

# Chapter 8

## The Nonlinear Hydro-climatic Process: A Case Study of the Tarim Headwaters, NW China

Jianhua Xu, Yaning Chen and Weihong Li

**Abstract** Some researchers have suggested that the hydro-climatic process is a complex system, with nonlinearity as its basic characteristic. But there is still a lack of effective means available to thoroughly discover the dynamics of hydro-climatic process at different time scales. Therefore, more studies are required to explore the nonlinearity of hydro-climatic process from different perspectives and using different methods. Based on the hydrologic and meteorological data in the areas of the Tarim headwaters, this chapter investigated the nonlinear hydro-climatic process by a comprehensive method including correlation dimension, R/S analysis, wavelet analysis, regression and artificial neural network modeling. The main findings are as follows:

- (1) The hydro-climatic process in the Tarim headwaters presented periodic, nonlinear, chaotic dynamics, and long-memory characteristics.
- (2) The correlation dimensions of the attractor derived from the AR time series for the Hotan, Yarkand, Aksu and Kaidu rivers were all greater than 3.0 and non-integral, implying that all four headwaters are dynamic chaotic systems that are sensitive to initial conditions, and that the dynamic modeling of hydro-climatic process requires at least four independent variables.
- (3) The Hurst exponents indicate that a long-term memory characteristic exists in hydro-climatic process. However, there were some differences observed, with

---

J. Xu (✉)

The Research Center for East-West Cooperation in China, The Key Lab of GIScience of the Education Ministry PRC, East China Normal University, 200241 Shanghai, China  
e-mail: jhxu@geo.ecnu.edu.cn

Y. Chen · W. Li

State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, 830011 Urumqi, Xinjiang, China  
e-mail: chenyn@ms.xjb.ac.cn

W. Li

e-mail: liwh@ms.xjb.ac.cn

Y. Chen (ed.), *Water Resources Research in Northwest China*,

DOI 10.1007/978-94-017-8017-9\_8, © Springer Science+Business Media Dordrecht 2014

the Aksu, Yarkand and Kaidu rivers demonstrating a persistent trait, and the Hotan River exhibiting an anti-persistent feature.

- (4) The variation pattern of runoff, temperature and precipitation was scale-dependent with time. Annual runoff (AR), annual average temperature (AAT) and annual precipitation (AP) at five time scales resulted in five variation patterns respectively.
- (5) The nonlinear variation of runoff is resulted from regional climatic change. The variation periodicity of AR is close with that of AAT and AP. The multiple linear regression (MLR) and back-propagation artificial neural network (BPANN) based on wavelet analysis reveal the correlations between annual runoff (AR) with annual precipitation (AP), annual average temperature (AAT) at different time scales.

**Keywords** Hydro-climatic process · Nonlinearity · Periodic · Nonlinear and chaotic dynamics · Tarim headwaters

## 8.1 Introduction

Theoretically, hydro-climatic process can be evaluated to determine if they comprise an ordered, deterministic system, an unordered, random system, or a chaotic, dynamic system, and whether change patterns of periodicity or quasi-periodicity exist. However, it is difficult to achieve a thorough understanding of the nonlinear hydro-climatic process (Cannon and McKendry 2002; Xu et al. 2010).

In the last 20 years, many studies have been conducted to evaluate climatic change and hydrological processes in the arid and semi-arid regions in northwestern China (Chen and Xu 2005; Wang et al. 2010; Xu et al. 2011a, b; Zhang et al. 2010). A number of studies have indicated that there was a visible transition in the hydro-climatic processes in the past half-century (Chen and Xu 2005; Chen et al. 2006; Shi et al. 2007; Wang et al. 2010). This transition was characterized by a continual increase in temperature and precipitation, added river runoff volumes, increased lake water surface elevation and area, and elevated groundwater levels. This transition may inquiries a series questions if these changes represent a localized transition to a warm and wet climate type in response to global warming, or merely reflect a centennial periodicity in hydrological dynamics. To date, these questions have not received satisfactory answers; therefore, more studies are required to explore the nonlinear characteristics of hydro-climatic process from different perspectives and using different methods (Xu et al. 2012).

Because of the above reasons, this chapter investigated the nonlinear hydro-climatic process in the Tarim headwaters by an integrated approach including correlation dimension, R/S analysis, wavelet decomposition, regression analysis and artificial neural network.

## 8.1.1 Materials and Methods

### 8.1.1.1 Study Area

The Tarim River basin covers an area of  $1.02 \times 10^6$  km<sup>2</sup> and is the largest continental river basin in China. The basin covers the entire southern part of Xinjiang in western China and is characterized by an abundance of rich natural resources and a fragile environment. This region has an extreme desert climate with an annual average temperature of 10.6 ~ 11.5 °C. In addition, the monthly mean temperature ranges from 20 to 30 °C in July and -10 to -20 °C in January and the highest and lowest temperatures are +43.6 °C and -27.5 °C, respectively. The accumulative temperature > 10 °C ranges from 4,100 to 4,300 °C. The average annual precipitation is approximately 116.8 mm for the entire area, with an uneven distribution of 200–500 mm in the mountainous area, 50 ~ 80 mm on the edges of the basin, and only 17.4–25.0 mm in the central portion of the basin. There is great temporal unevenness in precipitation within each year as well. More than 80% of the total annual precipitation falls between May and September in the high-flow season, and less than 20% of the total precipitation occurs between November and April.

The main channel of the Tarim River is 1,321 km in length. Naturally and historically the Tarim River basin consists of 114 rivers from nine drainage systems, which include the Aksu, Hotan, Yarkand, Qarqan, Keriya, Dina, Kaxgar, and Kaidu-Konqi rivers. The basin contains  $20.44 \times 10^6$  ha of arable lands and has a human population of  $8.26 \times 10^6$ . The mean annual natural surface runoff is  $3.98 \times 10^{10}$  m<sup>3</sup>, most of which originates from glaciers, snowmelt and precipitation in the surrounding mountains.

Intensive disturbances caused by human activities, particularly in response to excessive water resources exploitation, have brought about marked changes during the past 50 years. The drainage systems gradually disintegrated when the Weigan, Kaxgar, Dina, Keriya, and Qarqan rivers stopped flowing into the mainstream and were eventually disconnected from it. Today, there are only three drainage systems connected to the mainstream of the Tarim River, the Aksu River, Yarkand River and Hotan River. The Aksu River has two main tributaries (the Tongshigan and Kumalak) and originates from the Tianshan Mountains in the northwest portion of the basin. The Hotan River also has two main tributaries (the Kalaksh and Yulongkash) and originates in the Kunlun Mountains and flows through the southwestern portion of the basin. The Yarkand River originates from the Pamir Plateau and lies between the Aksu River and the Yarkand River (Fig. 8.1).

