# Multi-temporal runoff-sediment discharge relationships in the source region of the Yellow River

Honglin Xiao[1], Jinping Zhang[1], and Hongyuan Fang[1]

[1]Zhengzhou University

September 28, 2020

**Abstract**

To understand the runoff-sediment discharge relationship in the source region of the Yellow River, this study examined the annual runoff and sediment discharge data obtained from the Tangnaihai hydrometric station. The data were decomposed into multiple time scales through Complete Ensemble Empirical Mode Decomposition with adaptive noise (CEEMDAN). Furthermore, double cumulative curves were plotted and the cointegration theory was employed to analyze the microscopic and macroscopic multi-temporal correlations between the runoff and the sediment discharge and their detailed evolution. Multi-temporal component composite models were then constructed considering structural breaks. The simulation results were compared with the actual values to examine the accuracy of the models. The results suggested that the runoff and the sediment discharge variations in the source region of the Yellow River showed reasonable consistency as a whole. However, their relationship at different time scales varied slightly. The runoff-sediment discharge double cumulative curves in the multi-temporal components exhibited high goodness of fit. The curves of the intrinsic mode function 1 and 2 (IMF1 and IMF2) components provided a more satisfactory goodness of fit, whereas distinct breakpoints were present in those of IMF3 and IMF4. The variations in the runoff-sediment discharge relationship of the raw data series resulted from the different time scales. The medium- and long-term runoff-sediment discharge relationships were insignificant, which affected the raw data series. With the help of the variable structure cointegration composite model, the smallest average relative error for the simulated annual runoff (7.82%) was obtained. This composite model could more accurately reflect the long-term equilibrium and short-term fluctuating relationships between the runoff and the sediment discharge in the source region of the Yellow River.

# Multi-temporal runoff-sediment discharge relationships in the source region of the Yellow River

Honglin Xiao [a], Jinping Zhang [a, b,]11* Corresponding authors: Name: Jinping Zhang, E-mail address: jinping2000_zh@163.com; Name: Hongyuan Fang, E-mail address: 18337192244@163.com., Hongyuan Fang [a,*]

a: School of Water Conservancy Engineering, Zhengzhou University, High-tech District, No. 100 Science Road, Zhengzhou City, 450001, Henan Province, China.

b: Yellow River Institute for Ecological Protection & Regional Coordinated Development, Zhengzhou University, Zhengzhou City, 450001, Henan Province, China.

**Abstract:** To understand the runoff-sediment discharge relationship in the source region of the Yellow River, this study examined the annual runoff and sediment discharge data obtained from the Tangnaihai hydrometric station. The data were decomposed into multiple time scales through Complete Ensemble Empirical

Mode Decomposition with adaptive noise (CEEMDAN). Furthermore, double cumulative curves were plotted and the cointegration theory was employed to analyze the microscopic and macroscopic multi-temporal correlations between the runoff and the sediment discharge and their detailed evolution. Multi-temporal component composite models were then constructed considering structural breaks. The simulation results were compared with the actual values to examine the accuracy of the models. The results suggested that the runoff and the sediment discharge variations in the source region of the Yellow River showed reasonable consistency as a whole. However, their relationship at different time scales varied slightly. The runoff-sediment discharge double cumulative curves in the multi-temporal components exhibited high goodness of fit. The curves of the intrinsic mode function 1 and 2 (IMF1 and IMF2) components provided a more satisfactory goodness of fit, whereas distinct breakpoints were present in those of IMF3 and IMF4. The variations in the runoff-sediment discharge relationship of the raw data series resulted from the different time scales. The medium- and long-term runoff-sediment discharge relationships were insignificant, which affected the raw data series. With the help of the variable structure cointegration composite model, the smallest average relative error for the simulated annual runoff (7.82%) was obtained. This composite model could more accurately reflect the long-term equilibrium and short-term fluctuating relationships between the runoff and the sediment discharge in the source region of the Yellow River.

**Keywords:** runoff, sediment discharge, source region of the Yellow River, variable structure cointegration, multi-temporal scales

## 1 Introduction

Water plays an essential role in the rock cycle, especially in sediment transport. Furthermore, runoff and sediment transport are the results of interactions of various natural and human factors and the superposition of their effects (Li et al., 2017; Gu et al., 2019; Zhao et al., 2014). The Yellow River is known for its huge sediment load. However, the large amount of sediment carried by the river is continuously deposited along its riverbed, which often causes devastating floods. This leads to significant hazards to the people and the country (Guo et al., 2020a; Bai et al., 2019). Although this issue usually affects the midstream and downstream regions, the sediment contents in the water in the upstream regions are closely related to those in the midstream and downstream regions. Understanding upstream sediment content variation is the basis for analyzing those midstream and downstream. Furthermore, the source region of the Yellow River is the most important headwater region in the Yellow River Basin in terms of the amount of water contributed. The runoff variation in the source region critically affects and governs the variation in the available water resources in the entire river basin (Lu et al., 2020). Runoff and sediment discharge are the main output variables of river basins. They must show certain correlations during their evolution (Guo et al., 2020b; Varvani et al., 2019; Han et al., 2019; Tanzil et al., 2019). Because the source region plays a critical role in the Yellow River Basin, the variation in the runoff and sediment discharge in the region because of environmental changes significantly affect both the socioeconomic development (Zhang et al., 2008) and ecosystem maintenance (Yu et al., 2012) in the basin. Hence, to realize reasonable control of the runoff and sediment discharge in the Yellow River Basin and to understand the causes and mechanisms of runoff and sediment discharge variation, it is necessary to investigate the runoff-sediment discharge relationship and its detailed evolution in the source region.

