

# Integrating Wavelet Analysis and BPANN to Simulate the Annual Runoff With Regional Climate Change: A Case Study of Yarkand River, Northwest China

Jianhua Xu · Yaning Chen · Weihong Li · Qin Nie · Chunan Song · Chunmeng Wei

Received: 14 December 2012 / Accepted: 15 April 2014 /  
Published online: 5 May 2014  
© Springer Science+Business Media Dordrecht 2014

**Abstract** Selecting the Yarkand River as a typical representative of an inland river in northwest China, We identified the variation pattern of hydro-climatic process based on the hydrological and meteorological data during the period of 1957~2008, and constructed an integrated model to simulate the change of annual runoff (AR) with annual average temperature (AAT) and annual precipitation (AP) by combining wavelet analysis (WA) and artificial neural network (ANN) at different time scale. The results showed that the pattern of hydro-climatic process is scale-dependent in time. At 16-year and 32-year time scale, AR presents a monotonically increasing trend with the similar trend of AAT and AP. But at 2-year, 4-year, and 8-year time scale, AR exhibits a nonlinear variation with fluctuations of AAT and AP. The back propagation artificial neural network based on wavelet decomposition (BPANNBWD) well simulated the change of AR with AAT and AP at the all five time scales. Compared to the traditional statistics model, the simulation effect of BPANNBWD is better than that of multiple linear regression (MLR) at every time scale. The results also revealed the fact that the simulation effect at a larger time scale (e.g. 16-year or 32-year scale) is better than that at a smaller time scale (e.g. 2-year or 4-year scale).

**Keywords** Annual runoff · Regional climate change · Wavelet decomposition · Back-propagation artificial neural network · Inland river

## 1 Introduction

Water is a critical ecological element because it is scarce in an arid environment (Kim et al. 2003). Therefore, water source has severely restricted the sustainable development in the arid region of northwest China. For the arid region in northwest China, the water resource which can be utilized is mainly from the streamflow of inland rivers. So the runoff variation of an

---

J. Xu (✉) · Q. Nie · C. Song · C. Wei  
The Research Center for East–west Cooperation in China, Department of Geography, East China Normal University, Shanghai 200241, 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, Urumqi 830011, China

inland river has aroused more and more attention (Chen et al. 2009; Hao et al. 2008; Li et al. 2008; Xu et al. 2008a, 2009a; Wang et al. 2013).

As a scientific base for water utilization planning, hydrological prediction has been paid attention by the scientists in various countries. One of the traditional forecasting methods for hydrology system is the regressive analysis (Ding and Deng, 1988). But in the recent 20 years, various methods such as grey model (Deng 1989), functional-coefficient time series model (Shao et al. 2009), wavelet analysis (Smith et al. 1998; Chou 2007; Sang 2012), artificial neural network (Hsu et al. 1995; Hu et al. 2008), fractal and chaotic theory (Wilcox et al. 1991; Xu et al. 2008a, 2009a) have been widely applied in the hydrological prediction. Specially, some hybrid models (Kim et al. 2003; Lin 2006; Wang and Ding 2003; Sahay and Srivastava 2014; Yazar 2014) have been used.

Because the hydrology process is closely interconnected with climatic process, and greatly influenced by climate change, the key of hydrological prediction is to understand the hydro-climatic process. 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. Many case studies in different countries and regions have suggested that the hydro-climatic process is a complex system, with nonlinearity as its basic characteristic (Ibbitt and Woods 2004; Sivakumar 2007; Wang et al. 2008; Xu et al. 2011a, 2013a). But it is difficult to achieve a thorough understanding of the complex mechanism of nonlinear hydro-climatic process (Xu et al. 2010).

In the last 10 years, many studies have been conducted to evaluate climatic change and hydrological processes in the arid regions of northwestern China (Chen and Xu 2005; Wang et al. 2010; Xu et al. 2011a, 2011b; 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; 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 bring 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 understand the nonlinear hydro-climatic process from different perspectives by using different methods (Xu et al. 2013a).

To further understand the change of runoff in an inland river and its response to regional climate change in northwest China at different time scales, we selected the Yarkand River as a typical representative to analyze the variation patterns of annual runoff (AR), annual average temperature (AAT) and annual precipitation (AP) by using wavelet analysis (WA), and constructed an integrated model to simulate the variation of AR with AAT and AP by integrating wavelet decomposition (WD) and back propagation artificial neural network (BPANN).

