Wavelet Network Model and Its Application to the Prediction of Hydrology

Wensheng Wang, Jing Ding

Department of Hydrology and Water Resources, Hydraulic School of Sichuan University, Chengdu, Sichuan 610065, China, wangws70@sina.com

Abstract: Based on the multi-time scale and the nonlinear characteristics of the observed time series, a new hybrid model between wavelet analysis and artificial neural network (ANN): wavelet network model, has been suggested. The present model absorbs some merits of wavelet transform and artificial neural network. Case studies, the short and long term prediction of hydrological time series, have been researched. The comparison results revealed that the suggested model could increase the forecasted accuracy and prolong the length time of prediction. The wavelet network model is satisfied. [Nature and Science 2003;1(1):67-71].

Keywords: wavelet analysis; artificial neural network; wavelet network model; prediction of hydrology

1. Introduction

The accuracy prediction of hydrology and water resource can give important information for the city planning, land use, the design of civil project and water resource management. Hydrology system is influenced by many factors, such as weather, land with vegetal cover, infiltration, evaporation and transpiration, so it includes the good deal of stochastic dependent component, multi-time scale and highly nonlinear characteristics. In general, the hydrology system is predicted with regressive analysis, stochastic theory models (Ding, 1988) and Grey model method (Deng, 1992). In recent years, artificial neural network (ANN), fuzzy theory and chaos theory have been widely applied in the sphere of hydrology and water resource. The studies have demonstrated these forecasted approaches are not very satisfied in precision because of only considering some aspects of its property. In order to raise the forecasted precision and lengthen the forecasted time, the hybrid model based on some methods should be probed. In this paper, a new hybrid model: wavelet network model, which is combined with wavelet analysis and ANN, has been proposed.

A wavelet network model makes use of the merits of wavelet analysis and ANN, so it has excellent performance in simulation and forecast. Some case studies are presented (in section 5) in which wavelet network model has been developed using the suggested methodology (in section 4) to forecast hydrological time series in China. The results show the technique and the model are feasible. The conclusions of the study are given in section 6.

2. Study Review

Wavelet analysis has become a research hot point. Wavelet analysis has good time and frequency multi-resolution, and can effectively diagnose signal's main frequency component and abstract local information of the time series. It has huge advances in signal processing, image compress and encoding, tongue encoding, mode identification and nonlinear science fields. The researches and applications of wavelet analysis have already begun in hydrology and water resources. The document (Li, 1997) points out the potential applications of wavelet analysis to hydrology and water resources. Li et al (1999) probe long time interval forecast of hydrological time series with combing neural network models based on wavelet transform. Wang et al (2000) have proposed a wavelet transform stochastic simulation model, which generate synthetic streamflow sequences that are statistically similar to observed streamflow sequences. The multi-time scale characteristics of hydrological variable have been studied by Wang et al (2002). Wavelet analysis will make a new research approach for the system of hydrology and water resources and broaden the content of hydrology greatly.

ANN is highly flexible function approximator that has self-learning and self-adaptive feature. Many studies attempted to model runoff by ANN. For example, Half et al (1993) designed a three-layer feed-forward ANN using the observed rainfall hyetographs as inputs and hydrographs as output to predict runoff from a watershed. Tokar (1999) reported that their ANN model had better prediction accuracy and flexibility than statistical regression and simple conceptual models. Applications of ANN are widely reported in the hydrological literature (French, 1992; Raman, 1995; Hu, 2001; Qing, 2002). ANN models have shown their utility in a broad rang of water resources application and are powerful tool for forecasting and prediction.