At present, existing time-series research on runoff-sediment discharge relationships and their evolution characteristics are mostly on annual, seasonal, monthly, or daily scales (Zhang et al., 2006; Gao et al., 2016; Cui and Li., 2011; Jiang et al., 2017). Runoff and sediment discharge often show certain seasonal or periodic variation and such periodicity is usually multi-temporal(Zhang et al., 2019a; Ren et al., 2015; Prasad et al., 2019). This means that the runoff and sediment discharge variation in a certain time series does not follow certain fixed and simple patterns (such as those with constant periods). However, the variation includes different periodic changes and local fluctuations. This is one of the important evolution characteristics of complex non-linear systems. Runoff and sediment discharge in rivers show complex relationships not only for raw long-term time series. Complex fluctuation characteristics and relationships are also noted for time series on different time scales. Therefore, studies on runoff and sediment discharge should not only focus on the

macroscopic research on raw time series but also the detailed multi-temporal evolution characteristics of the series. Only in this manner, a comprehensive and in-depth understanding of the relationships between runoff and sediment discharge is possible. In recent years, multi-temporal analysis methods, such as wavelet analysis methods (Kuang et al., 2014; Nourani et al., 2019) and empirical mode decomposition (EMD) (Zhang et al., 2014), have been rapidly developed and combined with traditional hydrological methods to study the intrinsic relationships between hydrological variables and their evolution characteristics. These have become important approaches in hydrological research. However, pseudo-harmonics are found during decomposition using wavelet analysis methods, whereas EMD causes issues, such as mode mixing and end-point effects. Hence, the analysis results show certain deviations. Torres et al. (Torres et al., 2011; Colominas et al., 2014) proposed the Complete Ensemble Empirical Mode Decomposition with adaptive noise, which is an improved EMD algorithm. It can reasonably resolve the mode mixing problem of the original EMD and it is a relatively mature time-frequency analysis method.

Furthermore, runoff and sediment systems are highly complex, and they are influenced by various factors, such that the time series or multi-temporal component series of runoff and sediment discharge may be non-stationary (Chang et al., 2017). Nevertheless, previous studies on the time series of hydrological variables have usually assumed that the time series are stationary, and they have thereby constructed steady-state models. This may lead to pseudo-regression and certain errors in the analysis results. In economics, to avoid pseudo-regression during the construction of series models, cointegration theory has been proposed (Gu et al., 2017). Normally, the Engle–Granger two-step method (Engle and Granger., 1987) is adopted to determine whether long-term stable relationships exist. Runoff and sediment discharge do not only show long-term equilibrium relationships in their time series, but they also have short-term fluctuating relationships at different time scales. Unfortunately, most of the existing cointegration theory-based research on hydrological variables have investigated the entire time series (Zhang et al., 2013; Bello et al., 2018). Only a few studies have considered short-term fluctuating relationships with the help of multi-temporal analysis methods. Moreover, because of the effects of various factors, such as environmental and climate changes and human activities (Hu et al., 2019; Zhang et al., 2019b), structural breaks may be present for runoff and sediment discharge. This leads to variation in their relationships in raw time series or multi-temporal component series. Hence, it is necessary to combine the cointegration theory and multi-temporal analysis methods to analyze multi-temporal runoff-sediment discharge correlations and the structural breaks under the changing environment. The multi-temporal component model based on variable structure cointegration can be subsequently constructed to better reflect the runoff-sediment discharge relationship. This novel approach is herein adopted for the first time in the field.

This study first employed CEEMDAN to decompose the runoff and sediment discharge series of the source region of the Yellow River. Next, double cumulative curves were used to analyze the evolution characteristics and structural breaks of the multi-temporal runoff-sediment discharge correlations. Furthermore, the cointegration theory was used to analyze the runoff-sediment discharge relationships of different time series. For the time series with structural breaks, corresponding variable structure cointegration models were established. Their results were compared to examine their accuracy. Reasonable models were then selected to simulate and predict runoff.

## 2 Methodology and data

### 2.1 CEEMDAN

Suppose a signal is denoted as $x(t)$, and$b_i(t)$ is the white noise series with a mean value of 0 and a variance value of 1. $D_k(.)$ is the k-th order modal operator generated by EMD method.. To overcome the mode mixing caused by the previous two modes, $D_k(b_i(t))$ is used to extract the k-th mode. $E(.)$ is the operator that generates the local mean of the signals to be decomposed. The procedures of CEEMDAN are explained as follows:

(1) Add Gaussian white noise to the original series and the signal of the $i$ -th realization can be expressed as:

$x_i(t) = x(t) + \beta_0 D_1(b_i(t))(1)$

(2) Use the EMD method to calculate the mean value of the local mean of the signal $x_i(t)$ with the white noise. Then, the first residual value is obtained as:

$r_1(t) = \frac{1}{I}\sum_{i=1}^{I} E(x(t) + \beta_0 D_1(b_i(t)))(2)$

(3) With the first residual value, the first intrinsic mode function is calculated as:

$d_1(t) = x(t) - r_1(t)$ (3)

(4) Take the mean value of the local mean of$r_1(t) + \beta_1 D_2(b_i(t))$as the estimated value of the second residual value$r_2(t)$, then the second intrinsic mode function is:

$d_2(t) = r_1(t) - r_2(t)(4)$

(5) For k=3, ..., K, the k-th residual value is described as:

$r_k(t) = \frac{1}{I}\sum_{i=1}^{I} E(x(t) + \beta_{k-1} D_k(b_i(t)))(5)$

(6) From the k-th residual value, calculate the k-th intrinsic mode function:

$d_k(t) = r_{k-1}(t) - r_k(t)(6)$

(7) Repeat steps (5) to (6) until the residual $r_k(t)$satisfies one of the following conditions (Adarsh and Reddy., 2018): it cannot be further decomposed by EMD method; it satisfies the*IMF* condition; the number of the local extrema is less than three.

