Desalination and Water Treatment
www.deswater.com
doi: 10.5004/dwt.2018.22363

129 (2018) 201–206
October

Yellow River runoff forecast based on Volterra adaptive filter

Yanbin Li[a,*], Xinli Zhu[a], Cui Tian[b], Xiaoya Chang[a], Feng You[a]

[a]*North China University of Water Resources and Electric Power, Henan, Zhengzhou, China, 450045, emails: Liyb101@sina.com (Y. Li), zhuxinli@ncwu.edu.cn (X. Zhu), Changxiaoya@ncwu.edu.cn (X. Chang), Youfeng@ncwu.edu.cn (F. You)*
[b]*Renhuai Water Affairs Bureau, Guizhou, China, 564500, email: Tiancui@ncwu.edu.cn (C. Tian)*

Received 24 February 2018; Accepted 8 April 2018

ABSTRACT

This paper reconstructs the phase space for the runoff time series of three typical stations in the middle and lower reaches of the Yellow River by using the C–C algorithm and proves that the runoff time series of the Yellow River has chaotic characteristics through calculating the largest Lyapunov index. The three-order Volterra adaptive filter model is used to predict the runoff time series of three stations, which has a better forecast effect. A new method is provided for the forecast of river runoff.

Keywords: Chaos; Phase space reconstruction; Volterra filter; Adaptive; Runoff forecast; The Yellow River

1. Introduction

The formation of river runoff is influenced by a variety of random factors such as climate, rainfall, watershed characteristics, geographical environment, and human disturbance. And the mechanism of action of each factor is often difficult to accurately describe in precise mathematical languages. Therefore, the river runoff forecast is a complex nonlinear non-stationary system problem. The traditional runoff forecasting model cannot accurately describe the variation characteristics of river runoff, and the prediction accuracy is also declining naturally. This makes it possible to study new methods and improve the conventional methods and complementary use of various methods. Volterra series can effectively characterize nonlinear nonstationary system problems. Volterra adaptive filter can adaptively track the motion trajectory of chaotic system and has high prediction accuracy [1,2].

Based on the previous scholars' research [3–6], the runoff time series of the Yellow River has chaotic properties, but the calculation is cumbersome and the calculation speed is slower. The C–C algorithm [7] has the characteristics of fast calculation and simultaneous estimation of the delay time and the embedded dimension. This paper reconstructs the phase space for the runoff time series by using the C–C algorithm and calculates the largest Lyapunov index of runoff time series. On the basis of this, the third-order Volterra adaptive filter model is applied to the Yellow River runoff forecast for the first time. The forecast results show that the Volterra adaptive filter model can effectively reflect the future trend of runoff series and achieve high prediction accuracy, which can provide the basis for the Yellow River runoff forecast.

2. Volterra adaptive filter prediction principle

2.1. Phase space reconstruction of time series

The judgment and prediction of chaotic time series are based on phase space reconstruction. According to Takens' embedding theorem, for a given time series, if $m \geq 2d + 1$ (d is the correlation dimension of the dynamical system, and m is the embedded dimension), the attractor can be recovered in the m-dimensional reconstruction space. Phase trajectory in reconstructed space and differential of motivation system are homeomorphism, in addition, the reconstructed space

* Corresponding author.

Presented at the 3rd International Conference on Recent Advancements in Chemical, Environmental and Energy Engineering, 15–16 February, Chennai, India, 2018.

1944-3994/1944-3986

is topological equivalent to the primitive dynamical system [8]. Assuming that the time series is $\{x(t)\}$, $t = 1, 2, \cdots, n$, n is the number of time series, If the embedded dimension is m and the time delay is f, the phase space is reconstructed as follows:

$$Xm(t) = [x(t), x(t+f), x(t+2f)\ldots x(t+(m-1)f)] \quad t = 1, 2, \ldots, N; N = n-(m-1)f \tag{1}$$

For a specific time series, the Lyapunov index quantitatively describes the property of exponential divergence between adjacent orbits in phase space, which is used to reflect the initial sensitivity of the chaotic motion. In the actual system chaotic recognition, only the maximum Lyapunov index is usually estimated. It indicates that the system has chaotic properties when $L > 0$. The main methods are small data method [8], Wolf method, [9] and so on.