As mentioned above, glaciers, snowmelt and precipitation in the surrounding mountains are the major sources of runoff in the Tarim River. Specifically, glacier melt and snowmelt comprise 48.2% of the total runoff of the river. Interannual runoff variability is small, with a coefficient of variation ranging from 0.15 to 0.25 and maximum and minimum modular coefficients of 1.36 and 0.79, respectively. Additionally, seasonal runoff is unevenly distributed, with runoff during the flood

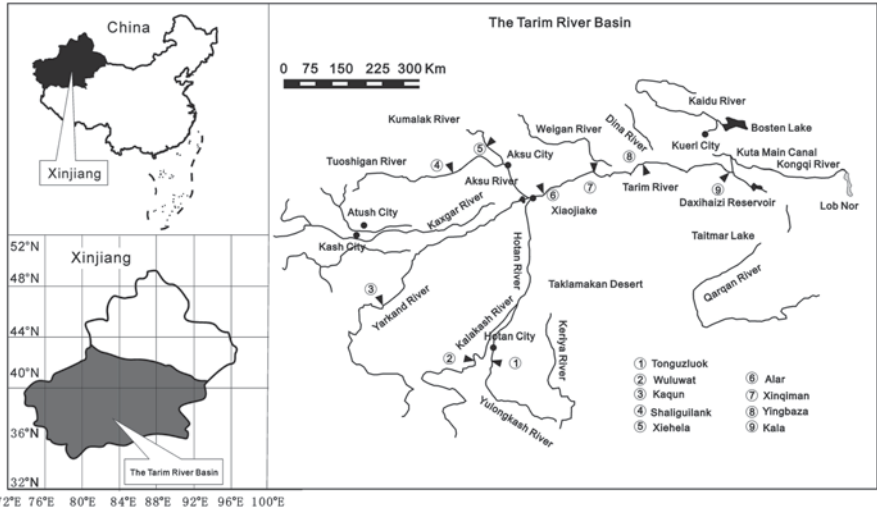


Fig. 8.1 Location of study area

season (June–August) accounting for 60–80% of the total annual runoff (Chen and Xu 2005).

**8.1.1.2 Data**

To investigate hydro-climatic process of the Tarim headwaters, this study used the time series of annual runoff (AR) in the period of 1957~2008 from Xiehela and Shaliguilank hydrologic stations for the Aksu River, from the Kaqun hydrologic station for the Yarkand River, from the Tongzuluok and Wuluwat hydrologic stations for the Hotan River, and from the Dashankou hydrologic station for the Kaidu River. We also used annual average temperature (AAT) and annual precipitation (AP) from 28 meteorological stations in the Tarim River Basin for the same study period.

**8.1.1.3 Methodology**

(1) Correlation dimension

The correlation dimension method is usually applied to analyze a series and determine if the process exhibits a chaotic dynamic characteristic (Sivakumar 2007; Xu et al. 2009a). Consider  $x(t)$ , the time series of annual runoff, and suppose it is generated by a nonlinear dynamic system with  $m$  degrees of freedom. To restore the dynamic characteristic of the original system, it is necessary to construct an appropriate series of state vectors,  $X^{(m)}(t)$ , with delay coordinates in the  $m$ -dimensional

phase space according to the basic ideas initiated by Grassberger and Procaccia (1983):

$$X^{(m)}(t) = \{x(t), x(t + \tau), \dots, x(t + (m - 1)\tau)\} \tag{8.1}$$

where  $m$  is the embedding dimension and  $\tau$  is an appropriate time delay.

The trajectory in the phase space is defined as a sequence of  $m$  dimensional vectors. If the dynamics of the system can be reduced to a set of deterministic laws, the trajectories of the system converge toward a subset of the phase space, which is called an ‘‘attractor’’. Many natural systems do not conform with time to a cyclic trajectory. Some nonlinear dissipative dynamic systems tend to shift toward the attractors for which the motion is chaotic, i.e. not periodic and unpredictable over long times. The attractors of such systems are called strange attractors. For the set of points on the attractor, using the G-P method (Grassberger and Procaccia 1983), the correlation-integrals are defined to distinguish between stochastic and chaotic behaviors.

The correlation-integrals can be defined as follows:

$$C(r) = \frac{1}{N_R^2} \sum_{j=1}^{N_R} \sum_{i=1}^{N_R} \Theta(r - |X_i - X_j|) \tag{8.2}$$

where  $r$  is the surveyor’s rod for distance,  $N_R$  is the number of reference points taken from  $N$ , and  $N$  is the number of points,  $X^{(m)}(t)$ . The relationship between  $N$  and  $N_R$  is  $N_R = N - (m - 1)\tau$ .  $\Theta(x)$  is the Heaviside function, which is defined as:

$$\Theta(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases} \tag{8.3}$$

The expression counts the number of points in the dataset that are closer than the radius,  $r$ , within a hypersphere of the radius,  $r$ , and then divides this value by the square of the total number of points (because of normalization). As  $r \rightarrow 0$ , the correlation exponent,  $d$ , is defined as:

$$C(r) \propto r^d \tag{8.4}$$

It is apparent that the correlation exponent,  $d$ , is given by the slope coefficient of  $\ln C(r)$  versus  $\ln r$ . According to  $(\ln r, \ln C(r))$ ,  $d$  can be obtained by the least squares method (LSM) using a log-log grid.

To detect the chaotic behavior of the system, the correlation exponent has to be plotted as a function of the embedding dimension (as shown as Fig. 3.2 in Sect. 3.1). If the system is purely random (e.g. white noise) the correlation exponent increases as the embedding dimension increases, without reaching the saturation value.

If there are deterministic dynamics in the system, the correlation exponent reaches the saturation value, which means that it remains approximately constant as the embedding dimension increases. The saturated correlation exponent is called the

correlation dimension of the attractor. The correlation dimension belongs to the invariants of the motion on the attractor. It is generally assumed that the correlation dimension equals the number of degrees of freedom of the system, and higher embedding dimensions are therefore redundant. For example, to describe the position of the point on the plane (two-dimensional system), the third dimension is not necessary because it is redundant. In addition, the correlation dimension is often fractal and represented as a non-integral dimension, which is typical for chaotic dynamical systems that are very sensitive to initial conditions.

(2) R/S analysis

R/S analysis, which is also called rescaled range analysis, is usually applied to analyze long-term correlation characteristics of a time series (Xu et al. 2008a; Li et al. 2008). The principle of R/S analysis is briefly introduced as follows (Mandelbrot and Wallis 1969; Turcotte 1997):

Considering a time series  $x(t)$ , such as the annual runoff sequence of a certain river, for any positive integer  $\tau \geq 1$ , the mean value series is defined as:

$$\langle x \rangle_\tau = \frac{1}{\tau} \sum_{t=1}^{\tau} x(t) \quad \tau = 1, 2, \dots \tag{8.5}$$

The accumulative deviation is:

$$X(t, \tau) = \sum_{u=1}^t (x(u) - \langle x \rangle_\tau) \quad 1 \leq t \leq \tau \tag{8.6}$$

The extreme deviation is:

$$R(\tau) = \max_{1 \leq t \leq \tau} X(t, \tau) - \min_{1 \leq t \leq \tau} X(t, \tau) \quad \tau = 1, 2, \dots \tag{8.7}$$

The standard deviation is:

$$S(\tau) = \left[ \frac{1}{\tau} \sum_{t=1}^{\tau} (x(t) - \langle x \rangle_\tau)^2 \right]^{\frac{1}{2}} \quad \tau = 1, 2, \dots \tag{8.8}$$

When analyzing the statistic rule of  $R(\tau)/S(\tau) \triangleq R/S$ , H E Hurst discovered a relational expression,

$$R/S \propto \left( \frac{\tau}{2} \right)^H \tag{8.9}$$

which can be used to identify the Hurst phenomenon in the time series, where  $H$  is known as the Hurst exponent. It is evident that  $H$  is given by the slope coefficient of  $R/S$  versus  $\tau/2$ . According to  $(\ln \tau/2, \ln (R/S))$ ,  $H$  can be obtained by the least squares method (LSM) in a log-log grid.