Thus, the original signal $x(t)$ can be decomposed into the $k$ -number of *IMF* components and one trend term$r_K(t)$:

$x(t) = \sum_{k=1}^{K} d_k(t) + r_K(t)(7)$

### 2.2 Double cumulative curves

Double cumulative curves are often plotted to examine the correlations between variables and their variation. For example, they are adopted to analyze the consistency, trends, and intensity of the changes in hydrological variables. (Zhang et al., 2017; Jiang et al., 2012) The construction of the double cumulative curve for two variables, A and B, is described as follows. It is assumed that A denotes the independent variable, whereas B is the dependent variable. The time series covers N years. The values of different years are expressed as $A_i$ and$B_i$. The annual cumulative values were obtained for variables A and B to provide a new cumulative series $A_i^{'}$ and$B_i^{'}$, where $i = 1, 2, 3 \cdots N$. They are expressed as follows:

$A_i^{'} = \sum_{i=1}^{N} A_i$ (8)

$B_i^{'} = \sum_{i=1}^{N} B_i$ (9)

In the double cumulative curve, the continuous cumulative value of variable A is plotted against that of B. If a straight line results, the variables vary proportionally with each other, and the slope of the line provides a constant ratio.

Nevertheless, if the slope of the curve changes at a certain point, a structural break exists. At the structural break, the relationship between the two variables is altered. The year in which the slope changes is the time of the structural break (Peng et al., 2013). Kohler (Kohler., 1949) believed that double cumulative curves give accurate results only if the dependent and independent variables exhibit strong correlations, direct proportionality, and reasonable comparability during the period of interest.

### 2.3 Cointegration theory

#### 2.3.1 Cointegration model

The cointegration theory comes from the field of economics. Whether long-term cointegration relationships exist for non-linear series can be determined by applying this theory (Gao et al., 2019; Boukhelkhal and Bengana., 2018; Heberling et al., 2015). This theory was proposed by Engle and Granger in 1987 to process a non-stationary series. Only when the variables of the non-stationary series become integrated at the same order after differentiation, valid cointegration regression models can be constructed. In general, cointegration relationships are tested through two steps: (1) The time series are subjected to unit root tests (augmented Dickey–Fuller, ADF tests) to determine whether the variable series are stationary. (2) The Engle–Granger two-step method is adopted to analyze whether cointegration relationships exist between the variables. The method is illustrated below:

First, for observed data series $\{x_t\}$ and $\{y_t\}$, ordinary least squares (OLS) are used to estimate $\alpha$ and $\beta$ and to calculate the residual series $\varepsilon_t$. A cointegration regression equation for the two series is given as follows:

$$y_t = \alpha + \beta x_t + \varepsilon_t \quad (10)$$

Second, the stationarity of the residual series is tested.

The residual series $\varepsilon_t$ is subjected to the ADF test. If it is a stationary series, then $\{x_t\}$ and $\{y_t\}$ exhibit a cointegration relationship. On the contrary, if the residual series is non-stationary (a unit root exists), then no cointegration relationship is present for $\{x_t\}$ and $\{y_t\}$.

### 2.3.2 Variable structure cointegration model

Because of the effect of the external environment, variables may show structural breaks. Hence, the long-term stable relationships between variables may vary. The cointegration relationships before and after structural breaks reflect the original and current long-term stable relationships, respectively. Thus, when there are significant structural variations between variables, cointegration analysis has to consider structural breaks as well (Singh, 2015; Vicente, 2014). When economic structures or policy systems are altered, parametric cointegration is normally adopted.

The point of the structural break is first determined. To construct a variable structure cointegration model, it is then assumed that the structural break is mainly caused by series $x_t$. A virtual variable is introduced:

$$D_t = \{$$

$$\begin{matrix} 0, t \le T_\tau \\ 1, t > T_\tau \end{matrix} \quad (11)$$

where, $T_\tau$ denotes the time of the structural break.

The cointegration parameter variations of a variable structure cointegration model can be primarily divided into the following three scenarios:

Scenario 1: Variable structure cointegration because of a constant term shift

In this case, only the variation in the constant term c of the model is considered. The following resulted:

$$y_t = c_1 + D_t c_2 + \alpha^T x_t + \varepsilon_t, \quad t = 1, 2, 3 \cdots T (12)$$

where, $c_1$ is the constant term before the shift and $c_2$ is the amount of the shift.

Scenario 2: Variable structure cointegration because of shifts in both the constant term and trend term

The variation in both the constant term and trend term is considered. This gives the following:

$$y_t = c_1 + D_t c_2 + \beta t + \alpha^T x_t + \varepsilon_t, \quad t = 1, 2, 3 \cdots T (13)$$

where, $\beta$ denotes the coefficient of the time trend term.

Scenario 3: State switch variable structure cointegration model

In this case, the variation in the constant term, trend term, and cointegration vector term are taken into consideration.

$$y_t = c_1 + D_t c_2 + \beta t + \alpha_1^T x_t + D_t \alpha_2^T x_t + \varepsilon_t, \quad t = 1, 2, 3 \cdots T (14)$$

### 2.4 Data sources

The source region of the Yellow River refers to the region from the headwater of the river to the Tangnaihai hydrometric station. It is a unique geographical location, which makes it the most important water-contributing region in the Yellow River Basin. It is also known as the "water tower" of the river (Figure 1). The Tangnaihai hydrometric station monitors a basin area of $1.22 \times 10^5 \mathrm{km}^2$. The area has an average annual runoff of $2.051 \times 10^{10}$ $\mathrm{m}^3$. This station plays a critical role in monitoring and regulating the flow of the Yellow River. This paper analyzed the annual runoff and sediment discharge data measured at the Tangnaihai station during 1960-2013, as illustrated in Figure 2.