## 2 Study Area and Data

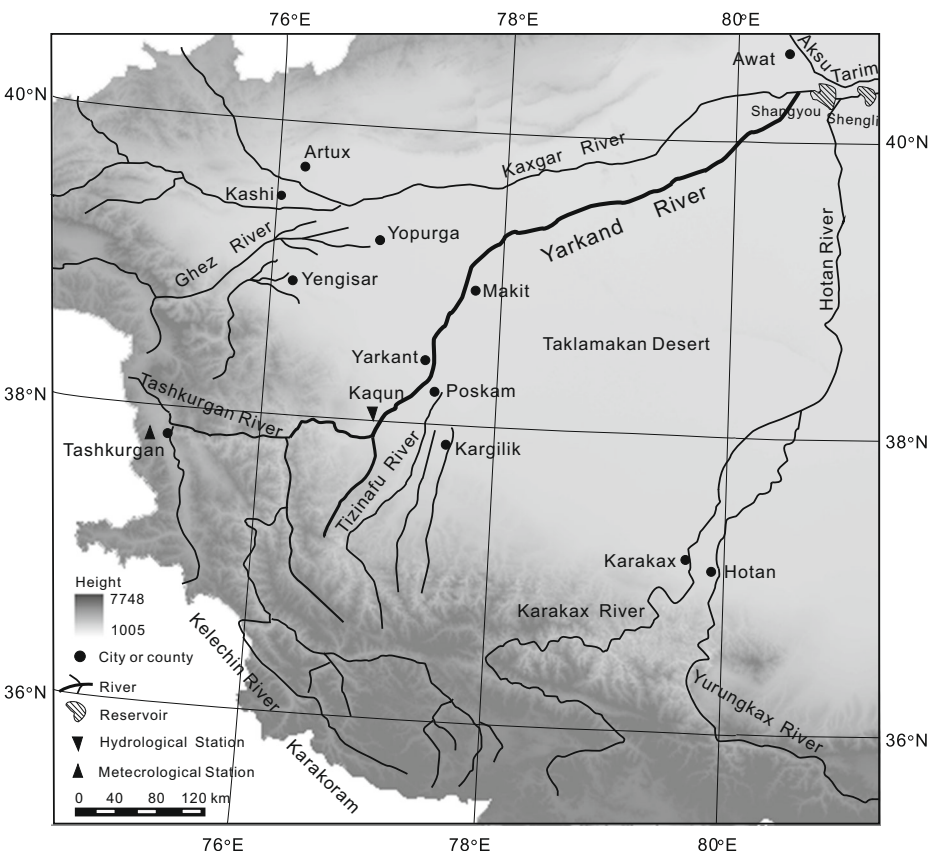
### 2.1 Study Area

Originating from the surrounding mountains, one of the headwaters of the Tarim River, the Yarkand River is a typical inland river in northwest China, which has been relatively undisturbed by human activities. Therefore, to understand the response of runoff in an inland river to regional climate change in northwest China, we specially selected the Yarkand River as a case for modelling and demonstration in this study.

The Yarkand River (Fig. 1) is located in the southeastern region of Xinjiang Uygur Autonomous Region, with a length of 1097 km. The Yarkand River ( $35^{\circ}40' \sim 40^{\circ}31' \text{N}$ ,  $74^{\circ}28' \sim 80^{\circ}54' \text{E}$ ) has a total basin area of  $9.89 \times 10^4 \text{ km}^2$ , including  $6.08 \times 10^4 \text{ km}^2$  as the mountain area, which accounts for 61.5 %, and  $3.81 \times 10^4 \text{ km}^2$  as the plain area, which takes up 38.5 % (Sun et al. 2006). The main stream of the Yarkand River originates from Karakoram Pass in the north slope of Karakoram Mountain, where is full of towering peaks and glaciers. The Yarkand River is a typical inland river, and there is rare precipitation and no water recharge to the stream in the plain area. The multi-year average runoff in the Yarkand River consists of 64.0 % from glacial ablation, 13.4 % from rain and snow supply, and 22.6 % from groundwater supply, respectively (Sabit and Tohti 2005; Liu et al. 2008).

## 2.2 Data

As a headwater of the Tarim River, the streamflow is rarely disturbed by human activities, whereas mainly affected by climatic factors, especially temperature and precipitation in the mountainous (Hao et al. 2008; Tao et al. 2011). In order to analyze the stream flow of the Yarkand River and its response to regional climate change, we use the data of runoff as well as temperature and precipitation. The runoff data were from the Kaqun hydrological station, and



**Fig. 1** Location of the study area

temperature and precipitation data were from Tash Kurghan meteorological station. The two stations are located in the source areas of the river; the amount of water used by humans is minimal compared to the total discharge. Therefore, the observed hydrological and meteorological records reflect the natural conditions.

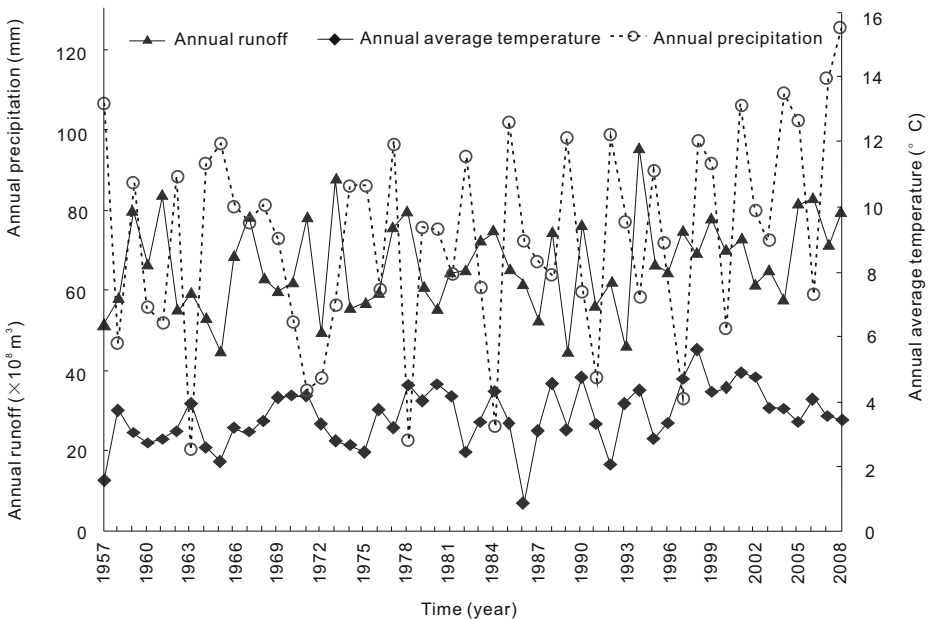
To investigate the annual variation of runoff and its relation with regional climate change, this study used the raw data of annual runoff (AR), annual average temperature (AAT) and annual precipitation (AP) from 1957 to 2008, which were revealed in Fig. 2.

