M., Ovcharuk, V., Radevski, I., Rogger, M., Salinas, J. L., Sauquet, E., Šraj, M., Szolgay, J., Viglione, A., Volpi, E., Wilson, D., Zaimi, K. & Živković, N. 2017 Changing climate shifts timing of European floods. *Science* **357**, 588–590. doi: 10.1126/science.aan2506.

Chau, K. W. 2017 Use of meta-heuristic techniques in rainfall-runoff modelling. *Water* **9**, 186. doi:10.3390/w9030186.

Chou, C. M. 2011 Wavelet-based multi-scale entropy analysis of complex rainfall time series. *Entropy* **13**, 241–253. doi:10.3390/e13010241.

Chou, C. M. 2014 Complexity analysis of rainfall and runoff time series based on sample entropy in different temporal scales. *Stochastic Environmental Research and Risk Assessment* **28**, 1401–1408. doi:10.1007/s00477-014-0859-6.

Giulia, N., Francesco, S. & Linda, S. 2011 Impact of EMD decomposition and random initialisation of weights in ANN

hindcasting of daily stream flow series: an empirical examination. *Journal of Hydrology* **406**, 199–214. doi:10.1016/j.jhydrol.2011.06.015.

Hao, Z., Sun, D., Wu, M. W. & Zheng, J. Y. 2019 Does El Niño play an early signal role for the south-flood north-drought pattern over eastern China? *Theoretical & Applied Climatology* **137**, 217–227. doi:10.1007/s00704-018-2590-6.

He, Z. B., Wen, X. H., 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:10.1016/j.jhydrol.2013.11.054.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q. A., Yen, N.-C., Tung, C. C. & Liu, H. H. 1998 The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society* **454**, 899–955. doi:10.1098/rspa.1998.0193.

Huang, N. E., Shen, Z. & Long, S. R. 1999 A new view of nonlinear water waves: the Hilbert spectrum. *Annual Review of Fluid Mechanics* **31**, 417–457. doi:10.1146/annurev.fluid.31.1.417.

Huang, S. Z., Chang, J. X., Huang, Q. & Chen, Y. T. 2014 Monthly streamflow prediction using modified EMD-based support vector machine. *Journal of Hydrology* **511**, 764–775. doi:10.1016/j.jhydrol.2014.01.062.

Jeon, J.-H., Park, C.-G. & Engel, B. A. 2014 Comparison of performance between genetic algorithm and SCE-UA for calibration of SCS-CN surface runoff simulation. *Water* **6**, 3433–3456. doi:10.3390/w6113433.

Jin, J. L., Shen, S. X., Pan, Z. W., Wang, M. W. & Cui, Y. 2017 Advances in theoretical and applied research on set pair analysis method for water resources system. *Journal of North China University of Water Resources and Electric Power (Natural Science Edition)* **38** (4), 54–66. doi:10.3969/j.issn.1002-5634.2017.04.008. (in Chinese).

Karthikeyan, L. & Kumar, D. N. 2013 Predictability of nonstationary time series using wavelet and EMD based ARMA models. *Journal of Hydrology* **502**, 103–119. doi:10.1016/j.jhydrol.2013.08.030.

Kim, D. & Oh, H. S. 2009 EMD: a package for empirical mode decomposition and Hilbert spectral analysis. *Royal Journal* **1**, 40–46. doi: 10.32614/RJ-2009-002.

Lee, T. & Ouarda, T. B. M. J. 2011 Prediction of climate non-stationary oscillation processes with empirical mode decomposition. *Journal of Geophysical Research: Atmospheres* **116**, D06107. doi:10.1029/2010JD015142.

Lloyd, H. C. & Tommy, S. W. 2011 Runoff forecasting for an asphalt plane by artificial neural networks and comparisons with kinematic wave and autoregressive moving average models. *Journal of Hydrology* **397**, 191–201. doi:10.1016/j.jhydrol.2010.11.030.

Nourani, V. 2017 An emotional ANN (EANN) approach to modeling rainfall-runoff process. *Journal of Hydrology* **544**, 267–277. doi:10.1016/j.jhydrol.2016.11.033.

Nourani, V., Baghanam, H. A., Adamowski, J. & Gebremichael, M. 2013 Using self-organizing maps and wavelet transforms for space–time pre-processing of satellite precipitation and runoff data in neural network based rainfall–runoff modeling. *Journal of Hydrology* **407**, 28–40. doi:10.1016/j.jhydrol.2012.10.054.

Pan, Z. W., Wang, Y. H., Jin, J. L. & Liu, X. W. 2017 Set pair analysis method for coordination evaluation in water resources utilizing conflict. *Physics & Chemistry of the Earth* **101**, 149–156. doi:10.1016/j.pce.2017.05.009.

Potocki, K., Gilja, G. & Kunstek, D. 2017 An overview of the applications of wavelet transform for discharge and suspended sediment analysis/Pregled primjene valicne transformacije u analizi protoka i suspendiranog nanosa. *Tehnicki Vjesnik – Technical Gazette* **24**, 1561–1569.