2.2. The C–C method selects the reconstructed spatial parameters

In this paper, the delay time τ and the embedded window width τ_w are estimated by C–C method. Because the delay time mainly depends on the embedded dimension and the embedded window width, the embedding dimension can be obtained according to $\tau_w = (m-1)\tau_d$. The time series $\{x(t)\}$ $(t = 1, 2, \cdots, n)$ is divided into M disjoint time series, the length is int(n/M), where int is the rounding function. For the general natural number M:

$$\begin{aligned} &\{x(1), x(M+1), x(2M+1), \ldots\} \\ &\{x(2), x(M+2), x(2M+2), \ldots\} \\ &\quad\vdots \\ &\{x(M), x(M+M), x(2M+M), \ldots\} \end{aligned} \tag{2}$$

Calculate the statistical component of each subsequence:

$$S(m,N,r,\tau) = \frac{1}{t}\sum_{l=1}^{t}\{C_1(m,N,r,\tau) - [C_1(m,N,r,\tau)]^m\} \tag{3}$$

In Eq. (3), C_l is the correlation integral of the Lth subsequence of the time series, defined as (4).

$$C_1(m,N,r,\tau) = \frac{1}{M(M-1)}\sum M_{i,j=1}\theta\left[r - \|X(i) - X(j)\|\right] \tag{4}$$

In Eq. (4), θ() represents the Heaviside unit function; $X(i)$ represents the ith time series. The local maximum interval can take the zero point of S()or the time point that is the smallest difference from the full radius r. Select the corresponding minimum radius and maximum radius r, the definition of the difference is as follows:

$$\Delta S(m,N,t) = \max[S(m,N,r_i,t)] - \min[S(m,N,r_i,t)]\, i \neq j \tag{5}$$

According to the statistical principle, the ranges of m are between 2 and 5, r is between $\sigma/2$ and 2σ, and σ is the mean square of the time series. Calculate the Eq. (6) as follows:

$$\begin{cases} \Delta\bar{S}(t) = \frac{1}{4}\sum_{m=2}^{5}\Delta\bar{S}(m,N,t) \\ \bar{S}(t) = \frac{1}{16}\sum_{j=2}^{4}\sum_{m=2}^{5}S(m,N,r_i,\tau) \\ S_{cor}(t) = \Delta\bar{S}(t) + |\bar{S}(t)| \end{cases} \tag{6}$$

In Eq. (6), $\Delta\bar{S}(t)$ is the first local maximum time point corresponding to the first minimum; $\bar{S}(t)$ is the mean of the statistic c of all subsequences $S(m,N,r_i,\tau)$; independent of the first overall maximum time window is $S_{cor}(t)$ minimum corresponding to the time series, known as the delay time window.

3. Adaptive prediction model based on Volterra filter

A large number of studies and experiments have proved that most of the nonlinear systems can be characterized by Volterra series. Volterra adaptive prediction model can accurately predict many chaotic sequences. This kind of prediction requires small amount of training data, easy to implement hardware and software. It can adaptively track the trajectory of chaotic system and has high prediction precision [10]. Therefore, this paper applies the Volterra adaptive filter forecasting model to predict the natural flow data of the three representative stations in the upper, middle and downstream of the Yellow River for the first time.