——————————place Figure 1 here——————————-

——————————place Figure 2 here——————————-

## 3 Analysis results

### 3.1 Multi-temporal analysis

CEEMDAN was adopted for the multi-temporal decomposition of the annual runoff and sediment discharge series measured at the Tangnaihai station during 1960–2013. Finally, five layers were obtained. There are four were IMF components and one was a trend term (residual (RES) component). The decomposition results are shown in Figure 3.

——————————place Figure 3 here——————————-

Figure 3 reveals that the annual runoff and sediment discharge series measured at Tangnaihai station in the source region of the Yellow River showed complex multi-period variations and fluctuations. Although they have different amplitude fluctuations, the runoff and sediment discharge components of the same frequency vary simultaneously. This indicates that the runoff and sediment discharge in the source region are reasonably correlated with each other on both the macroscopic and microscopic scales. The IMF1 component shows the greatest amplitude, the shortest period, and the highest frequency. The RES component reflects the macroscopic variations in the runoff and sediment discharge and it demonstrates that both variables decline gradually with time.

For all the decomposition series, their periods and amplitude variation were analyzed. The results are summarized in Table 1.

——————————place Table 1 here——————————-

According to Table 1, the quasi-periods of the IMF runoff components are the same as those of the corresponding IMF sediment discharge components. The IMF1 components for both the runoff and sediment discharge have high frequencies and short quasi-periods. The IMF2 components have medium frequencies and quasi-periods. Furthermore, IMF3 and IMF4 are low-frequency components and they have medium/long and long periods, respectively. Overall, the minimum, maximum, and average amplitudes of the IMF components decreased progressively. As the decomposition scale increased, the amplitude fluctuations of the components declined and the periods lengthened. Lastly, the annual runoff and sediment discharge declined gradually macroscopically.

### 3.2 Multi-period evolution of the runoff-sediment discharge relationship

#### 3.2.1 Correlation of the raw series

Double cumulative curves were plotted to analyze the long-term variation and years with structural breaks for the relationship between the raw annual runoff series (X) and the raw annual sediment discharge series

(Y) observed at Tangnaihai station, as illustrated in Figure 3.

———————————place Figure 4 here———————————-

Figure 4 shows that the runoff and sediment discharge measured at Tangnaihai station were significantly positively correlated. All the data points were located close to the trendline. The annual cumulative runoff and the annual cumulative sediment discharge formed a straight line in the Cartesian coordinate system. The coefficient of determination $R^2$ equaled 0.9958, indicating that the runoff and sediment discharge in the source region were highly correlated. The overall fitted curve had a slope of 6.53, suggesting that $6.53 \times 10^3$ t of sediments would be generated for every $1 \times 10^8$ $m^3$ of runoff. The curve during 1989–2004 shifted slightly upward, indicating more sediment was created per unit runoff. This was because the Longyangxia Reservoir started to operate in 1987, leading to river closure for some channels. This has gradually affected the water and sediment transport processes upstream, i.e., at Tangnaihai station. The slope variation suggested that the influence on the sediment discharge was greater than that of runoff. The curve during 2005–2013 shifted slightly downward, indicating less sediment was discharged per unit runoff. During this period, there was less runoff and upstream vegetation was affected by water and sediment retaining projects in the region. Hence, consistent variation in the runoff and sediment discharge was changed, and thus, their relationship was also altered.

### 3.2.2 Multi-temporal correlation

The negative values of the multi-temporal components were eliminated with the help of equivalent substitution. Subsequently, double cumulative curves were plotted for the multi-temporal components of runoff and sediment discharge to study their multi-temporal correlations, detailed evolution, and structural breaks. The double cumulative curves of the multi-temporal components are shown below.

———————————place Figure 5 here———————————-

From Figure 5:

(1) Higher goodness of fit is noted in the runoff-sediment discharge double cumulative curves of the IMF1 and IMF2 components. The corresponding $R^2$ values are 0.9996 and 0.9995, respectively. The curve of IMF3 shows structural breaks in 1978, 1989, 2001, and 2005, and that in 2005 was the most prominent. Similarly, structural breaks were noted in 1994, 1997, and 2001 for the curve of IMF4, and the most prominent occurred in 2001. Furthermore, the double cumulative curve of the RES component demonstrated that the runoff and sediment discharge showed consistent variation macroscopically. In short, the runoff-sediment discharge double cumulative curves were different at different time scales. At some time scales, more distinct structural breaks were present, indicating significant variation in the runoff-sediment discharge relationship.

(2) The slopes of the trendlines for the multi-temporal components declined gradually. This suggests that, microscopically, the components had different amplitudes and fluctuations. Local characteristics of the runoff and sediment discharge variation became observable only when the components were analyzed separately. Thus, different slopes were found for the trendlines of different double cumulative curves and the runoff-sediment discharge correlations varied at different time scales. For the IMF1 and IMF2 components, higher goodness of fit was observed in their double cumulative curves. However, lower goodness of fit was noted for the curves of the IMF3 and IMF4 components. This indicated that medium- and high-frequency components exhibited stronger runoff-sediment discharge correlations, whereas low-frequency ones demonstrated weaker correlations. Additionally, the RES component reflected the macroscopic variation of water and sediment systems. Reasonable goodness of fit was observed for its double cumulative curve, demonstrating a relatively strong runoff-sediment discharge correlation at the macroscopic scale. Hence, the characteristics of the macroscopic variations in the runoff and sediment discharge reflected by the RES component should be examined to study the runoff-sediment discharge relationship and its evolution.

