Journal of Hydrology 544 (2017) 343–351

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

# Wavelet-cointegration prediction of irrigation water in the irrigation district

Jinping Zhang [a,*], Yanyan Li [a], Yong Zhao [b,*], Yang Hong [c]

[a] School of Water Conservancy & Environment, Zhengzhou University, Zhengzhou 450001, China
[b] State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Beijing 100038, China
[c] School of Civil Engineering and Environmental Sciences, University of Oklahoma, Norman, OK 73019, USA

ARTICLE INFO

*Article history:*
Received 15 July 2016
Received in revised form 4 November 2016
Accepted 19 November 2016
Available online 23 November 2016
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Fi-John Chang, Associate Editor

*Keywords:*
Wavelet-cointegration prediction
Irrigation water
Rainfall
Crop water requirement

ABSTRACT

Affected by various influencing factors, many time series analysis methods are used to predict irrigation water. However, it is often assumed to be stationary for the hydrologic time series which will result in an insignificant "spurious regression". Actually, irrigation water is largely determined by rainfall and crop water requirement in the irrigation district. With the wavelet analysis method, the varying fluctuation characteristics of rainfall, irrigation water, and crop water requirement with multi-temporal scales are exhibited. The cointegration equilibrium relationship amongst their original time series and decomposed time series are also revealed by the cointegration theory. Combining the wavelet analysis method with the cointegration theory, the wavelet-cointegration prediction model of irrigation water is proposed. The results show that the model has the available prediction accuracy, and the relative errors of all predicted years are less than 5%, except 2004 and 2012.



## 1. Introduction

For irrigation districts, irrigation water is a main economic indicator. The rational prediction of irrigation water is important to the limited water resources, and also has a great significance for the irrigation water planning and agricultural development. Actually, irrigation water is affected by many factors, such as climate changes, water engineering, and crop planting, which results in the inter-annual variation of irrigation water. So it is difficult to construct a determined model to measure.

In practice, irrigation quota is often used to estimate irrigation water. This method is simple, but it needs more practical experience (Bai and Li, 2003). So it is adoptable for the rough estimation of irrigation water. Some scholars explored the mechanism of crop water requirement to determine irrigation water. They calculated crop evapotranspiration with the Penman formula, and considered rainfall, farm seepage, etc. to predict irrigation water according to the water balance in the irrigation district ( Allen et al., 1998; Kuo et al., 2006; Trajkovic and Kolakovic, 2009; Usman et al., 2015; Guo and Shen, 2016). This method needs a great amount of basic data, including meteorological, soil, crop planting, etc. and sometimes it needs a hydrological model, so this may have problems in practical application. Moreover, the irrigation water calculation by FAO needs a crop coefficient, which is usually determined by irrigation experiments. Because it is affected by the actual soil, climate and growth conditions, etc., the crop coefficient often varies, so it must be modified. Currently, researchers apply the time series analysis methods into the prediction of irrigation water and try to find the development patterns of either irrigation water itself or the relations between irrigation water and other influencing factors with the long data series. Considering the complexity and randomization of the influencing factors on irrigation water, some uncertainty analysis methods (e.g. grey theory, fuzzy mathematics, artificial neural network, etc.) are introduced (Elshorbagy and Parasuraman, 2008; Kim and Kim, 2008; Zou et al., 2010; Zhang et al., 2014; Abdullah et al., 2015; Mattar et al., 2015; Si et al., 2015). With these combination methods, the irrigation water prediction accuracy is improved.

However, no matter how much progress the previously studies made, these studied results were all with the same hypothesis: the hydrologic time series are stationary. Actually, the statistic characteristics of hydrological variables always change in time, so the hydrological time series is non-stationary. However, the traditional econometric method assumes that the stationary and normality exist in hydrological time series. This assumption will result in "spurious regression." It means that there are no dependencies

* Corresponding authors.
*E-mail addresses:* jinping2000_zh@163.com (J. Zhang), iwhrzhy@sohu.com (Y. Zhao).

http://dx.doi.org/10.1016/j.jhydrol.2016.11.040

between variables originally, but the regression results indicate the presence of dependencies. Fortunately, the wavelet analysis provides a powerful tool for non-stationary signals (Xu and Liu, 2013). With the wavelet analysis method, the hydrological time series can be decomposed into different temporal scales with the relatively stable statistics characteristics to reflect its fluctuation features and development patterns (Compagnucci et al., 2000; Gaucherel, 2002; Kahya and Kalayci, 2008; Can et al., 2005; Battista and Visini, 2006; Parmar and Bhardwaj, 2013). Usually, the wavelet analysis method can be integrated with other mathematical methods and technologies (Avci, 2007; He et al., 2008; Chenini and Khemiri, 2009; Rajaee et al., 2010; Yilmaz and Oysal, 2010; Abiyev, 2011; Zhang et al., 2013; Venkata et al., 2013; Aman, 2015). Moreover, the cointegration theory proposed by Engle and Granger (1987) can overcome this shortcoming. The cointegration theory can deal with the non-stationary time series of variables, and it describes the short-term dynamic fluctuations and long-term balance relationship between variables. The cointegration theory has been used in econometrics for decades (Loomis, 1996; Levin et al., 2002; Han et al., 2004; Basher and Sadorsky, 2006; Chen et al., 2010; Sebri, 2015; Ghosh and Kanjilal, 2016), and also in hydrological fields in recent years (Yoo and Yang, 1999; Cole, 2004; Seung-Hoon, 2007), but it is rare for the relationship among rainfall, irrigation water, and crop water requirement.