3. The Theory of Wavelet Analysis

3.1 Wavelet Transform

Wavelet analysis is multi-resolution analysis in time and frequency domain, and is the important milestone of the Fourier Transform. Wavelet function $\psi(t)$ is called mother wavelet, which has shock properties and can reduce zero rapidly. It can be defined as $\int_{-\infty}^{+\infty}\psi(t)dt = 0$ mathematically. $\psi_{a,b}(t)$ can be acquired through compressing and expanding $\psi(t)$:

$$\psi_{a,b}(t) = |a|^{-1/2}\psi(\frac{t-b}{a}) \quad b \in \mathrm{R}, a \in \mathrm{R}, a \neq 0 \quad (1)$$

Where $\psi_{a,b}(t)$ is successive wavelet; a is scale or frequency factor, b is time factor; R is the domain of real number.

If $\psi_{a,b}(t)$ satisfies equation (1), for the energy finite signal or time series $f(t) \in L^2(R)$, successive wavelet transform of $f(t)$ is defined as:

$$W_\psi f(a,b) =< f, \psi_{a,b} >= |a|^{-\frac{1}{2}} \int_R f(t)\overline{\psi}(\frac{t-b}{a})dt \quad (2)$$

Where $\overline{\psi}(t)$ is complex conjugate functions of ψ(t). Equation (2) describes that wavelet transform is the decomposition of $f(t)$ under different resolution level (scale). In other words, the essence of wavelet transform is to filter wave for $f(t)$ with different filter.

In real application successive wavelet is often discrete. Let $a = a_0^j$, $b = kb_0a_0^j$, $a_0>1$, $b_0 \in R$, k, j are integer number. Discrete wavelet transform of f(t) is written as:

$$W_\psi f(j,k) = a_0^{-j/2}\int_R f(t)\overline{\psi}(a_0^{-j}t - kb_0)dt \quad (3)$$

When a_0=2, b_0=1, equation (3) becomes binary wavelet transform:

$$W_\psi f(j,k) = 2^{-j/2}\int_R f(t)\overline{\psi}(2^{-j}t - k)dt \quad (4)$$

$W_\psi f(a,b)$ or $W_\psi f(j,k)$ can reflect the characteristics of original time series in frequency (a or j) and time domain (b or k) at the same time. When a or j is small, the frequency resolution of wavelet transform is low, but the time domain resolution is high. When a or j becomes large, the frequency resolution of wavelet transform is high, but the time domain resolution is low. That is, wavelet analysis is a mathematic microscope.

3.2 The Algorithm of Wavelet Transform

In real world observed time series are discrete, such as rainstorm process, flood process, monthly streamflow process, and daily runoff sequence. So discrete wavelet transform must be selected for decomposition and reconstruction of time series. There are many discrete wavelet transform algorithm, such as Mallat algorithm (Mallat, 1989; Mallat, 1989) and A Trous algorithm (Shensa, 1992; Aussum, 1997). A Trous algorithm has been adopted in the paper.

Let Z(t) (or C0(t)) denote the original discrete time series. A Trous decomposition algorithm as following:

$$C_i(t) = \sum_{\ell=-\infty}^{+\infty} h(\ell)C_{i-1}(t + 2^i\ell) \quad (i = 1,2,\cdots) \quad (5)$$

$$W_i(t) = C_{i-1}(t) - C_i(t) \quad (i = 1,2,\cdots) \quad (6)$$

Where $h(\ell)$ is the discrete low-pass filter; $C_i(t)$, $W_i(t)$ (i=1,2, ⋯) are scale coefficient (background information) and wavelet coefficient (detail information) at the resolution level *i* respectively. $W_1(t), W_2(t), \cdots, W_P(t)$ *and* $C_P(t)$ are called discrete wavelet transform with the resolution level *P*. In equation (5), extending of boundaries may be carried out in different ways. We took an intuitively acceptable approach by taking *C(n+k)=C(n-k)*.

The wavelet coefficients, $W_i(t)$ (i=1,2,⋯), provide the "detail" signal, which can capture small features of interpretational value in the data; the "residual" term $C_P(t)$ represents the data's "background" information. Because of simplicity of $W_1(t), W_2(t), \cdots, W_P(t), C_P(t)$ (we can see from section 4), some interesting characteristics, such as period, hidden period, dependence, jump, can be diagnosed easily through wavelet components $W_1(t), W_2(t), \cdots, W_P(t), C_P(t)$.