3.1. Volterra series expansion

Assuming that the input of the nonlinear discrete dynamical system is $X(n) = [x(n) + x(n-\tau), \ldots, x(n-(m-1)\tau)]$, the output based on one step is $y(n) = \hat{x}(n+\tau) = F(X(n))$. In this case, the second-order truncation error of the obtained nonlinear system is as follows[11]:

$$\hat{x}(n+1) = h_0 + \sum_{m_1=0}^{m-1}h_1(m_1)x(n-m_1\tau) + \sum_{(m_1=0)}^{(m-1)}\sum_{(m_1=m_2)}^{(m-1)}h_2(m_1,m_2)x(n-m_1\tau)x(n-m_2\tau) \tag{7}$$

Eq. (7) is represented by an finite impulse response (FIR) filter, that is, a Volterra adaptive filter. The filter coefficients and input vectors are as follows:

$$H(n) = [h_0, h_1(0), h_1(m_1-1), h_2(0,0), h_2(0,1), \cdots, h_2(m-1,m-1)]^T \tag{8}$$

$$U(n) = [1, x(n), x(n-\tau), x(n-(m-1)\tau), x^2(n), x(n)x(n-\tau), \cdots, x^2(n-(m-1)\tau)]^T \tag{9}$$

For high-order Volterra series filters, a sparse expansion of the predictive filter model [12] can be used to represent, specifically described as follows:

$$\begin{aligned}\hat{x}(n+1) = h_0 + \sum_{i=0}^{m-1} h_1(m_1)x(n-m_1\tau) + \\ \sum_{i=0}^{m-1} h_2(0,i)x^2(n-i\tau) + \sum_{i=0}^{m-1} h_2(1,i)x(n)x(n-i\tau) + \\ \sum_{i=0}^{m-1} h_3(0,i)x^3(n-i\tau) + \sum_{i=0}^{m-1} h_3(1,i)x^2(n)x(n-i\tau) + \\ \sum_{i=0}^{m-1} h_3(2,i)x\ (n)x^2(n-i\tau) + \cdots + \sum_{i=0}^{m-1} h_k(0,i)x^k(n-i\tau) + \\ \sum_{i=0}^{m-1} h_k(1,i)x^{(k-1)}(n)x(n-i\tau) + \sum_{i=0}^{m-1} h_k(2,i)x(n)x^{(k-1)} \\ (n-i\tau)\end{aligned} \tag{10}$$

At this point, the coefficient vector and the input vector are as follows:

$$H(n) = [h_0, h_1(1), \ldots, h_1(m-1), h_2(0,0), \ldots, h_2(0,m-1), h_2(1,1), \ldots, h_2(1,m-1), \ldots, h_3(0,0), \ldots, h_k(0,0), h_k(0,m-1), \ldots, h_k(1,1), \ldots, h_k(2,1), \ldots, h_k(2,m-1))]^T \tag{11}$$

$$U(n) = [1, x(n), x(n-\tau), \ldots, x(n-(m-1)\tau), x^2(n), \ldots, x^2(n-(m-1)\tau), x(n)x(n-\tau), \ldots, x(n)x(n-(m-1)\tau), x^3(n), \ldots, x^k(n), \ldots, x(n)x^{k-1}(n-\tau), x(n)x^{k-1}(n-(m-1)\tau)]^T \tag{12}$$

The above equation can be expressed as follows:

$$\hat{x}(n+1) = H^T(n)U(n) \tag{13}$$

4. Volterra adaptive filter prediction model

In a variety of adaptive algorithms for linear adaptive prediction models, time-domain orthogonal adaptive algorithms can be used directly in the adaptation of nonlinear models [12] to adjust the coefficient vectors. The nonlinear Volterra adaptive filter structure is shown in Fig. 1.

In this paper, the third-order Volterra adaptive filter is used to predict the natural flow time series of three stations, the forecasting step is divided into the following:

- The training sample is selected and the training sample data are reconstructed by phase space to obtain the input vector of the Volterra filter.
- Initialize the coefficient vector and use the time orthogonal adaptive algorithm to train the coefficient vector until it converges.
- The phase space included in the last time of the test point of the test sample is input to the trained Volterra filter to obtain the prediction result.

4.1. Basic information

The model was used to forecast and analyze the natural runoff sequences from 1919 to 1991 in the upper reaches of the Yellow River, Lanzhou station, middle reaches of Sanmenxia station and downstream Huayuankou station. The runoff data from the Yellow River Water Resources Commission Hydrology Bureau and the process of the runoff of the three stations are shown in Figs. 2 and 3.