Hurst et al. (1965) once demonstrated that if  $x(t)$  is an independently random series with limited variance, the exponent,  $H=0.5$ , and  $H$  ( $0 < H < 1$ ) is dependent on a correlation function  $C(t)$ :

$$C(t) = 2^{2H-1} - 1 \tag{8.10}$$

When  $H > 0.5$  and  $C(t) > 0$ , the process has a long, enduring characteristic, and the future trend of the time series will be consistent with the past. In other words, if the past showed an increasing trend, the future will also show an increasing trend. When  $H < 0.5$  and  $C(t) < 0$ , the process has an anti-persistence characteristic, and the future trend of the time series will be opposite from the past. In other words, if the past showed an increasing trend, the future will assume the reducing trend. When  $H = 0.5$  and  $C(t) = 0$ , the process is stochastic; in other words, there is no correlation or only a short-range correlation in the process (Ai and Li 1993).

(3) Wavelet analysis

Wavelet transformation has been shown to be a powerful technique for characterization of the frequency, intensity, time position, and duration of variations in climate and hydrological time series (Torrence and Compo 1998; Smith et al. 1998; Chou 2007; Xu et al. 2009b). Wavelet analysis can also reveal the localized time and frequency information without requiring the time series to be stationary, as required by the Fourier transform and other spectral methods.

A continuous wavelet function  $\Psi(\eta)$  that depends on a nondimensional time parameter  $\eta$  can be written as (Labat 2005):

$$\Psi(\eta) = \Psi(a, b) = |a|^{-1/2} \Psi\left(\frac{t-b}{a}\right) \tag{8.11}$$

where,  $t$  denotes time,  $a$  is the scale parameter and  $b$  is the translation parameter.  $\Psi(\eta)$  must have a zero mean and be localized in both time and Fourier space (Farge 1992). The continuous wavelet transform (CWT) of a discrete signal,  $x(t)$ , such as the time series of runoff, temperature, or precipitation, is expressed by the convolution of  $x(t)$  with a scaled and translated  $\Psi(\eta)$ ,

$$W_x(a, b) = |a|^{-1/2} \int_{-\infty}^{+\infty} x(t) \Psi^*\left(\frac{t-b}{a}\right) dt \tag{8.12}$$

where, \* indicates the complex conjugate, and  $W_x(a, b)$  denotes the wavelet coefficient. Thus, the concept of frequency is replaced by that of scale, which can characterize the variation in the signal,  $x(t)$ , at a given time scale.

Selecting a proper wavelet function is a prerequisite for time series analysis. The actual criteria for wavelet selection include self-similarity, compactness, and smoothness (Ramsey 1999). For the present study, symlet 8 was chosen as the wavelet function according to these criteria.

The nonlinear trend of a time series,  $x(t)$ , can be analyzed at multiple scales through wavelet decomposition on the basis of the discrete wavelet transform (DWT). The DWT is defined taking discrete values of  $a$  and  $b$  (Banakar and Azeem 2008). The full DWT for signal,  $x(t)$ , can be represented as (Mallat 1989):

$$x(t) = \sum_k \mu_{j_0,k} \phi_{j_0,k}(t) + \sum_{j=1}^{j_0} \sum_k \omega_{j,k} \psi_{j,k}(t) \tag{8.13}$$

where  $\phi_{j_0,k}(t)$  and  $\psi_{j,k}(t)$  are the flexing and parallel shift of the basic scaling function,  $\phi(t)$ , and the mother wavelet function,  $\psi(t)$ , and  $\mu_{j_0,k}$  ( $j < j_0$ ) and  $\omega_{j,k}$  are the scaling coefficients and the wavelet coefficients, respectively. Generally, scales and positions are based on powers of 2, which is the dyadic DWT (Sun et al. 2006).

Once a mother wavelet is selected, the wavelet transform can be used to decompose a signal according to scale, allowing separation of the fine-scale behavior (detail) from the large-scale behavior (approximation) of the signal (Bruce et al. 2002). The relationship between scale and signal behavior is designated as follows: low scale corresponds to compressed wavelet as well as rapidly changing details, namely high frequency; whereas high scale corresponds to stretched wavelet and slowly changing coarse features, namely low frequency. Signal decomposition is typically conducted in an iterative fashion using a series of scales such as  $a=2, 4, 8, \dots, 2^L$ , with successive approximations being split in turn so that one signal is broken down into many lower resolution components.

(4) Regression analysis based on wavelet decomposition

For understanding the relationship between annual runoff with its related climatic factors at different time scales, we employed Multiple linear regression (MLR) based on wavelet decomposition. This method fits multiple linear regression equation (MLRE) between AR with AAT and AP by using multiple linear regression (MLR) based on the results of wavelet approximation (Xu et al. 2008b).

The multiple linear regression model is

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_kx_k \tag{8.14}$$

where,  $y$  is dependent variable,  $x_i$  the independent variables;  $a_i$  is the regression coefficient, which is generally calculated by method of least squares (Xu 2002). In this study, the dependent variable is the annual runoff (AR) and the independent variables are related climatic factors, such as the annual average temperature (AAT) and annual precipitation (AP), etc.

(5) BPANN based on wavelet decomposition

In order to disclose the relationship between the annual runoff with its related climatic factors at different time scales, this study employed the back propagation artificial neural network (BPANN) based on the results of wavelet decomposition. We first approximated the variation patterns of runoff and its related climate factors, such as AR, ATT and AP using wavelet decomposition on the basis of the discrete wavelet transform (DWT) at different time scales. Then the relationship between AR with AAT and AP were revealed by using BPANN based on the wavelet approximation.



In back-propagation networks, a number of smaller processing elements (PEs) are arranged in layers: an input layer, one or more hidden layers, and an output layer (Hsu et al. 1995). The input from each PE in the previous layer ( $x_i$ ) is multiplied by a connection weight ( $w_{ji}$ ). These connection weights are adjustable and may be likened to the coefficients in statistical models. At each PE, the weighted input signals are summed and a threshold value ( $\theta_j$ ) is added. This combined input ( $I_j$ ) is then passed through a non-linear transfer function ( $f(\cdot)$ ) to produce the output of the PE ( $y_j$ ). The output of one PE provides the input to the PEs in the next layer. This process can be summarized in equations as follows (Maier and Dandy 1998):

$$I_j = \sum w_{ji}x_i + \theta_j \quad (8.15)$$

$$y_i = f(I_j) \quad (8.16)$$

Our ANN model is a three-tier structure: an input X with two variables (i.e. AAT and AP) is linearly mapped to intermediate variables (called hidden neurons) H, which are then nonlinearly mapped to the output y (i.e. AR).

By comparing the advantages and disadvantages of artificial neural network transfer functions (Dorofki et al. 2012), we selected the activation function as hyperbolic tangent sigmoid transfer function as follows:

$$f(I) = \frac{(1 - e^{-I})}{(1 + e^{-I})} \quad (8.17)$$

where  $f(I)$  represents transfer function, and  $I$  represents input.

The purpose of the model is to capture the relationship between a historical set of model inputs and corresponding outputs. As mentioned above, this is achieved by repeatedly presenting examples of the input/output relationship to the model and adjusting the model coefficients (i.e. the connection weights) in an attempt to minimize an error function between the historical outputs and the outputs predicted by the model. This calibration process is generally referred to as ‘training’. The aim of the training procedure is to adjust the connection weights until the global minimum in the error surface has been reached.

The back-propagation process is commenced by presenting the first example of the desired relationship to the network. The input signal flows through the network, producing an output signal, which is a function of the values of the connection weights, the transfer function and the network geometry. The output signal produced is then compared with the desired (historical) output signal with the aid of an error (cost) function.

The model parameters are optimized by minimizing the mean square error given by the cost function:

$$E = \langle \|y - y^{obs}\|^2 \rangle \quad (8.18)$$

Where  $y^{obs}$  is the observed data,  $\langle \cdot \rangle$  denotes a sample or time mean.

