

Hybrid Wavelet and Chaos Theory for Runoff Forecasting

Chen Xi, Jiang Chuanwen, Wang Yu, Zhou Jian
Department of Electrical Engineering, Shanghai Jiaotong University
Shanghai 200030, P.R.China

Abstract: -This paper introduced a method of decomposing non-stationary runoff time series. By wavelet decomposing, the runoff time series is decomposed into stationary time series and stochastic time series, and AR(n) model be imposed for forecasting stationary time series. By studying chaos characteristic of stochastic time series, this paper put forward a nonlinear chaos dynamics-forecasting model to dispose runoff time series with high-embedded dimension. It can effectively decrease the Lyapunov exponential sum in added dimensions of reconstruction set when the dimensions of reconstructed space are increased. Finally, the forecasting result is reconstructed based on wavelet theory. The forecasting result of original runoff time series is achieved. The method is high precision and feasible through example test.

Key Words:- Wavelet; Runoff; Chaos; Forecasting

1 Introduction

The nonlinear theories of runoff forecasting, such as the applications of neural network and chaos, have recently made considerable progress [1]-[4]. However, it is still necessary to study how to explore the characteristic of runoff under multi-scaling so that the more accurate model can be established. The wavelet transform is a kind of signal time-frequency analysis method. It has the characteristic of multi-resolution analysis. It can manifest the local characteristic of time domain and frequency domain. It is also a sort of localized time-frequency analysis method that both of time frame and frequency frame can be changed while the size of frame is fixed. By means of the wavelet transform, it is very effective to achieve the adjusting regulation of complicated time series and to resolve evolvement characteristic of time series at different time scaling. The local domain forecasting method of chaos is a very useful method, which is well developed in recent years [5][6]. Unfortunately, it is not so valid to predict

attractors with high-embedded dimensions by using local domain forecasting method.

This paper develops a new method for disposing runoff time series with high-embedded dimension based on wavelet transform theory. Firstly, non-stationary runoff time series is decomposed into stationary time series and stochastic time series. Secondly, different models are employed to forecast them respectively. Especially, by studying chaos characteristic of stochastic time series, a nonlinear chaos dynamics-forecasting model is put forward to dispose runoff time series with high-embedded dimensions. Finally, the forecasting result is reconstructed based on wavelet theory, and then the forecasting result of original runoff time series is achieved. The method is high precision and feasible through example test.

2 Theory of wavelet decomposition and reconstruction

Assuming $\psi(t) \in L^2(R)$, its Fourier transform is $\hat{\psi}(\omega)$. $\psi(t)$ is defined as a basic wavelet or

mother wavelet if the following full reconstruction condition is met:

$$C_\psi = \int_R \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty \quad (1)$$

The following equation (2) is obtained after the basic wavelet $\psi(t)$ is stretched and translated. $\psi_{a,b}(t)$ is named a wavelet series.

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

$a, b \in R; a \neq 0$

Where a and b are indexes for stretch and translation respectively.

If let $a = a_0^j, b = ka_0^j b_0$, then the corresponding discrete wavelet is:

$$\psi_{j,k}(t) = a_0^{-j/2} \psi(a_0^{-j} t - kb_0) \quad (3)$$

It is necessary to change the parameters a and b so that wavelet transform possesses changeable time-resolution and frequency-resolution, which accommodate the non-stationary signal. Let $a = 2^j, b = 2^j k$, then the binary wavelet is obtained:

$$\psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j} t - k) \quad (4)$$

The binary wavelet transform is:

$$\begin{aligned} W_{2^j} f(k) &= \langle f(t), \psi_{2^j}(k) \rangle \\ &= \frac{1}{2^j} \int_R f(t) \psi^*(2^{-j} t - k) dt \end{aligned} \quad (5)$$

The corresponding inverse transform is:

$$f(t) = \sum_{j \in Z} \int W_{2^j} f(x) \psi_{2^j}(2^{-j} t - k) dt \quad (6)$$

The binary wavelet does not destroy the signal translation invariable in time domain because it only has discrete parameter a but continuous parameter b in time domain. The signal can be decomposed into high frequency series and low frequency series because of the different scaling. In this paper, runoff time series is decomposed and reconstructed by using binary wavelet based on the above theory. Mallat algorithm^[7] is adopted to decompose and reconstruct signal. The signal is decomposed step by step. The low frequency signal having been acquired at previous step is decomposed into high frequency series and low frequency series every step. After

the Nth step has been fulfilled, the original signal is decomposed into:

$$X = D_1 + D_2 + \dots + D_N + A_N \quad (7)$$

Where D_1, D_2, \dots, D_N are the high frequency signals that are acquired at the first step and at the second step until the Nth step respectively. A_N is the low frequency signal that is acquired at the Nth step.