(3) Overall, some points of structural breaks on the microscopic scale in the runoff-sediment discharge double cumulative curves of the multi-temporal components appeared as ordinary points on the macroscopic scale in the curve of the raw series. Hence, accurate identification of structural break points is closely related

to the time scales employed during the analysis. When the raw series was examined, characteristics of the macroscopic scale may bury or neutralize those of the microscopic scale. Thus, when structural breaks are concerned, more attention should be paid to the details. For relatively complex raw series, medium- and/or high-frequency modal components can be utilized to study the runoff-sediment discharge relationships and their detailed evolution. Research on runoff-sediment discharge relationships can focus on relevant short- and/or medium-term observations. When the relationships show change and information (for example, years), structural breaks have to be determined and low-frequency components can be employed.

### 3.3 Multi-temporal runoff-sediment discharge cointegration relationships

#### 3.3.1 ADF test

First, ADF tests were performed for the raw runoff and sediment discharge series. The results are summarized in Table 2.

———————————place Table 2 here———————————-

From the table, unit roots exist for X and Y before differentiation at the 1%, 5%, and 10% significance levels. Thus, the series are non-stationary. Contrarily, after the first-order differentiation, there are no unit roots for both series, and they are now stationary series. Therefore, X and Y are first-order integrated series, i.e.,$X \sim I(1)$ and $Y \sim I(1)$.

Subsequently, the multi-temporal runoff and sediment discharge components were subjected to stationarity tests. The ADF test results are listed.

———————————place Table 3 here———————————-

———————————place Table 4 here———————————-

According to the above test results, unit roots do not exist for the undifferentiated series of the IMF1, IMF2, and RES components of the runoff and sediment discharge at the 1%, 5%, and 10% significance levels. The series of these components are stationary. Additionally, relatively satisfactory goodness of fit and higher $R^2$values were noted for the double cumulative curves of the IMF1, IMF2, and RES components. In these plots, the cumulative runoff and sediment discharge gave a straight line, and they showed consistent variation. It is believed that the equations describing the runoff-sediment discharge relationships of the IMF1, IMF2, and RES components can be determined using linear regression. Furthermore, structural breaks are noted in the double cumulative curves of IMF3 and IMF4. The series of these components are found to be first-order integrated series according to the ADF tests. Thus, for these components, OLS was employed to estimate the regression parameters and to determine the cointegration equations.

#### 3.3.2 Multi-temporal regression equation

Linear regression was used to determine the runoff-sediment discharge regression equations for IMF1, IMF2, and RES. The corresponding scatter plots and linear regression equations are illustrated below.

———————————place Figure 6 here———————————-

From the scatter plots, high goodness of fit of the linear regression lines are noted for the components of runoff and sediment discharge measured at Tangnaihai station. The $R^2$ values for the IMF1, IMF2, and RES components were 0.8105, 0.8357, and 0.9975, respectively. These indicated that the runoff and sediment discharge were significantly correlated with these components. At these scales, linear regression equations can be used to express the runoff-sediment discharge correlations. Therefore, the Engle–Granger two-step method was necessary only for the cointegration regression for the raw series and the IMF3 and IMF4 components. After that, the residual series of these series were subjected to the ADF tests. The results are listed in Table 5.

———————————place Table 5 here———————————-

Table 5 shows that all the residual series were stationary. Thus, it was concluded that the raw, IMF3, and IMF4 series all show cointegration relationships.

By employing linear regression and the cointegration theory, regression equations were obtained for the raw series and the multi-temporal component series. They are summarized in Table 6.

———————————place Table 6 here———————————-

In general, higher goodness of fit was noted for the regression model for the raw series. This indicated that the runoff and sediment discharge were significantly correlated at the macroscopic scale. Yet, their relationships at different time scales varied slightly. Among the series subjected to linear regression, the RES component demonstrated the highest goodness of fit. This suggested that the runoff and sediment discharge showed a relatively strong correlation in terms of their overall variation. Among those subjected to cointegration regression, the IMF3 component series (indicative of a medium/long-time scale) and the IMF4 component series (long-time scale) showed lower goodness of fit. This demonstrated weak runoff-sediment discharge cointegration relationships at these time scales.

When the cointegration test results were integrated with the double cumulative curves, the runoff and sediment discharge generally observed at Tangnaihai station consistently varied with each other proportionally. However, the relationship variation of the raw runoff and sediment discharge series was caused by multi-temporal variations. The weak runoff-sediment discharge relationships on the medium/long and long-time scales would affect that of the raw series.

### 3.3.3 Variable structure cointegration model

Because structural breaks were observed for the IMF3 and IMF4 runoff-sediment discharge relationships, the corresponding cointegration relationships were weaker. Hence, variable structure cointegration models were constructed to determine reasonable models by considering structural breaks. The most prominent break was noted for the IMF3 component in 2005 and that for the IMF4 component was observed in 2001. In model construction, these points were treated as structural breaks and virtual variables $D_{1t}$ and $D_{2t}$ were introduced.

For the IMF3 component, a virtual variable was introduced as follows:

$$D_{1t} = \{ \begin{matrix} 0, t \leq 2005 \\ 1, t > 2005 \end{matrix} \quad (15)$$

Three models corresponding to different scenarios (section 2.3.2) were examined to determine the most reasonable model. The models considering three different variation scenarios are summarized in Table 7.

———————————place Table 7 here———————————-

The $R^2$ values of the models corresponding to scenarios 1, 2, and 3 were 0.606516, 0.788878, and 0.800332, respectively. These values were significantly higher than that of the original cointegration model (0.410943). The highest goodness of fit was achieved by model 3, therefore, it is believed that this model can more accurately depict the IMF3 runoff-sediment discharge relationship.

The virtual variable introduced into the model for the IMF4 component is given as follows:

$$D_{2t} = \{ \begin{matrix} 0, t \leq 2001 \\ 1, t > 2001 \end{matrix} \quad (16)$$

Table 8 shows the models considering the three scenarios:

———————————place Table 8 here———————————-

Under the three scenarios, the models give $R^2$ values of 0.792911, 0.807911, and 0.841282, respectively. They were much higher than that of the original cointegration model for the IMF4 component (0.342336). More specifically, model 3 demonstrated the highest goodness of fit. Therefore, it could more accurately describe the runoff-sediment discharge relationship variation for IMF4 component.