### 3 Methods

In order to simulate the annual runoff and its response to regional climate change at different time scales, we constructed an integrated model by integrating wavelet analysis (WA) and artificial neural network (ANN). We firstly identified the variation pattern of annual runoff and its related climatic factors by using WA at a given time scale, and then simulated the variation of annual runoff with its related climatic factors by using ANN based on WA at the corresponding time scale.

#### 3.1 Wavelet analysis

Wavelet transformation has been shown to be a powerful technique for characterizing the frequency, intensity, scale, and duration of variations in hydro-climatic process (Labat 2005; Chou 2007; Xu et al. 2009b; Sang 2012). Wavelet analysis can also reveal localized time and frequency information without requiring the signal time series to be stationary, as required by the Fourier transform and other spectral methods (Farge 1992; Torrence and Compo 1998).



**Fig. 2** The raw data of AR, AAT and AP for the study period

One of our tasks in this paper is to approximate the variation patterns of runoff and its related climatic factors by using wavelet decomposition and reconstruction at different time scales.

The principle of wavelet decomposition and reconstruction is as follows (Mallat 1989; Farge 1992; Torrence and Compo 1998). Considering a given signal  $X(t)$ , such as AR, AAT and AP, etc., which can be built up as a sequence of projections onto Father and Mother wavelets indexed by both  $k \{k=1, 2, \dots\}$  and  $s \{s=2^j, j=1, 2, \dots\}$ . The coefficients in the expansion are given by the projections

$$\begin{cases} s_{J,k} = \int X(t)\Phi_{J,k}(t)dt \\ d_{j,k} = \int X(t)\Psi_{j,k}(t)dt \end{cases} \quad j = 1, 2, \dots, J \tag{1}$$

where  $J$  is the maximum scale sustainable by the number of data points,  $\Phi_{j,k} = 2^{-j/2}\Phi\left(\frac{t-2^j k}{2^j}\right)$  is father wavelet, and  $\Psi_{j,k} = 2^{-j/2}\Psi\left(\frac{t-2^j k}{2^j}\right)$  is mother wavelet. Generally, father wavelet is used for the lowest-frequency smooth components, which requires wavelet with the widest support; mother wavelet is used for the highest-frequency detailed components. In other words, father wavelet is used for the major trend components, and mother wavelet is used for all deviations from the trend.

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: a low scale corresponds to compressed wavelet as well as rapidly changing details, namely high frequency, whereas a 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.

The representation of the signal  $X(t)$  now can be given by:

$$X(t) = S_J + D_J + D_{J-1} + \dots + D_j + \dots + D_1 \tag{2}$$

where  $S_J = \sum_k s_{J,k}\Phi_{J,k}(t)$  and  $D_j = \sum_k d_{j,k}\Psi_{j,k}(t), j = 1, 2, \dots, J$

In general, we have the relationship as

$$S_{j-1} = S_j + D_j \tag{3}$$

where  $\{S_j, S_{j-1}, \dots, S_1\}$  is a sequence of multi-resolution approximations of the function  $X(t)$  at ever-increasing levels of refinement. The corresponding multi-resolution decomposition of  $X(t)$  is given by  $\{S_j, D_j, D_{j-1}, \dots, D_j, \dots, D_1\}$ .

Selecting a proper wavelet function is a prerequisite for wavelet analysis. The actual criteria for wavelet selection include self-similarity, compactness, and smoothness (Ramsey 1999; Xu et al. 2004). Choosing the Symmlet as the basic wavelet, we experimented with alternative choices of scaling functions, and found the qualitative results from ‘Sym8’ are robust. Therefore, ‘Sym8’ were used for approximating the variation patterns of AR, AAT and AP at different time scales in this study.

### 3.2 Back Propagation Artificial Neural Network Based on Wavelet Decomposition

In order to simulate the variation of annual runoff with regional climate change at different time scales, we constructed an integrated approach, a back propagation artificial neural network based on wavelet decomposition (BPANNBWD). We first approximated the variation patterns of AR, AAT and AP by using wavelet decomposition on the basis of the discrete wavelet transform (DWT) at a given time scale, and then simulated the variations of AR with AAT and AP by using the back propagation artificial neural network based on the wavelet approximation at the corresponding time scale.