It is possible to reconstruct the original hydrological time series from wavelet components $\{W_1(t), W_2(t), \cdots, W_P(t), C_P(t)\}$. The wavelet reconstruction of the original time series, in term of wavelet coefficients, is given by

$$Z(t) = C_p(t) + \sum_{i=1}^{p} W_i(t) \quad (7)$$

Equation (7) provides a reconstruction formula for original time series. That is A Trous reconstructing algorithm. A Trous decomposition and reconstructing algorithm are simple and rapid. It's key is to determine the discrete low-pass filter.

4. Wavelet Network Model

4.1 Main Idea

First, original time series can be decomposed into a certain number of sub-time series $\{W_1, W_2, \cdots, W_P, C_P\}$ by wavelet transform algorithm. $W_1, W_2, \cdots, W_P$ are detail time series, and C_P is background time series. These play different role in the original time series and the behavior of each sub-time series is distinct. So the contribution to original time series varies from each other. Then, ANN is constructed in which the sub-time series at *t* time are input of ANN and the original time series at *t+T* time are output of ANN, where *T* is the time length of forecast. Last, the wavelet network model (WNM) is formed in which the weighs are learned with some algorithm. The key of wavelet network model is wavelet decomposition of time series and the construction of ANN.

4.2 Decomposition of Observed Time Series

The low-pass filter h, which is a B_3 spline, defined as $(\frac{1}{16},\frac{1}{4},\frac{3}{8},\frac{1}{4},\frac{1}{16})$, is used. This is of compact support and point-symmetric. First the resolution level P must be determined. In general there is $INT(\ell g\, n)$ resolution scale number, where n is the length of time series and INT stands for integer number, the $\ell g\, n$ is common logarithm.

The wavelet coefficients and scale coefficients of the monthly groundwater level time series derived from A Trous decomposition algorithm are shown in Figure 1. In Figure 1 $W_1(t)$ and $W_2(t)$ denote wavelet coefficients at the resolution level 1 and 2 respectively; $C_2(t)$ denotes scale coefficients at resolution level 2.

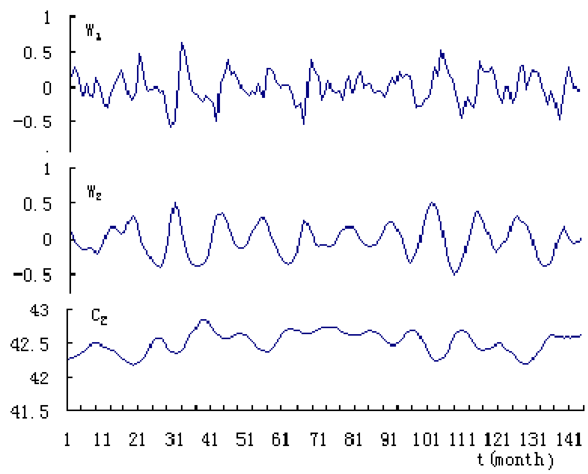

Figure 1. Wavelet Decomposed Process of Monthly Groundwater Level Time Series

4.3 The Structure of ANN

ANN is widely applied in the forecasting of hydrology and water resource. In ANN, BP network models are common to engineer. So called BP network models, that is the feed-forward artificial neural network structure and a back-propagation algorithm (BP). It has proved that BP network model with three-layer is satisfied for the forecasting and simulating in the science of water.

The input of BP network is $X=[W_1(t), W_2(t), \cdots, W_P(t), C_P(t)]^T$, and the nodes of input layer are P+1. The output is $Y=[L(t+T)]$, and the node is 1. The nodes of hidden layer are determined by trial and error. The network weights are learned by standard BP algorithm or self-adapted BP algorithm and so on. The details on BP algorithm are available in the references, hence these are not repeated here.