Because it can train any network as long as its weight, net input, and transfer functions have derivative functions (Kermani et al. 2005), we selected Levenberg–Marquardt (trainlm) as the training function in the computing environment of MATLAB.

#### (6) Coefficient of determination and Akaike information criterion

In order to identify the uncertainty of the estimated model for a given timescale, the coefficient of determination was calculated as follows:

$$CD = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (8.19)$$

where CD is the coefficient of determination;  $\hat{y}_i$  and  $y_i$  are the simulate value and actual data of runoff respectively;  $\bar{y}$  is the mean of  $y_i (i=1, 2, \dots, n)$ ;  $RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2$  is the residual sum of squares;  $TSS = \sum_{i=1}^n (y_i - \bar{y})^2$  is the total sum of squares. The coefficient of determination is a measure of how well the simulate results represent the actual data. A bigger coefficient of determination indicates a higher certainty and lower uncertainty of the estimates (Xu 2002).

To compare the relative goodness between the ANN and multiple linear regression (MLR) fit for a given timescale, we also used the measure of Akaike information criterion (AIC) (Anderson et al. 2000). The formula of AIC is as follows:

$$AIC = 2k + n \ln(RSS / n) \quad (8.20)$$

where  $k$  is the number of parameters estimated in the model;  $n$  is the number of samples;  $RSS$  is the same as in formula (15). A smaller AIC indicates a better model.

For small sample sizes (i.e.,  $n/k \leq 40$ ), the second-order Akaike Information Criterion ( $AIC_c$ ) should be used instead

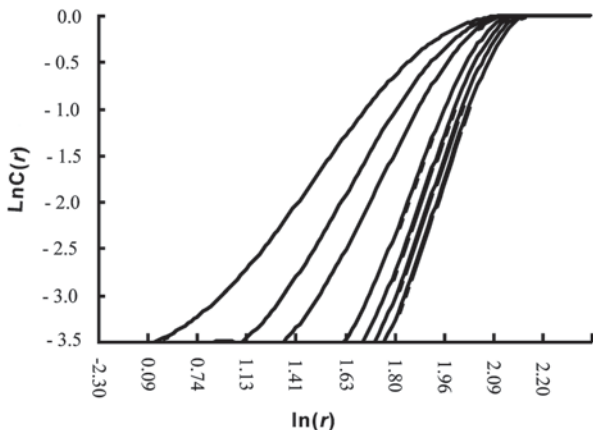
$$AIC_c = AIC + \frac{2k(k+1)}{n-k-1} \quad (8.21)$$

where  $n$  is the sample size. As the sample size increases, the last term of the  $AIC_c$  approaches zero, and the  $AIC_c$  tends to yield the same conclusions as the AIC (Burnham and Anderson 2002).

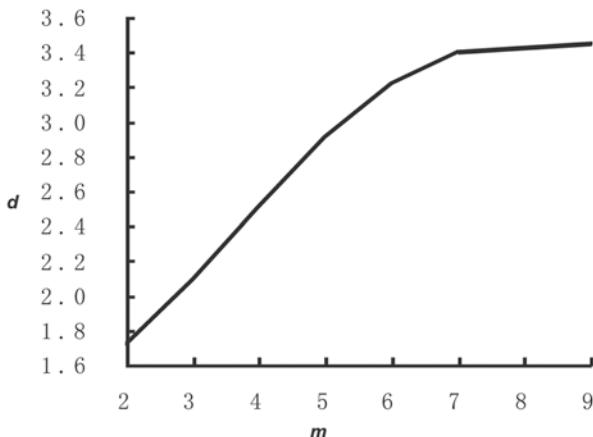
## 8.2 The Chaotic Dynamics of Runoff Process

This study employed the correlation dimension to demonstrate the dynamic characteristics of the runoff time series of four headwaters. The AR time series of Aksu River was first used to reconstruct the phase space, while the correlation dimension

**Fig. 8.2** A plot of  $\ln C(r)$  versus  $\ln(r)$



**Fig. 8.3** The correlation exponent ( $d$ ) versus embedding dimension ( $m$ )



of the attractor was calculated. Different values for the radius,  $r$ , were first selected to compute the values of the correlation-integrals,  $C(r)$ , which were used to plot the curves within a dual logarithmic coordinate system (Fig. 8.2). This diagram shows the relationship between  $\ln C(r)$  and  $\ln C(r)$  for the annual runoff with a number of different embedding dimensions,  $m$ . The slope coefficient of  $\ln C(r)$  versus  $\ln r$ , i.e. the correlation exponent,  $d$ , which was used to embed dimension  $m = 1, 2, \dots$ , was calculated using the least square method (LSM).

The diagram in Fig. 8.3 shows the gradual saturation process of the correlation exponent. It is evident that the correlation exponent increases with embedding dimension,  $m$ , and a saturated correlation exponent, the correlation dimension of attractor ( $D$ ), was obtained when  $m \geq 7$ .

The same procedure was used to derive the correlation dimensions of the attractors for other headwaters, i.e. the Kaidu, Yarkand and the Hotan River. The index values of all four headwaters are shown in Table 8.1.

**Table 8.1** The correlation dimensions for AR in the Tarim headwaters

Rivers	Kaidu River	Aksu River	Yarkand River	Hotan River
Embedding dimension (m)	7	7	7	7
Attractor dimension (D)	3.1104	3.2092	3.2118	3.2227

**Table 8.2** Hurst exponents for AR in the Tarim headwaters

River Name	Kaidu River	Aksu River	Yarkand River	Hotan River
H	0.7897	0.6991	0.5113	0.4367
C(t)	0.4942	0.3178	0.0158	-0.0840

The fact that none of the correlation dimensions is an integer indicates that the annual runoff time series of all four headwaters are chaotic dynamic systems that are very sensitive to the initial conditions. Because the value of the index is above 3 for all headwaters, at least four independent variables are needed to describe the dynamics of the annual runoff process in each river.

### 8.3 The Long-memory Characteristic of Runoff Process

All statistical methods used for time series analysis are based on the assumption that all data from a given time series are independent (i.e. conforming to the Gauss distribution); hence, the series is stochastic. When Hurst (1951) and Hurst et al. (1965) analyzed the water level of the Nile River, he found that time-series variables such as the river water level did not exhibit the stochastic characteristic, but instead showed the characteristic of durability. Based on the empirical findings of Hurst et al., Mandelbrot and Wallis (1968) led to a breakthrough regarding fundamental theories of traditional statistical methods. Specifically, they found that many time series no longer presented a random Brownian movement unrelated to the past, but instead showed a characteristic of long-term correlation (Comte and Renault 1996), which he called “fractal”.

This study employed the rescaled range (R/S) analysis method to characterize the fractal of annual runoff processes in the four headwaters of the Tarim River. Using the R/S analysis method, the Hurst exponent (H) and the correlation function ( $C(t)$ ) were computed for the annual runoff time series of each headwater during the period of 1957~2008 (Table 8.2).

The results shown in Table 8.2 suggest that the annual runoff time series for each of the Tarim’s headwaters possess the characteristic of long memory. The only difference among the headwaters is that the Kaidu, Aksu and Yarkand rivers demonstrate the persistent trait, whereas Hotan shows the anti-persistent feature. The H values for the Kaidu, Aksu and Yarkand River are greater than 0.5 and the C(t) values are greater than 0, which indicates that the future tendency of the annual runoff associated with these systems is consistent with the past runoff.