In the back-propagation artificial neural 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):

$$I_j = \sum w_{ji}x_i + \theta_j \quad (4)$$

$$y_i = f(I_j) \quad (5)$$

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:

$$f(I) = \frac{1 - e^{-I}}{1 + e^{-I}} \quad (6)$$

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

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:

$$E = \langle \|y - y^{obs}\|^2 \rangle \quad (7)$$

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.

It is evident that the BPANNBWD model is a multivariable simulation model for cause and effect, which is different from the hybrid models for hydrologic prediction of a time series such as the wavelet network (Wang and Ding 2003), wavelet neuro fuzzy model (Yarar 2014) and wavelet transform-genetic algorithm-neural network model (Sahay and Srivastava 2014). In the BPANNBWD model, variations of AAP and AP are regarded as the causes of the variation of AR, namely the variation of annual runoff is its response to regional climate change.

### 3.3 Test for Simulation Effect

To further compare the simulation effect of BPANNBWD with that of traditional statistics model such as multiple linear regression (MLR) at different time scales, we also constructed the multiple linear regression based on wavelet decomposition (MLRBWD). The principle of MLRBWD is as follows (Xu et al. 2008b, 2013c): the data series of AR, AAT and AP were firstly approximated by using wavelet decomposition (DW) at a chosen time scale; then the variations of AR with AAT and AP were estimated by using the multiple linear regression (MLR) based on the wavelet approximation at the corresponding time scale.

At each time scale, the variation of annual runoff with regional climate change was simulated by the multiple linear regression equation (MLRE) as follows:

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_kx_k \tag{8}$$

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.

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

$$CD = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \tag{9}$$

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 CD value is a measure of how well the simulate results represent the actual data.

Statistics tells us that a bigger CD indicates a higher certainty and lower uncertainty of the simulation model (Xu 2002).

To compare the relative goodness of models at different time scales, we also used the measure of Akaike information criterion (AIC). The formula of AIC is as follows (Anderson et al. 2000):

$$AIC = 2k + n \ln(RSS/n) \tag{10}$$

where  $k$  is the number of parameters estimated in the model;  $n$  is the number of samples; RSS is the same as in formula (9)

Akaike information criterion means that 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

$$AIC_c = AIC + \frac{2k(k+1)}{n-k-1} \quad (11)$$

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).

## 4 Results and discussion

### 4.1 Variation patterns of annual runoff and its related regional climate factors

Our previous study indicated that (Xu et al. 2010, 2011a), the annual average temperature (AAT) and annual precipitation (AP) are the most important two 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. 2013).

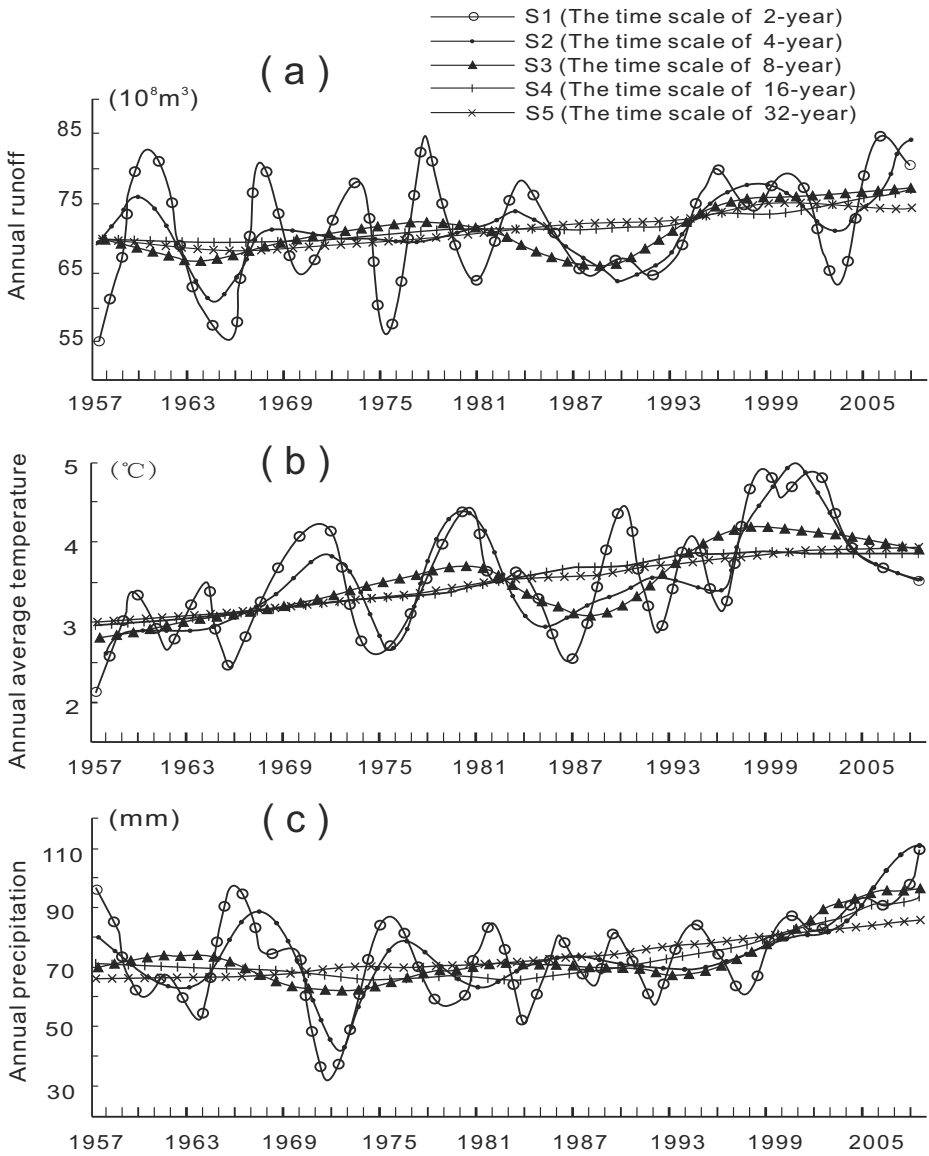
Because the original signal series of AR, AAT and AP presented fluctuations with high frequency, it is hard to identify any variation patterns from the raw data (Xu et al. 2011a, 2013a). For mining out the variation patterns of AR, AAT and AP, we need to represent the signal series by wavelet analysis at different time scales (Xu et al. 2008b, 2009b).