Combing the cointegration theory with the wavelet analysis method, Zhang et al. (2015) have proposed the multi-resolution cointegration prediction model for runoff and sediment load in Weihe River of China. However, as we know, the water issues of irrigation districts are distinct from that of rivers. In irrigation districts, "water supply" involves rainfall and irrigation water, while "water demand" refers to crop water requirement ($ET_c$). Affected by climate changes and human activities, these three variables have some relations among them. The irrigation water in irrigation districts is largely determined by rainfall and $ET_c$. Also from the view of hydrological cycle in irrigation districts, irrigation water and rainfall are "water input" and $ET_c$ is "water output". So their relations are relatively complex.

Therefore, the objective of this paper is that the fluctuation characteristics of rainfall, $ET_c$ and irrigation water with multi-temporal scales are exhibited using the wavelet analysis method firstly. Then with the cointegration theory, the short-term fluctuations and the long-term balance relationships of the original time series and the decomposed time series are all presented. Finally, combing the wavelet analysis method with the cointegration theory, the wavelet-cointegration prediction of irrigation water is established. So, this method not only presents the cointegration relations between variables with the different decomposed time scales, but also ensures the prediction accuracy of irrigation water with the reconstructed decomposed time series.

## 2. Materials and methods

### 2.1. Wavelet analysis

The wavelet function can be defined as: set $\varphi(t)$ belongs to a two-dimensional space $L^2(R)$, if its Fourier transform $\hat{\varphi}(w)$ satisfies the permitted condition $C_\varphi = \int_R |\hat{\varphi}(w)2/wdw < \infty|$, $\varphi(t)$ is called a mother wavelet or basic wavelet. The telescopic or translation of mother wavelet $\varphi(t)$ can be conducted as follows (Mallat, 1998):

$$\varphi_{a,b}(t) = 1/\sqrt{|a|}\varphi((t-b)/a)a, b \in R, a > 0 \tag{1}$$

where $a$ is the scale parameter, and $b$ is the translation parameter.

The continuous wavelet transform of any function $f(t) \in R^2$ is (Mallat, 1998):

$$W_f(a,b) = f(t), \varphi_{a,b}(t) = 1/\sqrt{|a|}\int f(t)\varphi((t-b)/a)dt \tag{2}$$

where $W_f(a,b)$ is called the wavelet coefficient, its reconstruction equation is calculated as (Mallat, 1998):

$$f(t) = \left(\frac{1}{C_\varphi}\right)\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(1/a^2)W_f(a,b)\varphi_{a,b}(t)db \tag{3}$$

The hydrological time series is often a discrete time series. Suppose the discretization of scale parameter $a$ and translation parameter $b$ are written as $a = a_0^j$, $a_0 > 1 j \in Z$; $b = ka_0^j b_0$, $b_0 > 0$, $k \in Z$, the discrete wavelet function can be described as (Mallat, 1998):

$$\varphi_{j,k}(t) = a_0^{-\frac{j}{2}}\varphi\left((t-ka_0^j b_0)/a_0^j\right) \tag{4}$$

The discrete wavelet transform wavelet coefficient can be expressed as (Mallat, 1998):

$$W_f(j,k) = f(t), \varphi_{j,k}(t) = \int_{-\infty}^{+\infty} f(t)\varphi_{j,k}(t)dt \tag{5}$$

Thus, its reconstruction equation is calculated as (Mallat, 1998):

$$f(t) = C_\varphi \sum_{j\in N, k\in N} W_f(j,k)\varphi_{j,k}(t) \tag{6}$$

### 2.2. Cointegration theory

#### 2.2.1. Cointegration concept

Cointegration theory is commonly used in econometrics to describe the dynamic relationship between two or more economic variables over the long run. According to Engle and Granger (1987), if a time series is non-stationary, but it becomes stationary after the $d$ differencing, the original time series is denoted as $I(d)$, which means it is integrated of order $d$. Suppose a time series defined as $X_t = (x_{1t}, x_{2t}, \ldots, x_{nt})^T$, if it satisfies the following conditions (Engle and Granger, 1987):

(1) $x_{it}$ Is $I(d)$ $(i = 1, 2, \ldots n)$, $d$ is integer;
(2) the presence of a non-zero vector makes that a particular combination of them $Z_t - \alpha^T X_t = \varepsilon_t$ is stationary at the level form, that is $Z_t - \alpha^T X_t = \varepsilon_t \sim I(0)$.

Thus, $X_t$ is cointegrated, where $\alpha$ is a cointegrated vector.

#### 2.2.2. Stationary test

Prior to the cointegrated test, the stationary of data series must be tested. The unit root test based on the augmented Dickey-Fuller test is widely employed to test the stationary of data series. The ADF test is as the following (Dickey and Fuller, 1979):

$$\Delta y_t = \alpha + \beta t + \delta y_{t-1} + \sum_{i=1}^{p}\xi_i \Delta y_{t-i} + \varepsilon_t \tag{7}$$

where $\Delta y_t = y_t - y_{t-1}$ is the first difference of $y_t$, $\alpha$, $\beta$, $\delta$, $\xi_i$ are parameters, $t$ is time, $p$ is the lag length and $\varepsilon_t$ is the pure white noise error term.

The optimal lag length $p$ is obtained by the smallest Akaike information criterion (AIC) or Schwartz criterion (SC) (Schwarz, 1978). The null hypothesis of $H_0$: $\gamma = 0$ means no cointegration. The ADF test involves the modified t-statistic allocated with the ordinary least squares (OLS) method, and it ensures the null hypothesis is accepted. The ADF test statistic is a negative number, and the more negative it is, the stronger the rejection of the hypothesis that there is a unit root. If the ADF test statistic is positive, one can automatically decide not to reject the null hypothesis of unit root. ADF test statistics are computed at significance levels