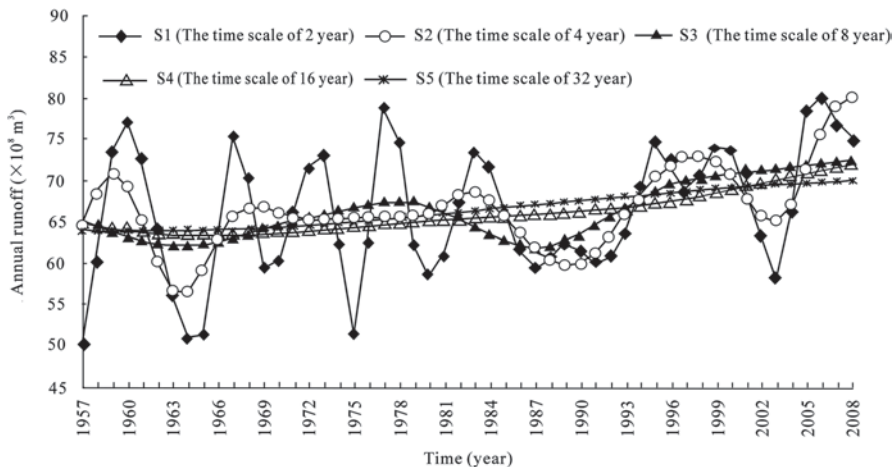


Fig. 8.4 The nonlinear variation patterns for AR in the Yarkand River at the different time scales

However, the H value is less than 0.5 and the C(t) value is less than 0 for the Hotan River, which implies that the future tendency of annual runoff is opposite to that of the past.

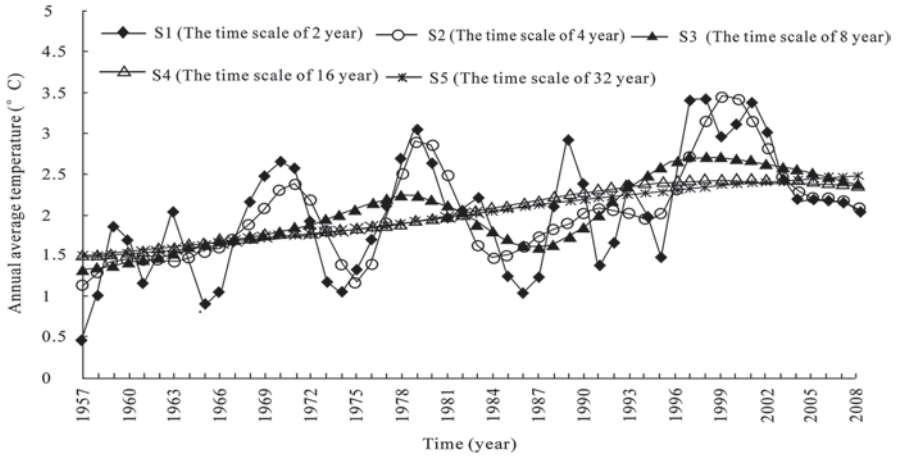
### 8.4 Variation Patterns of AR and its Related Regional Climate Factors

Our previous study indicated that (Xu et al. 2008b), the annual average temperature (AAT) and annual precipitation (AP) are the most important factors that related with the annual runoff (AR). The result was also supported by the other studies for the headwaters of the Tarim River Basin (Hao et al. 2008; Chen et al. 2009; Ling et al. 2012).

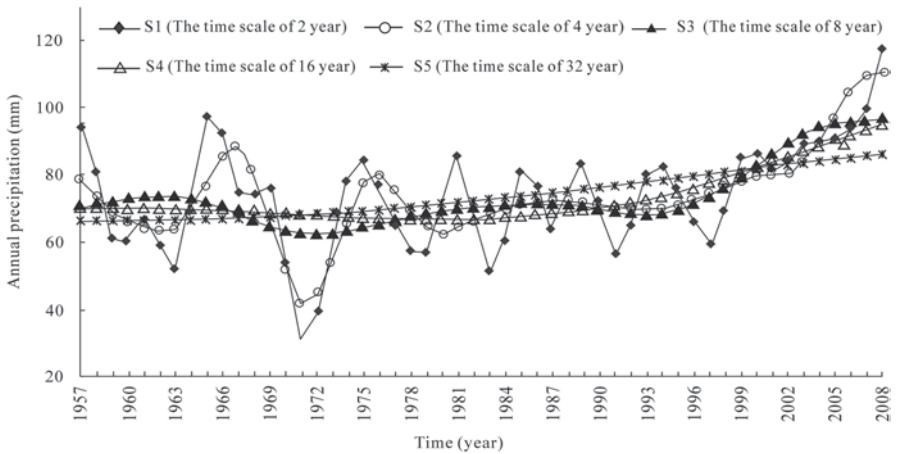
The raw data of AR, AAT and AP showed in fluctuating. It is difficult to identify any patterns simply based on the raw data. In order to show the scale-dependent with time for the hydro-climatic process, the wavelet analysis was used.

The nonlinear variation for AR in the Yarkand River and its related climate factors were analyzed at multiple-year scales through wavelet decomposition on the basis of the discrete wavelet transform (DWT).

The wavelet decomposition for the time series of AR in Yarkand River at five time scales resulted in five variants of nonlinear variations (Fig. 8.4). The S1 curve retains a large amount of residual noise from the raw data, and drastic fluctuations along the entire time span. These characteristics indicate that, although AR varied greatly throughout the study period, there was a hidden increasing trend. The S2 curve still retains a considerable amount of residual noise, as indicated by the



**Fig. 8.5** The nonlinear variation patterns for AAT in the Yarkand River Basin at the different time scales



**Fig. 8.6** The nonlinear variation patterns for AP in the Yarkand River Basin at the different time scales

presence of 4 peaks and 4 valleys. However, the S2 curve is much smoother than the S1 curve, which allows the hidden increasing trend to be more apparent. The S3 curve retained much less residual noise, as indicated by the presence of 2 peaks and 2 valleys. Compared to S2, the increase in runoff over time was more apparent in S3. Finally, the S5 curve presents an ascending tendency, whereas the increasing trend is obvious in the S4 curve.

Accordingly, Figs. 8.5 and 8.6 provide us a comparison for the nonlinear variations of AAT and AP at different time scales. The wavelet decomposition for the time series of AAT and AP in the Yarkand River Basin at five time scales resulted

in five nonlinear variations respectively. These five time scales are also designated as S1 to S5. The curves present an ascending tendency although drastic fluctuations in S1 and S2. Then, the curves are getting much smoother and the increasing trend becomes even more obvious as the scale level increases.

The upper analysis showed that the hydro-climatic process in Yarkand River was dependent on time scales.

Figure 8.7 showed the nonlinear variation patterns of AAT, AP and AR in the Kaidu River at different time scales, which also indicated that the hydro-climatic process was scale-dependent in time.

By using the same approach, we also came to the similar results for the Aksu and Hotan River.

## 8.5 Simulation for Streamflow with Regional Climate Change

### 8.5.1 Simulation by BPANN Based on Wavelet Decomposition

For the purpose of understanding hydro-climatic process, based on the results of wavelet decomposition at different time scales, the back-propagation artificial neural network (BPANN) was employed to simulate the relationship between AR with AAT and AP.

Our study considered a three-layer BPANN, i.e. input layer, output layer and hidden layer, to simulate the nonlinear relationship between AR with AAT, and AP at each time scales. The input layer contains two variables, i.e. AAT and AP, the output layer contains one variable i.e. AR, and the neuron number of hidden layer is 4. Table 8.3 shows the optimized parameters for BPANN in the Kaidu River.

The numerical work was carried out using MATLAB. We selected tansig as the transfer function, and trainlm as the training function to train network.