5. Example application

5.1. Runoff prediction of the Yellow River

The third-order Volterra adaptive filter model is used to forecast the natural flow data of Lanzhou, Sanmenxia and Huayuankou stations in the Yellow River, the forecast is divided into monthly natural flow forecast and annual natural flow forecast. For the natural average flow forecast model, the total number of samples is 876, a total of 868 samples of runoff time series data from January 1919 to April 1991 were selected as learning training samples and were used as input to the model. A total of eight samples of runoff time series data from May 1991 to December 1991 were selected as the samples for the model. For the natural average flow forecast model, the total number of samples of the model is 73, and 8 years after the selection of data for the retention test samples, the previous 65 years of data are select as training samples. Considering the length is too long, only the graphs of forecasting and analyzing the annual runoff series of Lanzhou station in the upper reaches of the Yellow River are given, and the results are shown in Figs. 4 and 5.

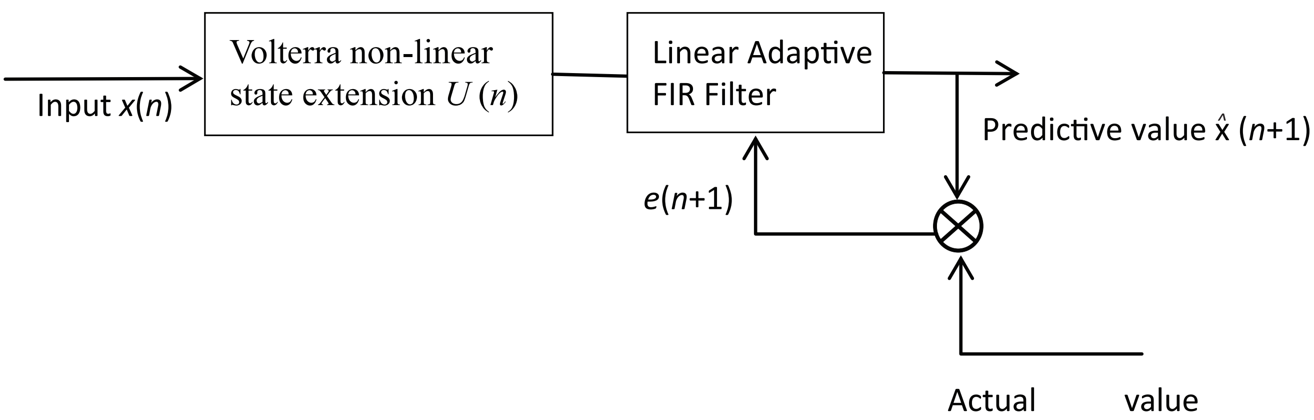

Fig. 1. Volterra adaptive filter structure.