For computing the wavelet decomposition of AR, AAT and AP, the five time scales are designated as S1 to S5, which represent 2-year, 4-year, 8-year, 16-year and 32-year time scale respectively. We used the method described in section 3.1 to represent the signal series of AR, AAT and AP, and the results are shown as Fig. 3.

Figure 3a revealed the five variation patterns for AR at the time scale of S1, S2, S3, S4, and S5 respectively. The S1 curve retains a large amount of residual noise from the raw data (see Fig. 2 for a comparison), and drastic fluctuations along the entire time span. Furthermore, the S1 curve also indicates that, although the annual runoff varied greatly throughout the study period, there was a hidden slightly increasing trend. The S2 curve still retains a considerable amount of residual, as indicated by the 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 retains much less residual, as indicated by the presence of 2 peaks and 2 valleys. Compared to S2, the increase in runoff over time is more apparent in S3. Finally, the S5 curve presents an ascending tendency, whereas the increasing trend is obvious in the S4 curve.

Accordingly, Fig. 3b and Fig. 3c provide us the cognition for comparing the variation patterns of AAT and AP at different time scales. Similar to AR, AAT and AP also presented five variation patterns at the time scale of S1, S2, S3, S4, and S5 respectively. The S1 and S2 curve showed a hidden slightly increasing trend with drastic fluctuations, whereas the curves are getting much smoother and the increasing trend becomes even more obvious with the time scale increases. It is evident that the S4 and S5 curve for AAT and AP exhibited a visible monotonically increasing trend similar to that for AR.





**Fig. 3** Variation patterns of AR, AAT, and AP at different time scales

In conclusion, the results as Fig. 3 revealed a fact that the pattern of hydro-climatic process is scale-dependent in time. At 16-year and 32-year time scale, AR presented a monotonically increasing trend with the similar trend of AAT and AP. But at 2-year, 4-year, and 8-year time scale, AR presented a nonlinear variation with fluctuations of AAT and AP.

#### 4.2 Simulation for Annual Runoff With Regional Climate Change

To simulate the variation of annual runoff with regional climate change at different time scales, we used the BPANNBDW method described in section 3.2 to individually construct a 3-layer

**Table 1** Optimized parameters of BPANN at different time scales

Time Scale	Neuron number of the hidden layer	Input variables	Output variables	Best epoch
S1	3	AAT, AP	AR	11
S2	5	AAT, AP	AR	4
S3	5	AAT, AP	AR	107
S4	5	AAT, AP	AR	4
S5	4	AAT, AP	AR	7

back-propagation artificial neural network based on the wavelet decomposition results of AR, AAT and AP for each time scale.

The network structure of the 3-layer back-propagation artificial neural network is “2-1-1”, which indicates that there is an input layer with two variables (i.e. AAT and AP), an output layer with one variables (i.e. AR), and a hidden layer in the network. The neuron number of the hidden layer in the network for each time scale is not same, which is 3, 5, 5, 5 and 4 at the time scale of S1, S2, S3, S4 and S5 respectively (Table 1).

Using the computing software, MATLAB, we selected the transfer function as tansig, and training function as trainlm 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, the all network models have reached the expected error target (0.001) with learning rate of 0.01. The optimized parameters of the BPANN to simulate the annual runoff with regional climate change at different time scales are showed in Table 1.

Table 2 listed the simulation error of BPANNBWD at the 2-year, 4-year 8-year, 16-year, and 32-year time scale. The results showed that, as the time scale increased from S1 to S5, the smaller the estimated error is, and the better the simulation effect is.

Figure 4 shows the simulated results by BPANNBWD on the five time scales, which take input variables: annual average temperature (AAT) and annual precipitation (AP), to simulate output variable: annual runoff (AR). In this figure, sub-figure (a), (b), (c), (d) and (e) show the comparison between original data of annual runoff and its simulation, at the 2-year, 4-year 8-year, 16-year, and 32-year time scale, respectively.

By comparing the CD and AIC value in Table 3, we can know the effect (good or bad) of different models at different time scales.