### 3.3.4 Model accuracy

With the help of the runoff-sediment discharge regression models established for the components, the runoff during the research period was simulated. The results are illustrated in Figure 7. The cointegration model constructed using the raw series was called an original model. The model built based on the multi-temporal components was named composite model 1. Lastly, composite model 2 refers to the variable structure cointegration model constructed using the multi-temporal components and considering structural breaks.

———————————place Figure 7 here———————————-

———————————place Figure 8 here———————————-

———————————place Figure 9 here———————————-

Figure 7 reveals that, in general, the models satisfactorily simulated the raw runoff. In particular, near the end of the research period, the composite model 2 showed higher accuracy. Its results showed the same variation trend as that of the raw series. From Figure 8, the average relative errors by the original model, the composite model 1, and the composite model 2 were 11.43%, 10.76%, and 7.82%, respectively. According to the Standard for Hydrological Information and Hydrological forecasting (GBT 22482-2008), an error tolerance of 20% is appropriate in hydrological forecasting. Hence, the relative errors of all the models fall within the acceptable range. In particular, the greatest average relative error resulted from the original model, whereas the smallest was generated by composite model 2 considering structural breaks. Figure 9 illustrates that this model yields relative errors less than 10% for 41 years out of the 54 years simulated. The relative errors by this model were 10%–20% for 9 years and they were greater than 30% for the remaining 4 years. Hence, most of the relative errors by this model were low, while few were high. Therefore, the resulting average relative error by composite model 2 was comparatively small. From Figure 8, at the beginning of the research period, the three models give the same relative error. However, the errors in the original model and the composite model 1 increased after 2000. The errors in 1997, 2002, and 2009 were the greatest. Furthermore, the composite model 2 provided more satisfactory runoff simulation results and smaller relative errors for these years. Because this model considered the local variation characteristics and structural breaks of the multi-temporal components, its overall simulation accuracy was higher. The model could more accurately reflect the long-term equilibrium and short-term fluctuating relationships between the runoff and sediment discharge in the source region of the Yellow River.

## 4 Conclusions

This paper combined CEEMDAN, double cumulative curves, and the cointegration theory to construct composite models that reflected the overall and multi-temporal relationships between the runoff and sediment discharge in the source region of the Yellow River. The models were then adopted for runoff simulation to evaluate their accuracy. The main conclusions were as follows:

(1) The raw series of both the runoff and sediment discharge of the source region of the Yellow River could be decomposed into five layers, including layers of four IMF components and one RES component. Significant runoff-sediment discharge correlations were noted at different time scales. The multi-temporal components of the same frequency varied simultaneously. The IMF components of the same variable had the same quasi-periods. Furthermore, the RES component revealed that the runoff and sediment discharge declined simultaneously with time.

(2) The multi-temporal runoff-sediment discharge double cumulative curves showed a high goodness of fit, and the multi-temporal components were significantly positively correlated. More specifically, those of the IMF1 and IMF2 components demonstrated higher goodness of fit, whereas those of IMF3 and IMF4 showed distinct

points of structural breaks. Therefore, for relatively complex raw series, high-frequency components could be employed to explore runoff-sediment discharge relationships and the characteristics of their detailed evolution. Thus, more attention could be paid to the relevant short- and/or medium-term observations. Furthermore, when attempts were made to obtain information on structural breaks, detailed evolution should be examined more carefully. When the runoff-sediment discharge relationships changed and the information (for example, years) on structural breaks had to be determined, low-frequency components could be employed.

(3) As demonstrated by the raw series, the runoff was reasonably correlated with the sediment discharge in the source region. However, their relationships at different time scales varied slightly. The relationship changes noted for the raw series were caused jointly by those at different time scales. Hence, the weak runoff-sediment discharge relationships on the medium/long and long-time scales would affect the relationship between the raw runoff and sediment discharge series.

(4) The runoff during the research period was simulated separately using the original runoff-sediment discharge cointegration model (based on the raw series), the composite model 1 (based on the multi-temporal components), and the composite model 2 (based on the multi-temporal components with the structural breaks). Their average relative errors were 11.43%, 10.76%, and 7.82%, respectively. The composite model 2 considered both the local variation characteristics of the multi-temporal components, as well as the structural breaks, and it showed higher simulation accuracy and provided smaller relative errors. It could more accurately reflect the long-term equilibrium and short-term fluctuating relationships between the runoff and sediment discharge in the source region of the Yellow River.

**Acknowledgements**

This research is supported by the National Key R&D Program of China (Grant No.2018YFC0406501), Outstanding Young Talent Research Fund of Zhengzhou University (Grant No.1521323002), Program for Innovative Talents (in Science and Technology) at University of Henan Province (Grant No.18HASTIT014), State Key Laboratory of Hydraulic Engineering Simulation and Safety, Tianjin University (Grant No. HESS-1717) and Foundation for University Youth Key Teacher of Henan Province (Grant No.2017GGJS006).

**References**

Adarsh, S., Reddy, M.J., 2018. Multiscale characterization and prediction of monsoon rainfall in India using Hilbert-Huang transform and time-dependent intrinsic correlation analysis. Meteorology and Atmospheric Physics, 130(6), 667-688.

Bai, T., Wei, J., Chang, O., Yang, W.W., Huang, Q., 2019. Optimize multi-objective transformation rules of water-sediment regulation for cascade reservoirs in the Upper Yellow River of China. Journal of Hydrology, 577, Article Number: UNSP 123987.

Bello, M.O., Solarin, S.A., Yen, Y.Y., 2018. The impact of electricity consumption on $CO_2$ emission, carbon footprint, water footprint and ecological footprint: The role of hydropower in an emerging economy. Journal of Environmental Management, 219, 218-230.