Rhif, M., Ben Abbes, A., Farah, I. R., Martínez, B. & Sang, Y. 2019 Wavelet transform application for/in non-stationary time-series analysis: a review. *Applied Science* **9**, 1345. doi:10.3390/app9071345.

Roushangar, K., Nourani, V. & Alizadeh, F. 2018 A multi scale time-space approach to analyze and categorize the precipitation fluctuation based on the wavelet transform and information theory concept. *Hydrology Research* **49**, 724–743. doi:10.2166/nh.2018.143.

Searcy, J. K. & Hardison, C. H. 1960 Double-Mass Curves Manual of Hydrology: Part 1 General Surface Water Techniques. United States Government Printing Office, Washington, DC, USA.

Wang, W. S., Xiang, H. L., Li, Y. Q. & Wang, S. J. 2008 A new approach to annual runoff classification based on set pair analysis. *Journal of Sichuan University* **40**, 29–33. (in Chinese).

Yan, F. S. 2013 A review on the applications of wavelet transform in hydrology time series analysis. *Atmospheric Research* **122**, 8–15. doi:10.1016/j.atmosres.2012.11.003.

Yang, P. C., Wang, G. L., Bian, J. C. & Zhou, X. J. 2010 The prediction of non-stationary climate series based on empirical mode decomposition. *Advances in Atmospheric Science* **27**, 845–854. doi: 10.1007/s00376-009-9128-x.

Yilinuer, A., Jilili, A., Michael, G., Gulnura, I. & Balati, M. 2021 Changes in intra-annual runoff and its response to climate variability and anthropogenic activity in the Lake Issyk-Kul Basin, Kyrgyzstan. *CATENA* **198**, 104974. doi:10.1016/j.catena.2020.104974.

Yu, P. S., Yang, T. C. & Chen, S. J. 2001 Comparison of uncertainty analysis methods for a distributed rainfall-runoff model. *Journal of Hydrology* **244**, 43–59. doi:10.1016/S0022-1694(01)00328-6.

Yu, Y., Zhang, H. & Singh, V. P. 2018 Forward prediction of runoff data in data-Scarce basins with an improved ensemble empirical mode decomposition (EEMD) model. *Water* **10**, 388. doi:10.3390/w10040388.

Zhang, Q., Wang, B. D., He, B., Peng, Y. & Ren, M. L. 2011 Singular spectrum analysis and ARIMA hybrid model for annual runoff forecasting. *Water Resources Management* **25**, 2683–2703. doi: 10.1007/s11269-011-9833-y.

Zhang, J. P., Zhao, Y. & Ding, Z. 2014 Research on the relationships between rainfall and meteorological yield in irrigation district. *Water Resources Management* **28**, 1689–1702. doi: 10.1007/s11269-014-0577-3.

Zhang, Y., Ding, A. J., Mao, H. T., Nie, W., Zhou, D. R., Liu, L. X., Huang, X. & Fu, C. B. 2016 Impact of synoptic weather patterns and inter-decadal climate variability on air quality in the North China Plain during 1980– 2013. *Atmospheric Environment* **124**, 119–128. doi:10.1016/j.atmosenv.2015.05.063.

Zhang, J. P., Zhao, Y. & Lin, X. 2017a Uncertainty analysis and prediction of river runoff with multi-time scales. *Water Science & Technology* **17**, 897–906. doi:10.2166/ws.2016.190.

Zhang, R. H., Min, Q. Y. & Su, J. Z. 2017b Impact of El Niño on atmospheric circulations over East Asia and rainfall in China: role of the anomalous western North Pacific anticyclone. *Science China Earth Sciences* **60**, 1124–1132. doi:10.1007/s11430-016-9026-x.

Zhang, L. X., Wu, P. L., Zhou, T. J. & Xiao, C. 2018 ENSO transition from La Niña to El Niño drives prolonged spring–summer drought over North China. *Journal of Climate* **5**, 3509–3523. doi:10.1175/JCLI-D-17-0440.1.

Zhang, J. D., Li, Y., Liu, C., Qu, Z., Li, F., Yang, Z., Jiang, L. & Fu, J. 2019 Application of set pair analysis in a comprehensive evaluation of water resource assets: a case study of Wuhan city, China. *Water* **11**, 1718. doi:10.3390/w11081718.

Zhao, K. Q. & Xuan, A. L. 1989 Pair theory-a new theory method of non-define and its applications. *Systems Engineering* **14**, 18–23.

Zhao, X., Chen, X., Xu, Y., Xi, D., Zhang, Y. & Zheng, X. 2017 An EMD-based chaotic least squares support vector machine hybrid model for annual runoff forecasting. *Water* **9**, 153. doi:10.3390/w9030153.

Zhu, B., Wang, H. F., Wang, W. S. & Li, Y. Q. 2007 Analysis of relation between flood peak and volume based on set pair analysis. *Journal of Sichuan University* **39**, 29–33. (in Chinese).