If D_1, D_2, \dots, D_N and A_N can be forecasted, the original signal will be able to be forecasted by means of the reconstruction algorithm.

3 The wavelet decomposition and characteristic analysis of runoff time series

The signal can be decomposed into stationary low frequency series and nonlinear high frequency series by using wavelet decomposition algorithm. For the runoff of drainage area in Eons state, the original series is decomposed through five steps. The results are showed in figure 1(a-f). Figure 1(a) shows the curve of the low frequency series, while figure 1(b-f) show the curves of high frequency series.

3.1 The forecasting model of stationary time series

Figure 1(a) shows that the low frequency series has distinct stationary characteristic and accordingly it can be forecasted through AR (n) model.

Assuming one series is x_1, x_2, \dots, x_N , and then the following equation is obtained:

$$x_i = \varphi_1 x_{i-1} + \varphi_2 x_{i-2} + \dots + \varphi_n x_{i-n} + \varepsilon_t \quad (8)$$

Where the parameter set $\varphi = [\varphi_1, \varphi_2, \dots, \varphi_n]$ can be estimated through least-squares algorithm according to the theory of time series.

3.2 The forecasting model of non-stationary time series

Figure 1(b-f) show that the high-frequency series is borne with a very strong nonlinear. It is necessary to study special forecasting models to deal with the high frequency series. We employ the chaos theory to forecast them in this paper. According to the chaos theory, we must decide whether the nonlinear time series has chaos

characteristic. Some methods, such as reconstruction of phase space, Lyapunov exponential spectrum and conjunction saturation exponential, are in common use to decide chaos characteristic of time series.

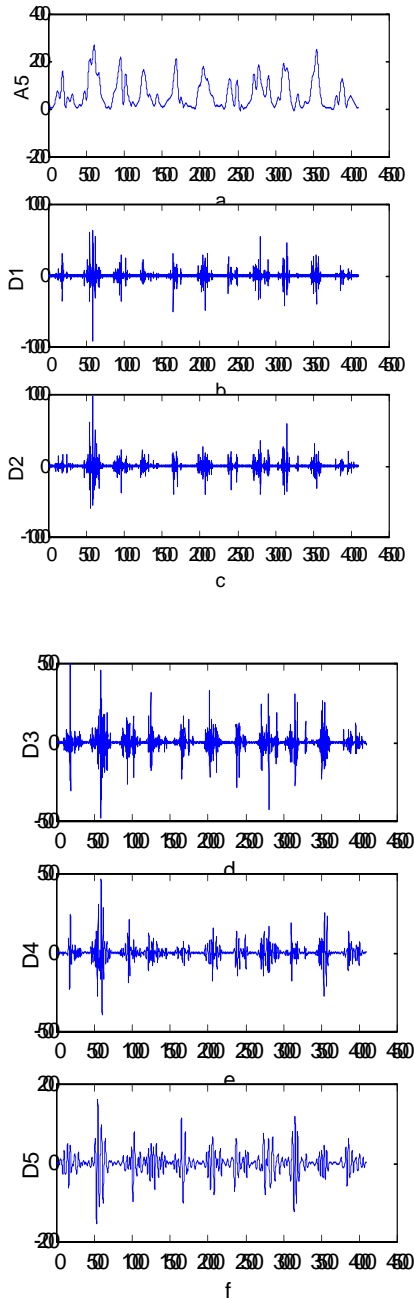


Fig. 1 Low frequency series and high frequency series for five steps

The theory of reconstruction of phase space is put forward by Takens in 1981^[8]. It is described as following:

Assuming an m -dimension self-control dynamics system can be represent with equation (9).

$$\frac{dx_i}{dt} = f_i(x_1, x_2, \dots, x_m) \quad i=1,2,\dots,m \quad (9)$$

Equation (9) is transformed into an m -rank nonlinear differential equation via elimination method.

$$x^{(m)} = f(x, x^{(1)}, \dots, x^{(m-1)}) \quad (10)$$

It corresponds to time evolution:

$$\vec{x}(t) = (x(t), x'(t), \dots, x^{(m-1)}(t)) \quad (11)$$

The above equation can be dispersed into equation (12) with substituting difference equation for differential.

$$\vec{x}(t) = (x(t), x(t+\tau), \dots, x(t+(m-1)\tau)) \quad (12)$$

Equation (12) is named reconstruction phase space of time delay. It retains the geometry structure, topology structure and dynamics characteristic with original system in a differential homoeomorphism sense. The theory is applied to reconstructing the time series of figure 1(f) in two-dimension and three-dimension phase space. The reconstruction diagrams are showed in figure 2 and figure 3 respectively.