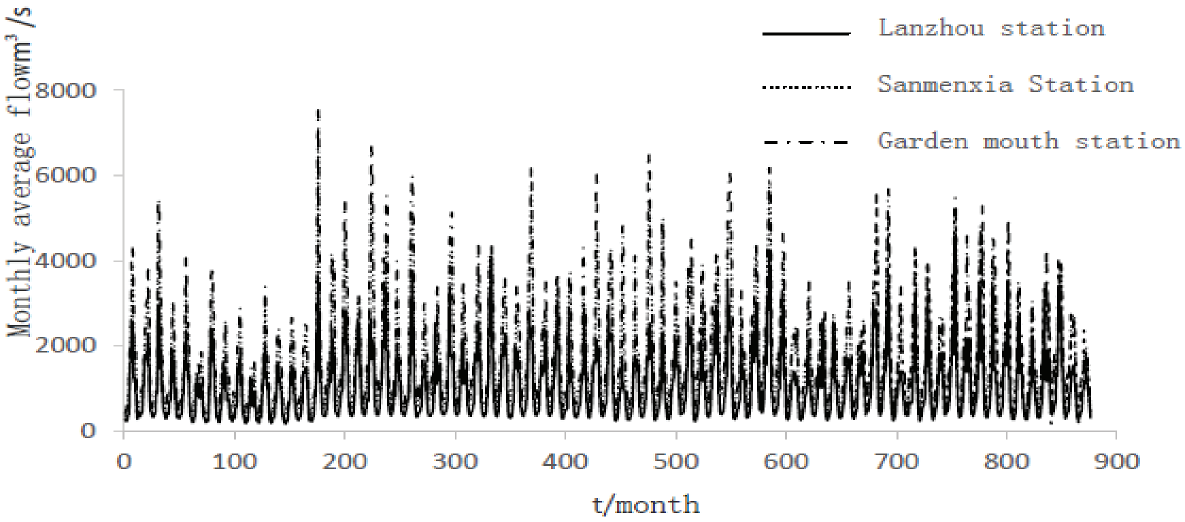

Fig. 2. Monthly average flow chart in three stations of the Yellow River.

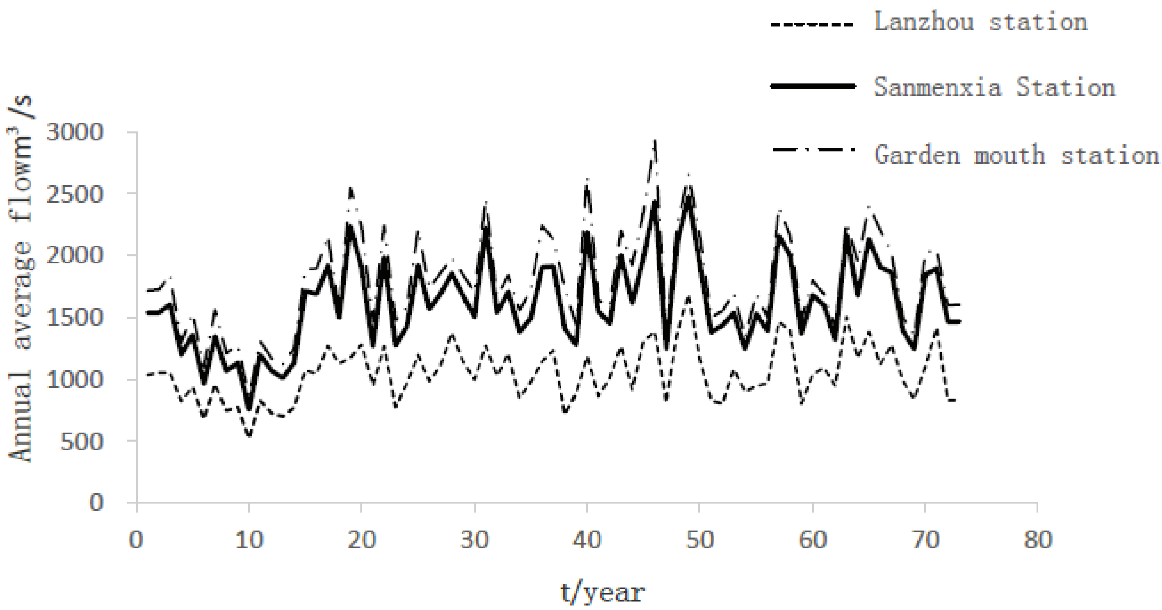

Fig. 3. Annual average flow chart in three stations of the Yellow River.

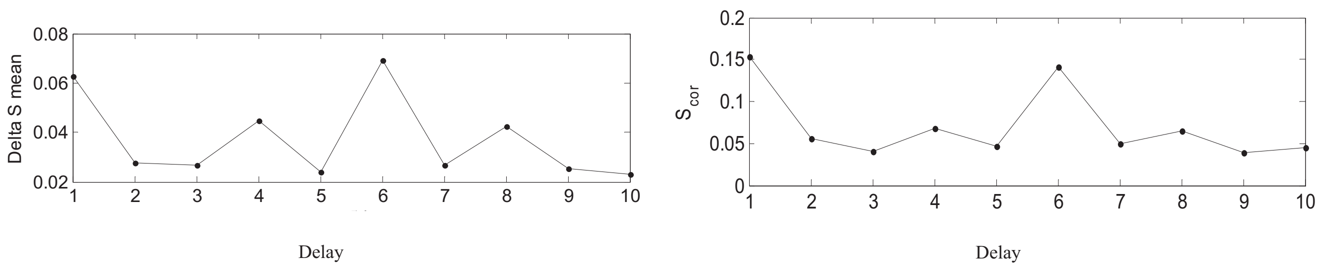

Fig. 4. Results the phase space of monthly average flow in Lanzhou station.