The CD value at the time scale of S1 and S2 (i.e. 2-year and 4-year scale) is 0.2103 and 0.4924 respectively, which is relatively lower; but that at the time scale of S3, S4 and S5 (i.e. 8-year, 16-year and 32-year scale) is high as 0.9904, 0.9955 and 0.9986 respectively. These results indicated that the certainty of estimates at the larger time scale (i.e. 8-year,

**Table 2** Simulation errors of BPANNBWD at different time scales

Time scale	S1	S2	S3	S4	S5
Average absolute error	5.8683	2.7563	0.2259	0.1131	0.0651
Average relative error	9.07 %	4.23 %	0.35 %	0.17 %	0.10 %

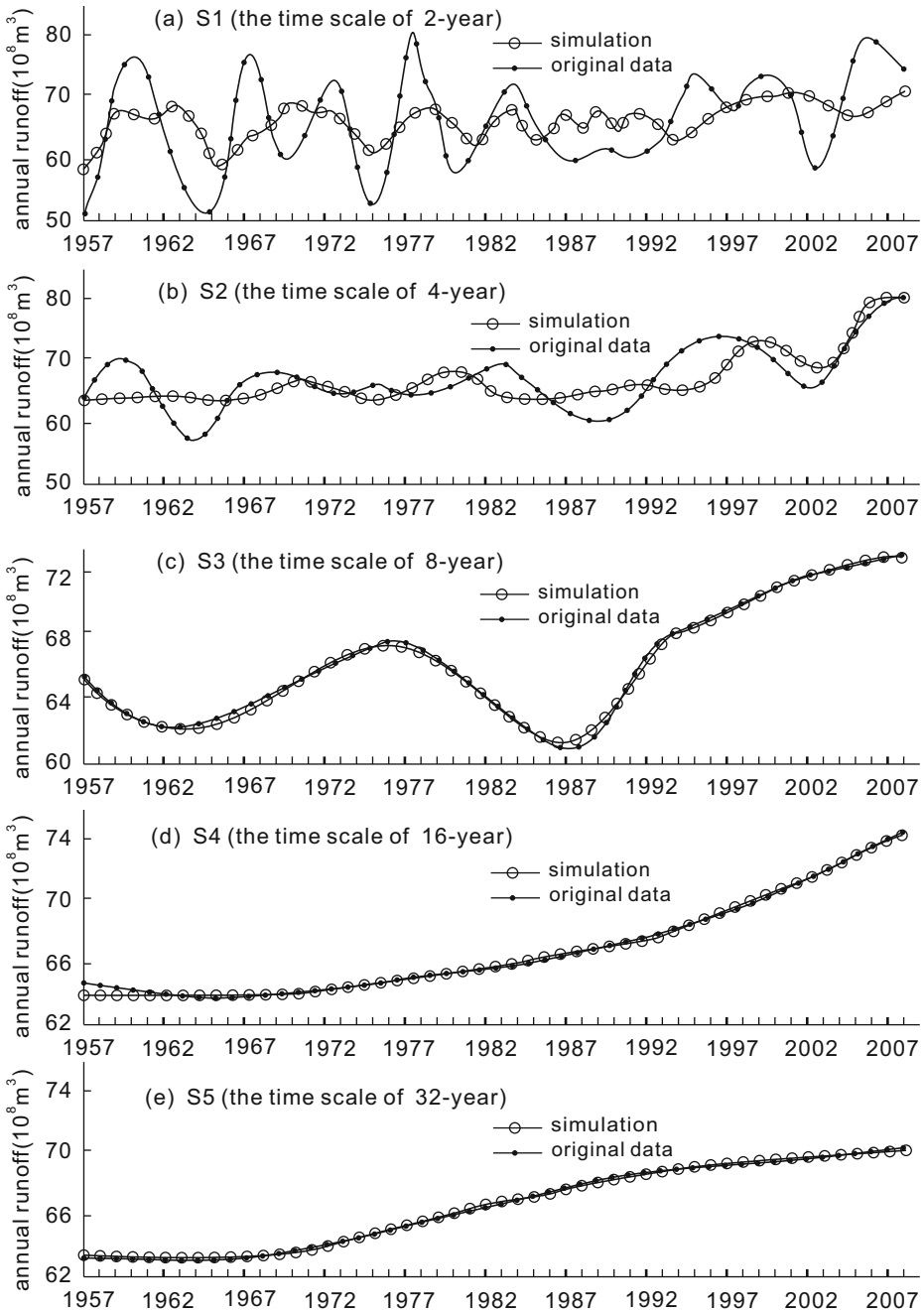


Fig. 4 Comparison between simulated values for AR and its original data at different time scale

16-year and 32-year scale) are markedly higher than the smaller time scale (i.e. 2-year and 4-year scale).

**Table 3** CD and AIC values of BPANNBWD at different time scales

Time Scale	CD	AIC
S1	0.361	204.542
S2	0.4924	131.308
S3	0.9904	-112.333
S4	0.9955	-186.356
S5	0.9986	-260.496

The AIC value at each time scale also tells us the similar results, the model at time scale of S5 is the best, that at time scale of S4 is better, that at time scale of S3 is moderate, that at time scale of S2 is the penult, and that at time scale of S1 is the worst.