Boukhelkhal, A., Bengana, I., 2018. Cointegration and causality among electricity consumption, economic, climatic and environmental factors: Evidence from North -Africa region. Energy, 163, 1193-1206.

Cui, B.L., Li, X.Y., 2011. Coastline change of the Yellow River estuary and its response to the sediment and runoff (1976-2005). Geomorphology, 127(1-2), 32-40.

Colominas, M.A., Schlotthauer, G., Torres, M.E., 2014. Improved complete ensemble EMD: A suitable for biomedical signal processing. Biomedical Signal Processing and Control,14(11), 19-29.

Chang, N.B., Yang, Y.J., Imen, S., Mullon, L., 2017. Multi-scale quantitative precipitation forecasting using nonlinear and nonstationary teleconnection signals and artificial neural network models. Journal of Hydrology, 548, 305-321.

Engle, R.F., Granger, C.W.J., 1987. Cointegration and Error Correction: Representation, Estimation, and Testing. Econometrica, 55,252-276.

Gu, C.J., Mu, X.M., Gao, P., Zhao, G.J., Sun, W.Y., Tan, X.J., 2019. Distinguishing the effects of vegetation restoration on runoff and sediment generation on simulated rainfall on the hillslopes of the loess plateau of China. Plant and Soil, 447 (1-2), 393-412.

Guo, Q.C., Zheng, Z., Huang, L.M., Deng, A.J., 2020a. Regularity of sediment transport and sedimentation during floods in the lower Yellow River, China. International Journal of Sediment Research, 35 (1), 97-104.

Guo, W.X., Li, Y., Wang, H.X., Cha, H.F., 2020b. Temporal variations and influencing factors of river runoff and sediment regimes in the Yangtze River, China. Desalination and Water Treatment, 174, 258-270.

Gao, G.Y., Ma, Y., Fu, B.J., 2016. Multi-temporal scale changes of streamflow and sediment load in a loess hilly watershed of China. Hydrological Processes, 30(3), 365-382.

Gu, A., Zhang, Y., Pan, B.L., 2017. Relationship between Industrial Water Use and Economic Growth in China: Insights from an Environmental Kuznets Curve. Water, 9(8), Article Number: 556.

Gao, F., Tong, M.L., Gao, Y.P., Yang, T.G., Zhao, C.S., 2019. The application of co-integration theory in ensemble pulsar timescale algorithm. Research in Astronomy and Astrophysics, 19(7), Article Number: 100.

Han, J.Q., Gao, J.N., Luo, H., 2019. Changes and implications of the relationship between rainfall, runoff and sediment load in the Wuding River basin on the Chinese Loess Plateau, Catena, 175, 228-235.

Hu, J.F., Zhao, G.J., Mu, X.M., Tian, P., Gao, P., Sun, W.Y., 2019. Quantifying the impacts of human activities on runoff and sediment load changes in a Loess Plateau catchment, China. Journal of Soils and Sediments, 19(11), 3866-3880.

Heberling, M.T., Nietch, C.T., Thurston, H.W., Elovitz, M., Birkenhauer, K.H., Panguluri, S., Ramakrishnan, B., Heiser, E., Neyer, T., 2015. Comparing drinking water treatment costs to source water protection costs using time series analysis. Water Resources Research, 51(11), 8741-8756.

Jiang, C., Zhang, L.B., Tang, Z.P., 2017. Multi-temporal scale changes of streamflow and sediment discharge in the headwaters of Yellow River and Yangtze River on the Tibetan Plateau, China. Ecological Engineering, 102, 240-254.

Jiang, S.H., Ren, L.L., Yong, B.; Fu, C.B., Yang, X.L., 2012. Analyzing the effects of climate variability and human activities on runoff from the Laohahe basin in northern China. Hydrology Research, 43(1-2), 3-13.

Kuang, C.P., Su, P., Gu, J., Chen, W.J., Zhang, J., Zhang, W.L., Zhang, Y.F., 2014. Multi-Time Scale Analysis of Runoff at the Yangtze Estuary Based on the Morlet Wavelet Transform Method. Journal of Mountain Science, 11(6), 1499-1506.

Kohler M. A., 1949. On the Use of Double-mass analysis for testing the consistency of meteorological records and for making required adjustments. Bull. Ann.Meteol. Soc, 30, 188-189.

Li, E.H., Mu, X.M., Zhao, G.J., Gao, P., Sun, W.Y., 2017. Effects of check dams on runoff and sediment load in a semi-arid river basin of the Yellow River. Stochastic Environmental Research and Risk Assessment, 31, 1791-1803.

Lu, Z.X., Feng, Q., Zou, S.B., Xie, J.L., Yin, Z.L., Li, F., 2020. The heterogeneity of hydrometeorological changes during the period of 1961-2016 in the source region of the Yellow River, China. Sciences in Cold and Arid Regions, 12 (2), 104-118.

Nourani, V., Molajou, A., Tajbakhsh, A.D., Najafi, H., 2019. A Wavelet Based Data Mining Technique for Suspended Sediment Load Modeling. Water Resources Management, 33(5), 1769-1784.

Prasad, R., Deo, R.C., Li, Y., Maraseni, T., 2019. Weekly soil moisture forecasting with multivariate sequential, ensemble empirical mode decomposition and Boruta-random forest hybridizer algorithm approach. Catena, 177,149-166.

Peng, S.Z., Liu, W.X., Wang, W.G., Shao, Q.X., Jiao, X.Y., Yu, Z.B., Xing, W.Q., Xu, J.Z., Zhang, Z.X., Luo, Y.F., 2013. Estimating the Effects of Climatic Variability and Human Activities on Streamflow in the Hutuo River Basin, China. Journal of Hydrologic Engineering, 18(4), 422-430.