Based on wavelet decomposition results of AR, AAT and AP from 1957 to 2008, we randomly extracted 80, 10 and 10% of the data as training, validation and testing samples, respectively. The results show that, at the time scales of S1, S2, S3, S4 and S5 (i.e. 2-year, 4-year, 8-year, 16-year, and 32-year), all network models have reached the expected error target (0.001) with learning rate of 0.01. The optimized parameters of the BPANN for the hydro-climatic process at different time scales are showed in Table 8.3.

Table 8.3 reveals that, as the time scale increased from S1 to S5, the estimated error decreases. The average absolute error and relative error for the simulation of AR at time scale of S1 are  $2.3163 \times 10^8 \text{ m}^3$  and 6.65% respectively, but those at time scale of S5 only are  $0.0617 \times 10^8 \text{ m}^3$  and 0.18% respectively.

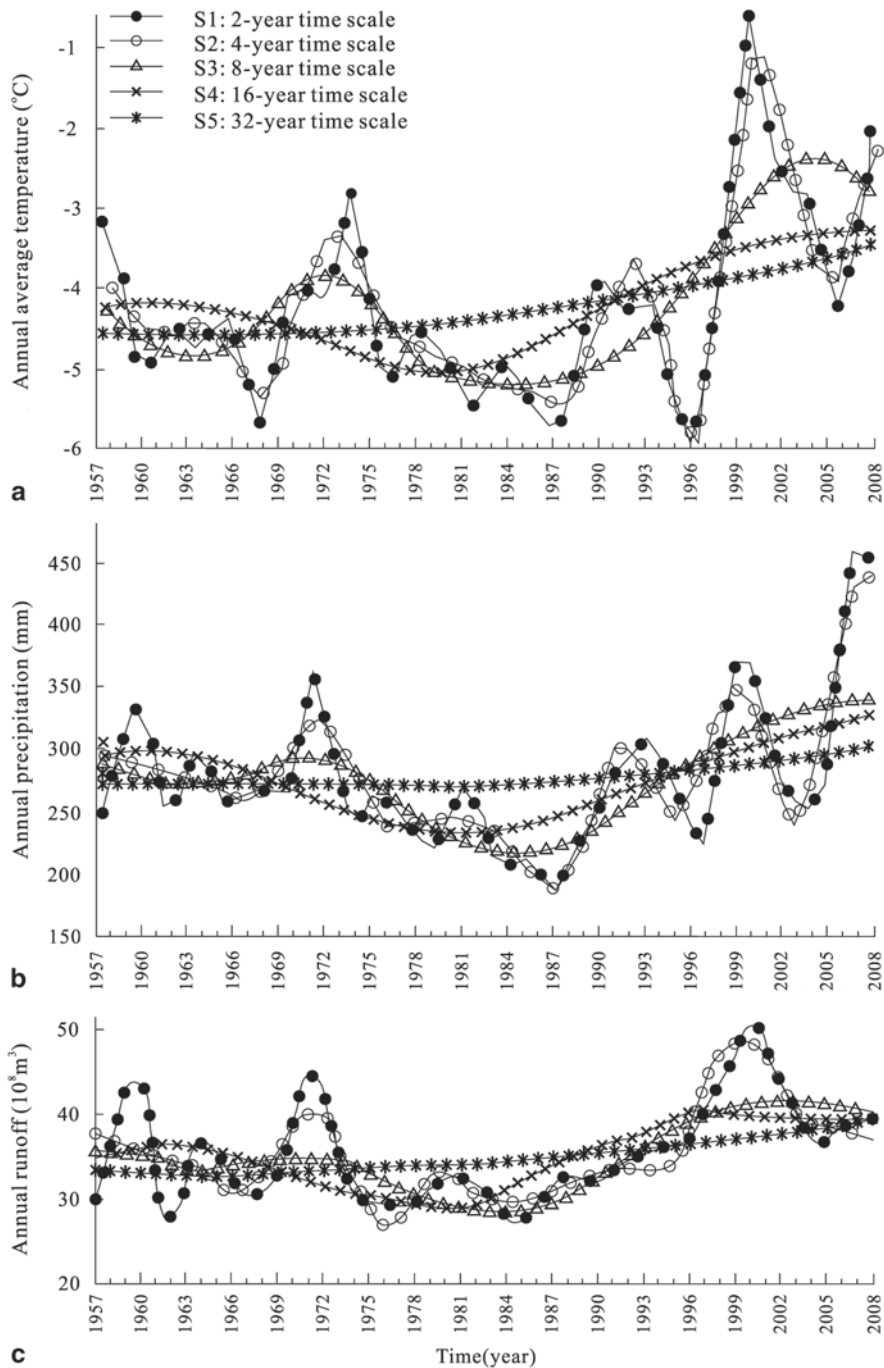


Fig. 8.7 The nonlinear variation patterns of AAT, AP and AR in the Kaidu River at different time scales



**Table 8.3** Basic parameters of the BPANN for hydro-climatic process in the Kaidu River at different time scales

Time scale	Neuron number of the hidden layer	Input variable	Output variable	Transfer function	Train function	Best epoch	Average absolute error	Average relative error (%)
S1	4	AAT, AP	AR	tansig	trainlm	4	2.3163	6.65
S2	4	AAT, AP	AR	tansig	trainlm	5	1.4335	4.25
S3	4	AAT, AP	AR	tansig	trainlm	6	0.9905	2.85
S4	4	AAT, AP	AR	tansig	trainlm	6	0.1938	0.57
S5	4	AAT, AP	AR	tansig	trainlm	10	0.0617	0.18

S1, S2, S3, S4 and S5 represent 2-year, 4-year, 8-year, 16-year and 32-year time scales, respectively. “tansig” is hyperbolic tangent sigmoid transfer function, and “trainlm” is a network training function that updates weight and bias values according to Levenberg–Marquardt optimization  
*AR* annual runoff, *AAT* annual average temperature, *AP* annual precipitation

**Table 8.4** MLREs for hydro-climatic process in the Kaidu River at different time scales

Time scale	Regression equation	R <sup>2</sup>	F	Average absolute error	Average relative error (%)
S1	$AR=2.251AAT+0.036AP+34.420$	0.541	28.932	2.9997	8.41
S2	$AR=4.098AAT-0.002AP+52.921$	0.745	71.584	1.9987	5.85
S3	$AR=2.922AAT+0.048AP+34.263$	0.917	269.877	0.9714	2.69
S4	$AR=5.644AAT+0.010AP+56.036$	0.978	1,066.793	0.4501	1.26
S5	$AR=5.216AAT-0.009AP+59.490$	0.999	30,814.085	0.0779	0.22

Significance level  $\alpha=0.01$ . S1, S2, S3, S4 and S5 represent 2-year, 4-year, 8-year, 16-year and 32-year time scales, respectively  
*AR* annual runoff, *AAT* annual average temperature, and *AP* annual precipitation

### 8.5.2 Simulation by MLR Based on Wavelet Decomposition

To compare the simulated results from BPANN with those from the regression model, a group of multiple linear regression equations (MLREs) were fitted based on the results of wavelet decomposition at different time scales (Table 8.4).

Table 8.4 shows that on the five time scales, AR has positive correlations with both AAT and AP at a high significant level of 0.01. In another words, although the runoff, temperature and precipitation displayed nonlinear variations, the runoff presented a linear correlation with the temperature and precipitation. In addition, both of the statistic F and coefficient of determination R<sup>2</sup> for each regression equation showed an increasing trend with the time scale. This pattern indicated that the impact of AAT and AP on AR is more significant at a larger time scale than at a smaller time scale.