In fact, our related study (Xu et al. 2008b, 2011a, 2011b, 2013b) revealed that the hydro-climatic process at a large time scale (e.g. 16-year or 32-year scale) is basically a linear process with monotonic trend, but at a small time scale (e.g. 2-year or 4-year scale) the process is essentially a nonlinear process with complicated causations. Therefore, the estimated precision at a large time scale (e.g. 16-year or 32-year scale) is high, whereas it is difficult to accurately predict at a small time scale (e.g. 2-year or 4-year scale).

#### 4.3 Comparing the Simulation Effect With the Traditional Statistics Method

To compare the simulation effect of BPANNBWD with that of the traditional statistics method such as multiple linear regression (MLR), we also used the MLRBWD method described in section 3.3 to fit a group of multiple linear regression equations (MLREs) for simulating the variations of AR with AAT and AP at the five time scales. Table 4 showed the four indices, average absolute error, average relative error, CD and AIC for the MLRBWD at the time scale of S1, S2, S3, S4 and S5 respectively.

Comparing table 2 to table 4, we found that the simulation errors, average absolute error and average relative error of BPANNBWD are both smaller than those of MLRBWD at every time scale. Comparing table 3 to table 4, we found that the CD value of BPANNBWD is bigger than that of MLRBWD at every time scale. Moreover, the AIC value of BPANNBWD is smaller than that of MLRBWD at every time scale. That is to say, all the four indices, average absolute error, average relative error, CD and AIC indicate that the simulated effect from BPANNBWD is better than that from MLRBWD at every time scale.

**Table 4** Simulation effect of MLRBWD at different time scales

Time scale	Average absolute error	Average relative error	CD	AIC
S1	6.266	9.665 %	0.2103	209.924
S2	3.250	4.959 %	0.416	143.263
S3	0.880	1.354 %	0.894	12.960
S4	0.329	0.502 %	0.975	-96.714
S5	0.109	0.165 %	0.996	-209.172

## 5 Conclusions

Selecting the Yarkand River as a typical representative, this study identified the variation pattern of annual runoff (AR), annual average temperature (AAT) and annual precipitation (AP) by using wavelet analysis (WA), and constructed an integrated model, i.e. the back propagation artificial neural network based on wavelet decomposition (BPANNBWD), to simulate the change of AR with AAT and AP at different time scales.

The main conclusions of this study are as follows:

1. The integrated approach combining wavelet analysis (WA) and artificial neural network (ANN) provides a way to understand the relationship between the runoff of an inland river and its related climatic factors in northwest China from a multi-scale perspective.
2. Variations of annual runoff (AR) with annual average temperature (AAT) and annual precipitation (AP) are scale-dependent in time. The variation patterns of AR, AAT and AP present basically linear trends at 16-year and 32-year time scale, but they exhibit nonlinear fluctuations at 2-year, 4-year and 8-year time scale.
3. The BPANNBWD well simulated the variation of AR with AAT and AP at the all five time scales. Compared to the traditional statistics method, the simulation effect of BPANNBWD is better than that of multiple linear regression (MLR) at every time scale. Moreover, the simulation effect at a larger time scale (e.g. 16-year or 32-year scale) is better than that at a smaller time scale (e.g. 2-year or 4-year scale).

**Acknowledgments** This work was supported by State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences.

## References

- Anderson DR, Burnham KP, Thompson WL (2000) Null Hypothesis Testing: Problems, Prevalence, and an Alternative. *J Wildl Manag* 64(4):912–923
- Bruce LM, Koger CH, Li J (2002) Dimensionality Reduction of Hyperspectral Data Using Discrete Wavelet Transform Feature Extraction. *IEEE Trans Geosci Remote Sens* 40(10):2331–2338
- Burnham KP, Anderson DR (2002) *Model Selection and Multimodel Inference: a practical information-theoretic approach (2nd Ed.)*. Springer-Verlag, New York, pp 49–97
- 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. *Quat Int* 208:53–61
- Chou CM (2007) Efficient Nonlinear Modeling of Rainfall-Runoff Process Using Wavelet Compression. *J Hydrol* 332:442–455
- Deng JL (1989) Introduction to Grey System Theory. *J grey syst* 1(1):1–24
- Ding J, Deng Y (1988) *Stochastic Hydrology*. Chengdu University of science and technology Press, Chengdu, China, pp 190–230 (in Chinese)
- Dorofki M, Elshafie AH, Jaafar O, Karim OA, Mastura S (2012) Comparison of Artificial Neural Network Transfer Functions Abilities to Simulate Extreme Runoff Data. *International Proceedings of Chemical, Biological and Environmental Engineering* 33:39–44