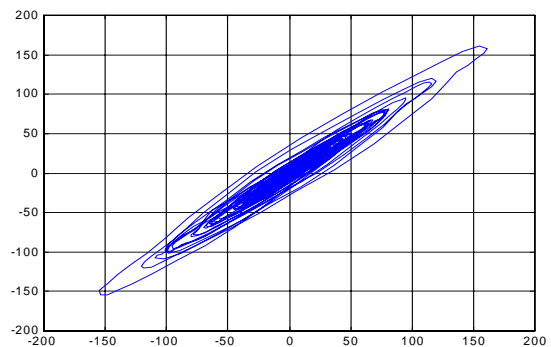


Fig. 2 Two-dimension reconstruction diagram of D5

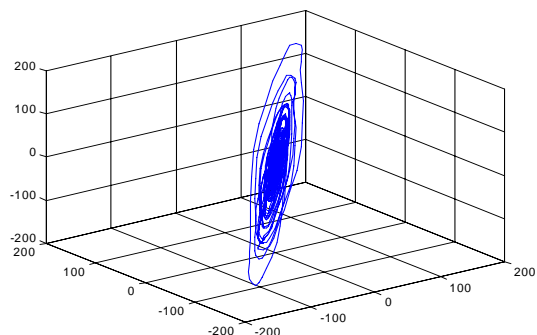


Fig.3 Three-dimension reconstruction diagram of D5

Figure 2 and Figure 3 show that the time series possesses fairly regular attractors structure. The maximum Lyapunov exponential of the time

series is acquired to judge its chaos characteristic by means of the following ways. It is well known that positive Lyapunov exponential means chaos. That is said that the important function of the exponential lies in judging chaos characteristic of the time series. The maximum Lyapunov exponential of the high frequency signal is achieved with the same method adopted by Wolf^[9]. The results are showed in table 1.

Table 1: The maximum Lyapunov exponential of high-frequency signal

Signal	λ_{\max}	Embedded dimensions
D1	0.125278	4
D2	0.382168	4
D3	0.410715	4
D4	1.302778	3
D5	1.632457	3

A chaos-forecasting model can be established according to the chaos characteristic of the time series.

The basic principle of traditional local domain chaos forecasting method is as following.

Firstly, for one time series $\{x_1, x_2, \dots, x_{N-1}\}$, it can be embedded into a D-dimension space by making use of reconstruction technology of phase space, and then one trajectory of it in D-dimension phase space can be constructed with equation (13).

$$\left. \begin{aligned} X_{D-1} &= (x_{(D-1)\tau}, x_{(D-2)\tau}, \dots, x_0) \\ X_D &= (x_{(D-1)\tau+1}, x_{(D-2)\tau+1}, \dots, x_1) \\ &\vdots \\ X_{N-1-(D-1)} &= (x_{(D-1)\tau+N-1}, x_{(D-2)\tau+N-1}, \dots, x_{N-1}) \end{aligned} \right\} \quad (13)$$

Secondly, assuming x_n is a known number, it is necessary to forecast $x_{n+\tau}$ (without loss of generality, let $\tau=1$). The latest trajectory point containing information of x_n is:

$$X_n = (x_n, x_{n-1}, \dots, x_{n-D+1}) \quad (14)$$

The next trajectory point that needs to be located in D-dimension space is

$\hat{X}_{n+1} = (\hat{x}_{n+1}, \hat{x}_n, \dots, \hat{x}_{n-D+2})$. Where the new information \hat{x}_{n+1} in \hat{X}_{n+1} can be regarded as one forecasting result of x_{n+1} . As a result, a mapping F_n will be created as a predictor to achieve: $\hat{X}_{n+1} = F_n(X_n)$.

The mapping F_n can be achieved through the following steps:

(1) According to equation (15), k of the nearest points among $N-D+1-(D-1)\tau$ points in D-dimension away from $X_n = (x_n, x_{n-1}, \dots, x_{n-D+1})$ can be found out.

$$|X_n - X_n^0| = \left[\sum_{i=1}^D (x_{n-D+i} - x_{n-D+i}^0)^2 \right]^{\frac{1}{2}} \quad (15)$$

Where equation (15) is the distance equation in D-dimension space.

(2) According to the iterative regularity of these k points, one polynomial P_n can be employed to fit F_n .