Table 8.4 also shows that, the simulated error of MLR is large at the time scale of S1 and S2 (i.e. 2-year and 4-year scale), moderate at the time scale of S3 (i.e. 8-year scale), and small at the time scale of S4 and S5 (i.e. 4-year and 5-year scale).

### 8.5.3 Comparison Between BPANN and MLR on Wavelet Decomposition

Figure 8.8 reveals the original data of AR in the Kaidu River and the simulated values by MLR and BPANN on wavelet decomposition at different time scales respectively.

Table 8.5 showed the coefficient of determination (i.e. CD) as well as the AIC value for BPANN and MLRE at each time scale.

A higher coefficient of determination and a lower AIC value indicate a better model. Overall, comparing the two modeling methods above, i.e. the multiple linear regression (MLR) and back-propagation artificial neural network (BPANN) based on wavelet analysis, we conclude that the BPANN is better than the MLR at each time scale. In other words, both MLR and BPANN successfully simulated the hydro-climatic process based on wavelet analysis, but the effect from BPANN is better than that from MLR.

By using the same approach, we also came to the similar results for other headwaters of the Tarim River, i.e. the Aksu, Yarkand and Hotan River.

**Summary** Based on the hydrologic and meteorological data in the areas of the Tarim headwaters, this chapter investigated the nonlinear hydro-climatic process by an integrated method including correlation dimension, R/S analysis, wavelet decomposition, regression analysis and artificial neural network modeling. The main findings are as follows:

- [1] The results of this study showed that the hydro-climatic process in the Tarim headwaters presented periodic, nonlinear, chaotic dynamics, and long-memory characteristics.
- [2] The correlation dimensions of the attractor for the time series of annual runoff in the four Tarim headwaters were all greater than 3.0 and non-integral, implying that all four headwaters are dynamic chaotic systems that are sensitive to initial conditions, and that the dynamic modeling requires at least four independent variables.
- [3] The Hurst exponents indicate that a long-term memory characteristic exists in hydrological process in the Tarim headwaters. However, there were some differences among the headwaters. The Kaidu, Aksu, and Yarkand River presented a persistent trait, But the Hotan River exhibited an anti-persistent feature.
- [4] The hydro-climatic process at a large time scale (e.g. 32-year scale) is basically a linear process, but at a small time scale (e.g. 2-year or 4-year scale) it is essentially a nonlinear process with complicated causations. Therefore, the estimated precision is high at a large time scale (e.g. 32-year scale) because the time series of runoff are monotonically related to long-term climatic changes. However, it is difficult to accurately predict a nonlinear hydro-climatic process at a small time scale (e.g. 2-year or 4-year scale).
- [5] Our integrated approach conducted by this study can be used to explore the hydro-climatic process of other inland rivers in northwest China.

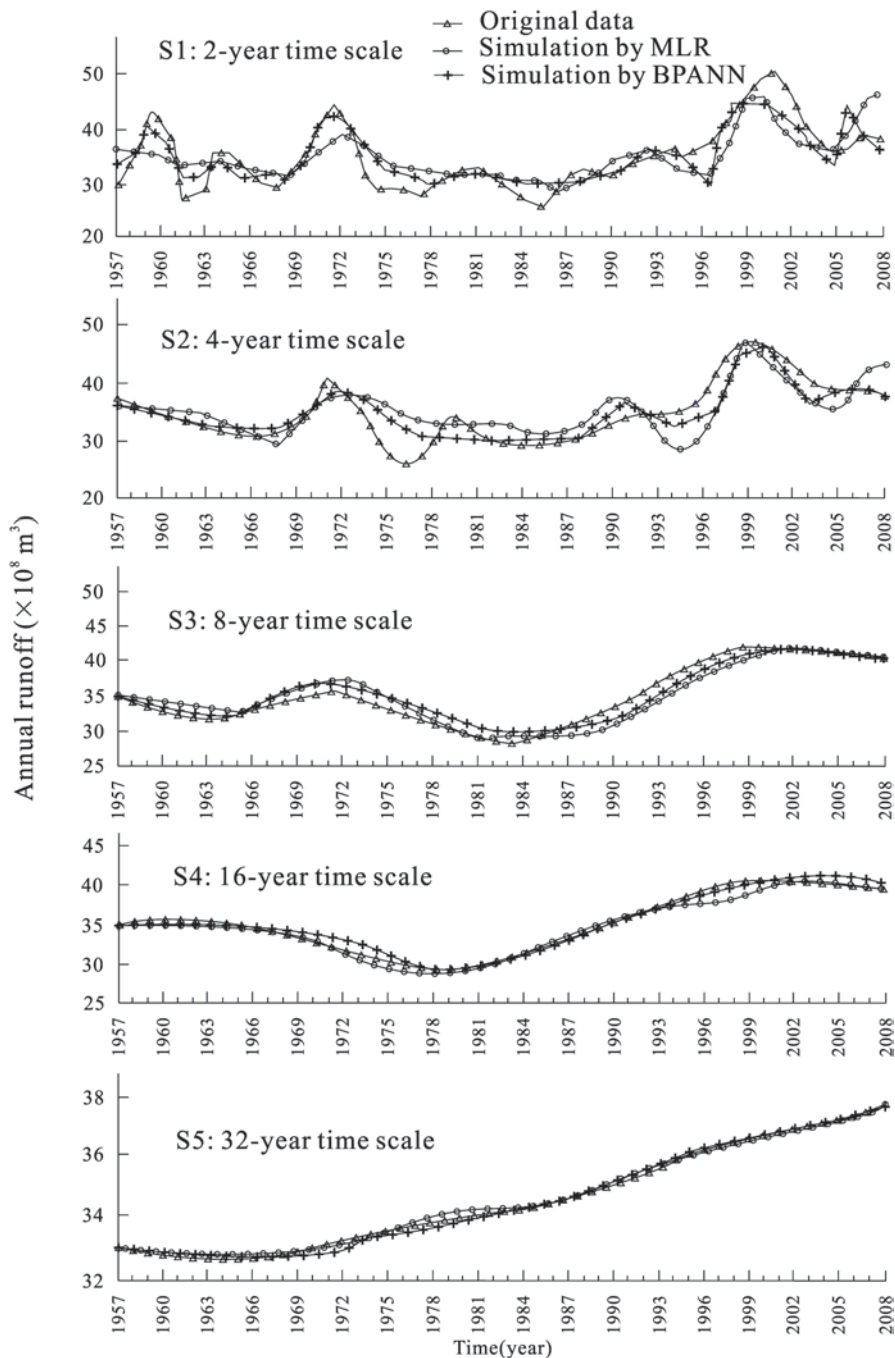


Fig. 8.8 Simulated results for AR by BPANN and MLR at the different time scales

**Table 8.5** Comparison between BPANN and MLR model in the Kaidu River at different time scales

Time scale		S1	S2	S3	S4	S5
MLRE	CD	0.541	0.556	0.916	0.977	0.978
	AIC	143.347	102.296	24.869	-63.774	-248.01
BPANN	CD	0.715	0.856	0.920	0.998	0.999
	AIC	118.4882	72.50646	22.28173	-151.494	-259.208