- Farge M (1992) Wavelet Transforms and Their Applications to Turbulence. *Annu Rev Fluid Mech* 24:395–457
- 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 Manag* 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
- Hu CH, Hao YH, Yeh TCJ, Pang B, Wu ZN (2008) Simulation of Spring Flows from a Karst Aquifer With an Artificial Neural Network. *Hydrol Process* 22:596–604
- Ibbitt R, Woods R (2004) Re-Scaling the Topographic Index to Improve the Representation of Physical Processes in Catchment Models. *J Hydrol* 293:205–218
- Kermani BG, Schiffman SS, Nagle HG (2005) Performance of the Levenberg–Marquardt Neural Network Training Method in Electronic Nose Applications. *Sensors Actuators B Chem* 110(1):13–22
- Kim T-W, Valdés JB, Yoo C (2003) Nonparametric Approach for Estimating Return Periods of Droughts in Arid Regions. *J Hydrol Eng* 8(5):237–246
- 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(23):4639–4646
- Lin CJ (2006) Wavelet Neural Networks With a Hybrid Learning Approach. *J Inf Sci Eng* 22(6):1367–1387
- Ling HB, Xu HL, Fu JY (2013) Temporal and Spatial Variation in Regional Climate and its Impact on Runoff in Xinjiang, China. *Water Resour Manag* 27(2):381–399
- Liu TL, Yang Q, Qin R, He YP, Liu R (2008) Climate Change Towards Warming-Wetting Trend and its Effects on Runoff at the Headwater Region of the Yarkand River in Xinjiang. *J Arid Land Res Environ* 22(9):49–53 (in Chinese)
- Mallat SG (1989) A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Trans Pattern Anal Mache Intell* 11(7):674–693
- Maier HR, Dandy GC (1998) The Effect of Internal Parameters and Geometry on the Performance of Back-Propagation Neural Networks: An Empirical Study. *Environ Model Softw* 13:193–209
- Ramsey JB (1999) Regression Over Timescale Decompositions: A Sampling Analysis of Distributional Properties. *Econ Syst Res* 11(2):163–183
- Sahay RR, Srivastava A (2014) Predicting Monsoon Floods in Rivers Embedding Wavelet Transform, Genetic Algorithm and Neural Network. *Water Resour Manag* 28(2):301–317
- Sang YF (2012) A Practical Guide to Discrete Wavelet Decomposition of Hydrologic Time Series. *Water Resour Manag* 26(11):3345–3365
- Shao QX, Wong H, Li M, Ip WC (2009) Streamflow Forecasting Using Functional-Coefficient Time Series Model With Periodic Variation. *J Hydrol* 368:88–95
- Sabit M, Tohti A (2005) An Analysis of Water Resources and its Hydrological Characteristic of Yarkend River Valley. *Journal of Xinjiang Normal University (Natural Sciences Edition)* 24(1):74–78 (in Chinese)
- 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 BG, Mao WY, Feng YR, Chang T, Zhang LP, Zhao L (2006) Study on the Change of Air Temperature, Precipitation and Runoff Volume in the Yarkant River Basin. *Arid Zone Research* 23(2):203–209 (in Chinese)
- Torrence C, Compo GP (1998) A Practical Guide to Wavelet Analysis. *Bull Am Meteorol Soc* 79(1):61–78
- Tao H, Gemmer M, Bai Y, Su B, Mao W (2011) Trends of Streamflow in the Tarim River Basin During the Past 50 Years: Human Impact or Climate Change? *J Hydrol* 400(1):1–9
- Wang W, Chen X, Shi P, van Gelder PHAJM (2008) Detecting Changes in Extreme Precipitation and Extreme Streamflow in the Dongjiang River Basin in Southern China. *Hydrol Earth Syst Sci* 12:207–221
- Wang HJ, Chen YN, Li WH, Deng HJ (2013) Runoff Responses to Climate Change in Arid Region of Northwestern China During 1960–2010. *Chin Geogr Sci* 23(3):286–300
- Wang WS, Ding J (2003) Wavelet Network Model and Its Application to the Prediction of Hydrology. *Nature and Science* 1(1):67–71
- 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
- Wilcox BP, Seyfried MS, Matison TH (1991) Searching for Chaotic Dynamics in Snowmelt Runoff. *Water Resour Res* 27(6):1005–1010
- Xu JH (2002) *Mathematical Methods in Contemporary Geography*. Higher Education Press, Beijing, pp 37–105 (in Chinese)
- Xu JH, Lu Y, Su FL, Ai NS (2004) R/S and Wavelet Analysis on Evolutionary Process of Regional Economic Disparity in China During the Past 50 Years. *Chin Geogr Sci* 14(3):193–201

- Xu JH, Chen YN, Li WH, Dong S (2008a) Long-Term Trend and Fractal of Annual Runoff Process in Mainstream of Tarim River. *Chin 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. *Chin 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. *Chin Geogr Sci* 19(4):306–313
- Xu JH, Li WH, Ji MH, Lu F, 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 (2013a) The Nonlinear Hydro-Climatic Process in the Yarkand River, Northwestern China. *Stoch Env Res Risk A* 27(2):389–399
- Xu JH, Chen YN, Li WH, Peng PY, Yang Y, Song CN, Wei CM, Hong YL (2013b) Combining BPANN and Wavelet Analysis to Simulate Hydro-Climatic Processes—a Case Study of the Kaidu River, North-West China. *Frontiers of Earth Science* 7(2):227–237
- Xu JH, Xu YW, Song CN (2013c) An integrative approach to understand the climatic-hydrological process: a case study of Yarkand River, Northwest China. *Advances in Meteorology* Vol. 2013: Article ID 272715, 9 pages. DOI:10.1155/2013/272715
- Yarar A (2014) A Hybrid Wavelet and Neuro-Fuzzy Model for Forecasting the Monthly Streamflow Data. *Water Resour Manag* 28(2):553–565
- 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