5. Cases Studies

5.1 Case One: Shallow Groundwater Level Forecast

There is 12-year (1983-1994) record of shallow monthly groundwater level in Beijing of China from the literatures (Lu, 1997), that is $\{Z(t), t=1,2,\ldots,144\}$. The first nine years time series are used for calibration /training of the model, and the remaining three years data are used for verification or testing purposes. In this research $n = 9\times 12$, then the scale number P=2.

Through A Trous algorithm, the groundwater level time series are decomposed into the sub-time series: $\{W_1(t), W_2(t), C_2(t)\}$, and are listed in Figure 1. Here three-layer network: input layer, hidden layer and output layer, is adopted. The number of nodes in hidden layer is equal to 3. So the structure of WNM is 3-3-1. The weight parameters of network are estimated by self-adapted BP algorithm. The number of training of WNM is 5000.

Given four forecasting periods (T=1 month, 2 month, 3 month, and 4 month), the fitting and forecasting results of groundwater level are shown in Table 1. The forecasted results of groundwater level of 1992~1994 are shown in Figure 2 (T=1 month) and Figure 3 (T=3 month). In order to compare, the results of calibration and verification of ARMA model (Lu, 1997) and threshold autoregressive model (TAR) based on genetic algorithm (Jing, 2000) are listed in Table 1. From Table 1 it can be seen that, for T=1 month, the calibration precision of WNM is almost good as ARMA and TAR, but the verification precision is better than the latter. At the same time, when the forecasting period (Table 1, Figure 3, Figure 4) is become long, the fitting and testing precision of WNM is also very higher than the other models (not listed).

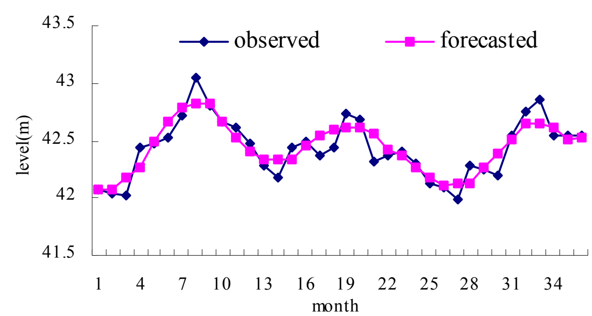

Figure 2. Comparison of the Original Groundwater Level Time Series and Forecasted Time Series of WNM (Where T=1 Month)

5.2 Case Two: Daily Discharge Forecast

Daily discharge data for Yangtze River basin of China at Cuntan Station are used. The years 1992-2001 are selected, the first 8 years data are used for build the wavelet network model, and the remaining two years data are used for verification of WNM. Here $n = 8\times 365$, then scale number P=3.

Table 1. Error Analysis of Validation and Verification of WNM

Model	Percent of calibration absolute error falling into the demarcation				Percent of verification absolute error falling into the demarcation			
	[0,0.1]	[0,0.2]	[0,0.3]	[0,0.4]	[0,0.1]	[0,0.2]	[0,0.3]	[0,0.4]
ARMA (T=1)	79.0	99.0	100	100	30.0	59.0	81.0	92.0
TAR (T=1)	51.0	82.3	94.8	96.9	47.2	80.6	91.7	100
WNM (T=1)	57.4	87.0	99.1	100	60.0	91.4	100	100
WNM (T=2)	73.2	98.2	100	100	73.5	100	100	100
WNM (T=3)	60.2	91.7	98.2	100	51.5	90.9	100	100
WNM (T=4)	52.8	83.3	94.4	98.2	43.8	71.8	93.8	100

Table 2. The Calibrated and Validated Results of Wavelet Network Model (%)