First, the C–C method is used to calculate the phase space reconstruction parameters of two time series, and the reconstruction results are shown in Figs. 4 and 5.

As can be seen from Figs. 4 and 5, The first minimum value of the monthly average flow data is 3, corresponding to the first local maximum time, that is, the optimal delay time τ = 3; And then by S_{cor} at τ = 9 to get the minimum, then corresponding to the time series independent of the first overall maximum time window, that is, the delay time window τ = 9.According to the embedded time window formula: $\tau_w = (m-1)\tau$, you can calculate the embedded dimension of 4; similarly, the optimal delay time of annual average flow is 4, and the embedding dimension is 2.

The Lyapunov–Wolf method is used to calculate the maximum Lyapunov index of the above data, and the number is 0.1286 and 0.3267, respectively. According to the chaos system theory, the maximum Lyapunov index obtained by the calculation is positive. It can be determined that both time series have chaotic properties. Then the third-order Volterra adaptive forecasting model is used to predict the above two sets of data, and the prediction results are shown in Figs. 6 and 7. Similarly, Sanmenxia station and Huayuankou station annual runoff forecast results are shown in Figs. 6 and 7.

5.2. Error analysis

"Hydrological information forecast" SL250-2000 provides 20% of the relative error between the predicted value, and the actual value is used as the permissible error. When the error of a forecast is less than the allowable error, it is a qualified forecast. The percentage of the ratio of the number of qualified forecasts to the total number of forecasts is the test pass rate, which reflects the overall accuracy of the forecast. The Volterra adaptive filter forecasting model is compared with the results of the annual runoff forecast of the Yellow River, and is shown in Table 1.

As can be seen from Table 1, for the monthly traffic forecast, the best forecast for the three sites is the Sanmenxia station, the relative error between the predicted value and the true value is the smallest, the minimum is 0.20%; for the average annual traffic forecast, the best prediction of the three sites is Lanzhou station, the relative error between the predicted value and the real value is the smallest, the minimum is 0.08%, the average error is 5.24%. The annual runoff

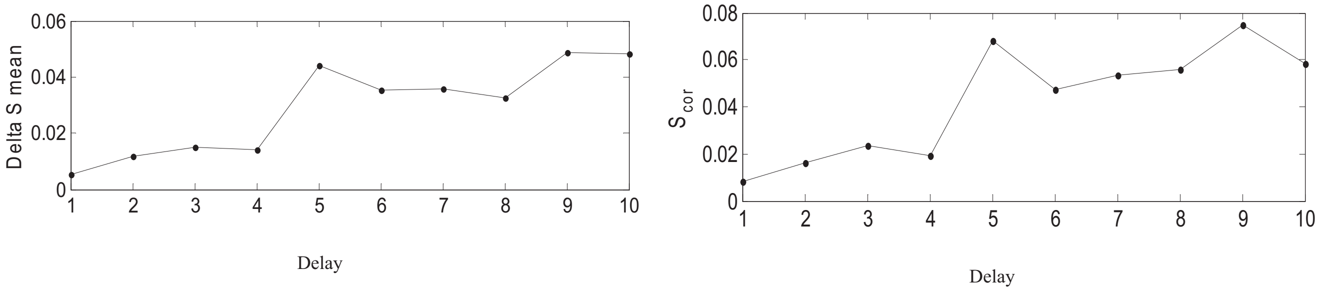

Fig. 5. Results the phase space of annual average flow in Lanzhou station.