CD is the coefficient of determination; AIC is Akaike information criterion

**Acknowledgments** This work was supported by National Basic Research Program of China (973 Program; No: 2010CB951003).

## References

- Ai NS, Li HQ (1993) The nonlinear science methods for quaternary studies. *Quatern Res* 2:109–120. (in Chinese)
- Anderson DR, Burnham KP, Thompson WL (2000) Null hypothesis testing: problems, prevalence, and an alternative. *J Wildl Manage* 64(4):912–923
- Banakar A, Azeem MF (2008) Artificial wavelet neuro-fuzzy model based on parallel wavelet network and neural network. *Soft Comput* 12:789–808
- Bruce LM, Koger CH, Li J (2002) Dimensionality reduction of hyperspectral data using discrete wavelet transform feature extraction. *IEEE T Geosci Remote* 40(10):2331–2338
- Burnham KP, Anderson DR (2002) *Model selection and multimodel inference: a practical information-theoretic approach* (2nd ed). New York: Springer-Verlag, pp 49–97
- Cannon AJ, McKendry IG (2002) A graphical sensitivity analysis for statistical climate models: application to Indian monsoon rainfall prediction by artificial neural networks and multiple linear regression models. *Int J Climatol* 22:1687–1708
- Chen YN, Xu ZX (2005) Plausible impact of global climate change on water resources in the Tarim River Basin. *Sci China (D)* 48(1):65–73
- Chen YN, Takeuchi K, Xu CC, Chen YP, Xu ZX (2006) Regional climate change and its effects on river runoff in the Tarim Basin, China. *Hydrol Process* 20:2207–2216
- Chen YN, Xu CC, Hao XM, Li WH, Chen YP, Zhu CG, Ye ZX (2009) Fifty-year climate change and its effect on annual runoff in the Tarim River Basin, China. *Quatern Int* 208:53–61
- Chou CM (2007) Efficient nonlinear modeling of rainfall-runoff process using wavelet compression. *J Hydrol* 332:442–455
- Comte F, Renault E (1996) Long memory continuous time models. *Journal of Econometrics* 73(1):101–149
- Dorofki M, Elshafie AH, Jaafar O, Karim OA, Mastura S (2012) Comparison of artificial neural network transfer functions abilities to simulate extreme runoff data. *Int Proc Chem Biol Environ Eng* 33:39–44
- Farge M (1992) Wavelet transforms and their applications to turbulence. *Annu Rev Fluid Mech* 24:395–457
- Grassberger P, Procaccia I (1983) Characterization of strange attractor. *Phys Rev Lett* 50(5):346–349
- Hao XM, Chen YN, Xu CC, Li WH (2008) Impacts of climate change and human activities on the surface runoff in the Tarim river basin over the last fifty years. *Water Resour Manage* 22(9):1159–1171
- Hsu K, Gupta HV, Sorooshian S (1995) Artificial neural network modeling of the rainfall-runoff process. *Water Resour Res* 31(10):2517–2530

- Hurst HE (1951) Long-term storage capacity of reservoirs. *Trans Am Soc Civ Eng* 116:770–808
- Hurst HE, Black RP, Simaika YM (1965) Long-term storage: an experimental study. Constable, London, pp 1–155
- Kermani BG, Schiffman SS, Nagle HG (2005) Performance of the Levenberg–Marquardt neural network training method in electronic nose applications. *Sensors Actuat B: Chem* 110(1):13–22
- Labat D (2005) Recent advances in wavelet analyses: part 1. A review of concepts. *J Hydrol* 314:275–288
- Li ZL, Xu ZX, Li JY, Li ZJ (2008) Shift trend and step changes for runoff time series in the Shiyang River basin, northwest China. *Hydrol Process* 22:4639–4646
- Ling HB, Xu HL, Fu JY (2012) Temporal and spatial variation in regional climate and its impact on runoff in Xinjiang, China. *Water Resour Manag*. doi:10.1007/s11269-012-0192-0
- Maier HR, Dandy GC (1998) The effect of internal parameters and geometry on the performance of back-propagation neural networks: an empirical study. *Environ Modell Softw* 13:193–209
- Mallat SG (1989) A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans Pattern Anal Mach Intell* 11(7):674–693
- Mandelbrot BB, Wallis JR (1968) Noah, Joseph and operational hydrology. *Water Resource Res* 4(5):909–918
- Mandelbrot BB, Wallis JR (1969) Robustness of the rescaled range R/S in the measurement of noncyclic long-run statistical dependence. *Water Resource Res* 5(5):967–988
- Ramsey JB (1999) Regression over timescale decompositions: a sampling analysis of distributional properties. *Econ Systems Res* 11(2):163–183
- Shi YF, Shen YP, Kang ES, Li DL, Ding YJ, Zhang GW, Hu RJ (2007) Recent and future climate change in northwest china. *Climatic Change* 80:379–393
- Sivakumar B (2007) Nonlinear determinism in river flow: prediction as a possible indicator. *Earth Surf Process Landf* 32(7):969–979
- Smith LC, Turcotte DL, Isacks BL (1998) Streamflow characterization and feature detection using a discrete wavelet transform. *Hydrol Process* 12:233–249
- Sun GM, Dong XY, Xu GD (2006) Tumor tissue identification based on gene expression data using DWT feature extraction and PNN classifier. *Neurocomputing* 69:387–402. doi:10.1016/j.neucom.2005.04.005
- Torrence C, Compo GP (1998) A practical guide to wavelet analysis. *Bull Am Meteorol Soc* 79(1):61–78
- Turcotte DL (1997) *Fractals and chaos in geology and geophysics* (2nd ed). New York: Cambridge University Press, pp 158–162
- Wang J, Li H, Hao X (2010) Responses of snowmelt runoff to climatic change in an inland river basin, northwestern China, over the past 50 years. *Hydrol Earth Syst Sci* 14(10):1979–1987
- Xu JH (2002) *Mathematical methods in contemporary geography*. Higher Education Press, Beijing, pp 37–105. (in Chinese)
- Xu JH, Chen YN, Li WH, Dong S (2008a) Long-term trend and fractal of annual runoff process in mainstream of Tarim river. *Chinese Geogr Sci* 18(1):77–84
- Xu JH, Chen YN, Ji MH, Lu F (2008b) Climate change and its effects on runoff of Kaidu river, Xinjiang, China: a multiple time-scale analysis. *Chinese Geogr Sci* 18(4):331–339
- Xu JH, Chen YN, Li WH, Ji MH, Dong S (2009a) The complex nonlinear systems with fractal as well as chaotic dynamics of annual runoff processes in the three headwaters of the Tarim River. *J Geogr Sci* 19(1):25–35
- Xu JH, Chen YN, Li WH, Ji MH, Dong S, Hong YL (2009b) Wavelet analysis and nonparametric test for climate change in Tarim river basin of Xinjiang during 1959–2006. *Chinese Geogr Sci* 19(4):306–313
- Xu JH, Li WH, Ji M, Lu FH, Dong S (2010) A comprehensive approach to characterization of the nonlinearity of runoff in the headwaters of the Tarim river, western China. *Hydrol Process* 24(2):136–146
- Xu JH, Chen YN, Lu F, Li WH, Zhang LJ, Hong YL (2011a) The nonlinear trend of runoff and its response to climate change in the Aksu river, western China. *Int J Climatol* 31(5):687–695

- Xu JH, Chen YN, Li WH, Yang Y, Hong YL (2011b) An integrated statistical approach to identify the nonlinear trend of runoff in the Hotan river and its relation with climatic factors. *Stoch Env Res Risk A* 25(2):223–233
- Xu JH, Chen YN, Li WH, Nie Q, Hong YL, Yang Y (2012) The nonlinear hydro-climatic process in the Yarkand river, northwestern China. *Stoch Env Res Risk A*. doi:10.1007/s00477-012-0606-9
- Zhang Q, Xu CY, Tao H, Jiang T, Chen D (2010) Climate changes and their impacts on water resources in the arid regions: a case study of the Tarim river basin, China. *Stoch Env Res Risk A* 24(3):349–358