Ren, H.R., Li, G.S., Cui, L.L., He, L., 2015. Multi-scale variability of water discharge and sediment load into the Bohai Sea from 1950 to 2011. Journal of Geographical Sciences, 25(1), 85-100.

Singh T., 2015. On the International Trade and Economic Growth Nexus in New Zealand. Economic Papers: A journal of applied economics and policy, 34(1-2), 92-106.

Tanzil, J.T.I., Goodkin, N.F., Sin, T.M., Chen, M.L., Fabbro, G.N., Boyle, E.A., Lee, A.C., Toh, K.B., 2019. Multi-colony coral skeletal Ba/Ca from Singapore's turbid urban reefs: Relationship with contemporaneous in-situ seawater parameters. Geochimica Etcosmochimica Acta, 250, 191-208.

Torres, M.E., Colominas, M.A., Schlotthauer, G., Flandrin, P., 2011. A complete ensemble empirical mode decomposition with adaptive noise. Proc. 36th IEEE Int. Conf. on Acoust, Speech and Signal Process. ICASSP 2011, Prague, Czech Republic, 4144-4147.

Varvani, J., Khaleghi, M.R., Gholami, V., 2019. Investigation of the Relationship between Sediment Graph and Hydrograph of Flood Events (Case Study: Gharachay River Tributaries, Arak, Iran). Water Resources, 46 (6), 883-893.

Vicente G.S., 2014. Assessing the Historical Water Flow Allocation in the Lower Rio Grande between Mexico and the United States. Journal of Borderlands Studies, 29(2), 275-289.

Yu, M.Z., Xu, X.X., Liu, P.L., Zheng, S.Q., 2012. The response of runoff and sediment on eco-environmental change of Yan'gou watershed in the loess hilly region. Journal of Food Agriculture & Environment, 10(1), 941-945.

Zhao, X.N., Chen, X.L., Huang, J., Wu, P.T., Helmers, M.J., 2014. Effects of vegetation cover of natural grassland on runoff and sediment yield in loess hilly region of China. Journal of the Science of Food and Agriculture, 94(3), 497-503.

Zhang, Q., Chen, G.Y., Su, B.D., Disse, M., Jiang, T., Xu, C.Y., 2008.Periodicity of sediment load and runoff in the Yangtze River basin and possible impacts of climatic changes and human activities. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 53(2), 457-465.

Zhang, Q., Xu, C.Y., Becker, S., Jiang, T., 2006. Sediment and runoff changes in the Yangtze River basin during past 50 years. Journal of Hydrology, 331(3-4), 511-523.

Zhang, J.P., Xiao, H.L., Zhang, X., Li, F.W., 2019a. Impact of reservoir operation on runoff and sediment load at multi-time scales based on entropy theory. Journal of Hydrology, 569, 809-815.

Zhang, J.P., Ding, Z.H., You, J.J., 2014. The joint probability distribution of runoff and sediment and its change characteristics with multi-time scales, Journal of Hydrology and Hydromechanics, 62(3), 218-225.

Zhang, J.P., Yuan, W.L., Guo, B.T., 2013. Study on prediction of stream flow based on cointegration theory. Water Resources and Power, 31(5), 18-20, 99.

Zhang, X.F., Yan, H.C., Yue, Y., Xu, Q.X., 2019b. Quantifying natural and anthropogenic impacts on runoff and sediment load: An investigation on the middle and lower reaches of the Jinsha River Basin. Journal of Hydrology-Regional Studies, 25, Article Number: UNSP 100617.

Zhang, L., Karthikeyan, R.; Bai, Z.; Wang, J., 2017. Spatial and temporal variability of temperature, precipitation, and streamflow in upper Sang-kan basin, China. Hydrological Processes, 31(2), 279-295.

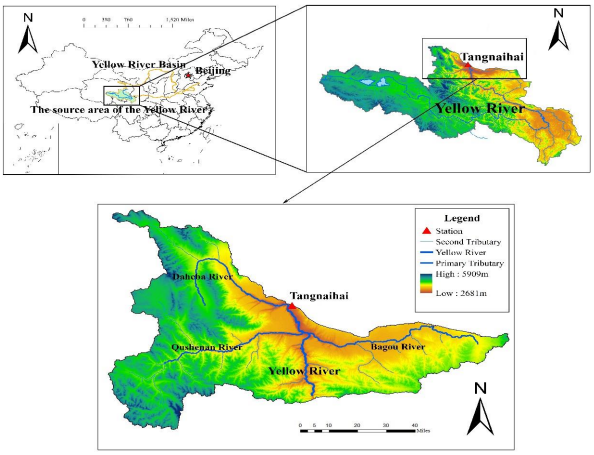


Figure 1 Location of the study area with the Tangnaihai station

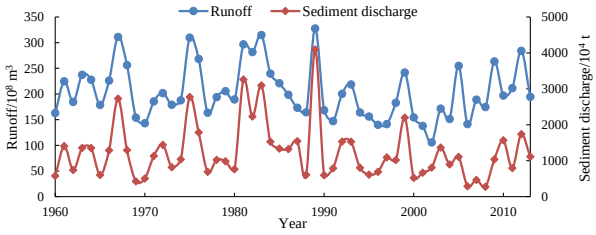


Figure 2 Variation in the runoff and sediment discharge series observed at the Tangnaihai station

## Hosted file

`Figures.pdf` available at https://authorea.com/users/361760/articles/483106-multi-temporal-runoff-sediment-discharge-relationships-in-the-source-region-of-the-yellow-river

## Hosted file

`Tables.pdf` available at https://authorea.com/users/361760/articles/483106-multi-temporal-runoff-sediment-discharge-relationships-in-the-source-region-of-the-yellow-river