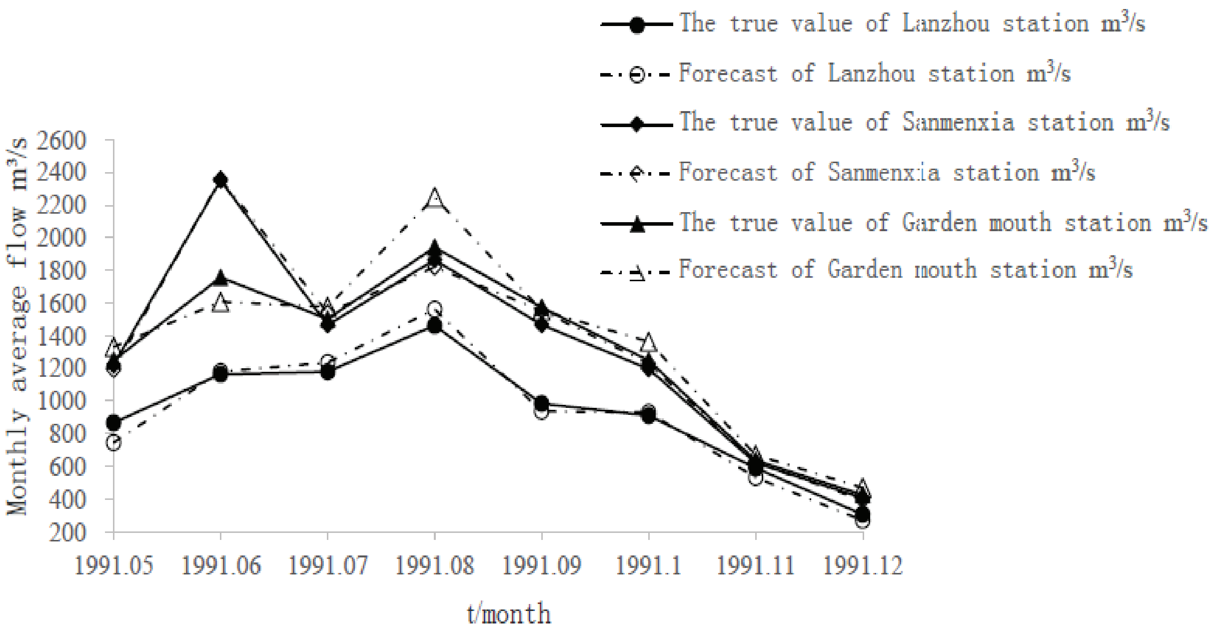

Fig. 6. The predicted results of monthly average flow in three stations of the Yellow River.

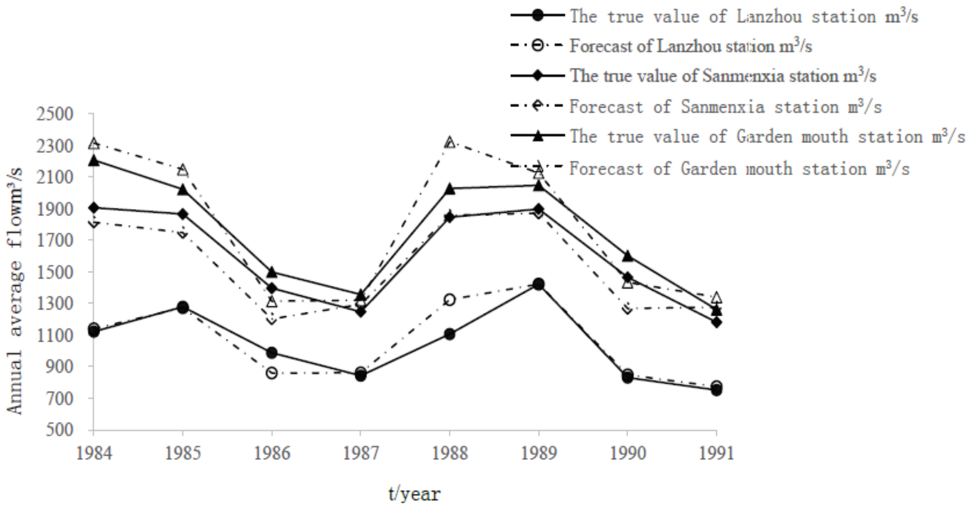

Fig. 7. The predicted results of annual average flow in three stations of the Yellow River.