If the polynomial P_n has been ascertained, we can employ equation (16) to achieve these coefficients of P_n . Consequently, X_{n+1} can be obtained: $\hat{X}_{n+1} = F_n(X_n)$

$$\begin{aligned} & \min_{1 \leq j \leq k} (F_n(X_{nj}) - P_n(X_{nj}))^2 \\ & = \min_{1 \leq j \leq k} (X_{nj+1} - P_n(X_{nj}))^2 \end{aligned} \quad (16)$$

The results are preferable by adopting the algorithm mentioned above when we forecast the attractors with low-embedded dimensions. However, it is not satisfied when we forecast the attractors with high-embedded dimensions^{[10]-[13]}. The main reason is that the higher of embedded dimension, the longer history of information will be needed. In the dynamics system with positive Lyapunov exponential spectrum, the longer the history is, the less effective of history information towards chaos time series is. The effect is exponentially decreased with time increasing. The large Lyapunov exponential in newly added dimensions of reconstruction set will make the arithmetic sum of full Lyapunov exponential spectrum increase, while the dimensions of reconstruction space increase. In order to reduce the arithmetic sum of full Lyapunov exponential spectrum in reconstruction set, equation (15) can be changed into:

$$|X_n - X_n^0| = \left[\sum_{i=1}^D \left(\frac{1}{2} \right)^{D-i+1} (x_{n-D+i} - x_{n-D+i}^0)^2 \right]^{\frac{1}{2}} \quad (17)$$

3.3 The test result

Firstly, we forecast low frequency series and high frequency series by means of AR(n) model and chaos model. Secondly, we construct them to achieve the forecast result of the original time series based on the wavelet theory. Runoff of Eons is forecasted in this paper and the result is shown in figure 4. Where the continuous curve and the dashed curve represent actual data and forecasting data respectively.

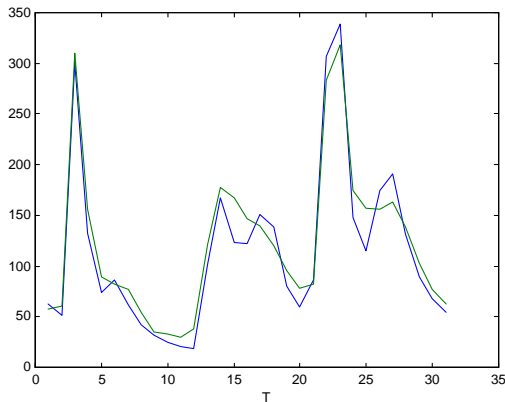


Fig. 4 The result of forecasting

4 Conclusion

It is difficult to find an effective model to forecast strong nonlinear runoff. In this paper, we decompose the time series into stationary time series and stochastic time series through considering different scaling based on wavelet transform. Stationary time series and stochastic time series are forecasted by means of different models respectively. The wavelet transform and chaos analysis are considered at the same time in the model of non-stationary series. The result shows the method is powerful.

References:

- [1] Lin Shuzhen, Liu Changling, You Baoshan. The Chaos Dynamics Research of Daily Runoff Time Series. *Taiwan Hydraulic*. 44(2), 1996:13-23
- [2] I.Rodriguez-Iturbe & B.F de Power, Chaos in Rainfall. *Water Resources Research*, 25(7), 1989:1667-1675
- [3] Jiang Chuanwen, Quan Xianzhang, Zhang Yongchuan. A Research on the Chaos Dynamic Model of Runoff. *Journal of Hydrodynamics*. Ver B12(2): 2000,85-91
- [4] Jiang Chuanwen, Li Shuai, Du Songhuai. Forecasting Runoff With Higher-Embedded Dimensions Using DMEP-Based Artificial Neural Networks. *Artificial Intelligence Applications and Innovations*. 486-497,2005
- [5] Martin Casdagli. Nonlinear Prediction of Chaotic Time Series. *Physica*, D35,1989
- [6] Jiang, Chuanwen, Li Tao. Forecasting Method Study on Chaotic Load Series with High Embedded Dimension. *Energy Conversion and Management*. 46(5),pp667-676, 2005
- [7] Mallat S,Zhang S. Characterization of Signals from Multiscale Edges. *IEEE. Trans. on pattern analysis and machine intelligence*,1992,14(7):710-732
- [8] Takens.F. Detecting Strange attractors in turbulence, in: *Dynamical Systems and Turbulence*,Warwick 1980,Lecture Notes in mathematics, No.898,edited by D.A.Rand and L.-S.Yong, Springer-Verlay, Berlin, 1981:366-381
- [9] Wolf A.etal. Determining Lyapunov Exponents From A Time Series. *Physica D*.16,1985(285-317)
- [10] Ma Shexiang, Liu Guizhong, Zeng Zhaohua. The Analysis and Forecasting of Non-stationary Time Series Based on Wavelet Analysis. *Journal of Systems Engineering*. 2000,(4)305-311
- [11] Jiang Chuanwen, Quan xianzhang, Zhang Yongchuan. Application of Chaos Theory in Power System Load Forecasting. *Journal of Wuhan Transportation University*. 23(6):1999,608-611.
- [12] Jiang Chuanwen, Hou Zhijian etal.An Improved Grey Procedure for Forecasting the Runoff. *Journal of Grey System*,UK,13(2):2001,179-182.
- [13] Francesco Lisi, Vigilio Villi. Chaotic Forecasting of Discharge Time Series: A Case Study. *JAWRA*, 37(2):2001.