Table 1
The predicted error analysis of the annual and monthly runoff in the Yellow River mainstream

Station	Monthly runoff forecast				Annual runoff forecast			
	Maximum relative error (%)	Minimum relative error (%)	Average error (%)	Inspection pass rate (%)	Maximum relative error (%)	Minimum relative error (%)	Average error (%)	Inspection pass rate (%)
Lanzhou station	14	1.48	6.98	100	19.66	0.08	5.24	100
Sanmenxia station	5.87	0.2	2.79	100	13.92	0.59	6.54	100
Garden mouth station	15.87	1.13	7.71	100	14.61	2.69	7.67	100

forecast of the three stations is qualified to meet the practical needs, indicating that the model can be used for the annual and annual runoff forecast of the Yellow River.

6. Conclusion

In this paper, the Volterra adaptive filter is introduced into the Yellow River runoff forecast, and the forecasting and analysis of the chaotic system is carried out by using the training data with small amount of training data, easy hardware, and software. The results show that the model has high consistency and good approximation ability, the prediction result is stable, easy to implement, and very fast. The model is suitable for the Yellow River runoff forecast, which enriches the river runoff forecasting method. But how to determine the optimal parameters of the Volterra adaptive filter and how to construct a more efficient, reliable, and general algorithm for calculating the Volterra filter parameters, so as to further improve the prediction accuracy and elongation of the Yellow River runoff forecast model. These problems still need to be studied in greater depth.

References

[1] Y.-Y. Meng, J.-P, Lu, Wind power chaos prediction based on Volterra adaptive filter, Power Syst. Prot. Control, 40 (2012) 91–96.
[2] Y.-H. Lei, C.-Y. Song, S.-Q. Ding, Nonlinear prediction of underwater reverberation based on a Volterra adaptive filter, J. Harbin Eng. Univ., 10 (2007) 1127–1130.
[3] W. Zhang, Q. Huang, C. Tong, Chaotic local-region multi-step forecasting model of flow time series and its application on the upper reach of the Yellow River, J. Hydroelec. Eng., 2007 11–15.
[4] X.I. Jiang, C. Liu, Y. Wang, H. Wang, Preliminary study on prediction of annual runoff change of Sanmenxia station in the Yellow River by using mixed regression model, Sci. China Ser. E: Technol. Sci., S1 (2004) 95–102.
[5] L.I. Yanbin, Research on chaotic characteristics and forecasting of the river runoff, Xi'an Univ. Technol., (2009) 45–57.
[6] Y. Li, Q. Huang, J. Xu, Z. Zhang, Support vector machine model of mid-long term prediction of runoff in stream, J. Hydroelec. Eng., 5 (2008) 28–32.
[7] X.Q. Zhang, J. Liang, Chaotic characteristic analysis and prediction model study on wind power time series, Acta Phys. Sinica, 19 (2012) 70–81.
[8] M. Han, Theory and Method of Chaotic Time Series Prediction, China Water Power Press, Beijing, (2007) 24–30.
[9] X.-Q. Zhang, J. Liang, Chaotic time series prediction model of wind power based on ensemble empirical mode decomposition-approximate entropy and reservoir, Acta Phys. Sinica, 5 (2013) 76–85.
[10] J.-S. Zhang, X.-C. Xiao, Predicting low-dimensional chaotic time, Acta Phys. Sinica, 49 (2000) 403–408.
[11] Z. Lu, Z. Cai, K. Jiang, Prediction of chaotic time series using singular value decomposition Volterra filter, J. Wuhan Univ. Technol., 4 (2007) 672–675.
[12] J.-S. Zhang, X.-C. Xiao, Prediction of chaotic time series by using adaptive higher-order nonlinear Fourier infrared filter, Acta Phys. Sinica, 49 (2000) 1222–1226.