Model	Percent of validation			Percent of verification		
	<10%	<20%	<30%	<10%	<20%	<30%
TAR(T=1d)	83.9	95.9	98.7	84.9	95.2	98.2
TAR(T=2d)	63.0	86.3	94.7	64.3	86.6	94.0
TAR(T=3d)	52.2	77.5	89.9	52.9	78.0	89.3
TAR(T=4d)	45.4	72.4	86.6	45.8	72.9	84.8
TAR(T=54d)	40.5	67.6	83.5	42.3	68.0	82.3
WNM(T=1d)	76.4	96.1	99.7	89.4	97.9	99.6
WNM(T=2d)	83.3	97.5	99.9	95.2	99.5	99.9
WNM(T=3d)	77.8	97.2	99.9	87.1	99.0	99.9
WNM(T=4d)	66.2	93.1	98.7	78.2	94.1	98.9
WNM(T=5d)	69.1	93.6	98.7	80.6	95.3	98.5

Through A Trous algorithm, the daily discharge time series are decomposed into the sub-time series: $\{W_1(t), W_2(t), W_3(t), C_3(t)\}$. Three layers network is adopted too. The number of nodes in hidden layer is equal to 4. So the structure of WNM is 4-4-1. The weight parameters of network are estimated by modified BP algorithm. The number of training of WNM is 2000.

Given five forecasting periods (T=1 day, 2 day, 3 day, 4 day and 5 day), the fitting and forecasting results of daily discharge are shown in Table 2. The forecasted results of daily discharge of 2001 year are shown in Figure 4 (T=1 day) and Figure 5 (T=3 day). In order to compare, the results of validation and verification of TAR model are listed in Table 2. It was noticed that WNM is better than TAR model.

6. Conclusions

This paper has reported a new hybrid model wavelet network model. It plays an important role in improving the precision and prolonging the forecasting time period or hydrology and water resource time series.

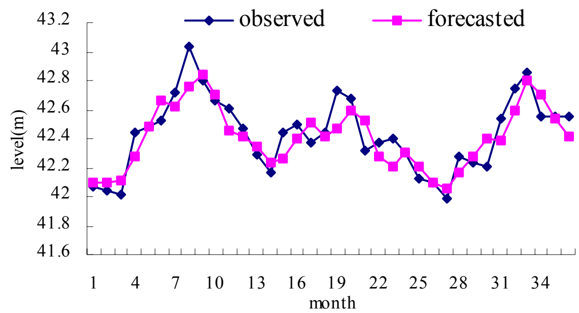

Figure 3. Comparison of the Original Groundwater Level Time Series and Forecasted Time Series of WNM (Where *T*=3 Month)

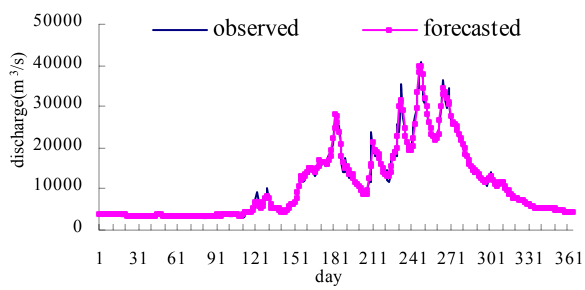

Figure 4. Comparison of the Original Daily Discharge Time Series and Forecasted Time Series of WNM (Where T=1 Day)

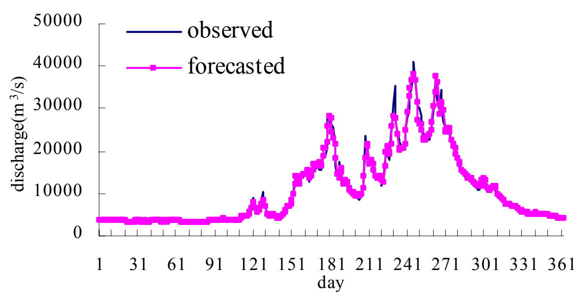

Figure 5. Comparison of the Original Daily Discharge Time Series and Forecasted Time Series of WNM (Where *T*=3 day)

Calibration and verification of wavelet network model for prediction of hydrology and water resource in case studies have shown that the method is functional. The suggested strategy is suit to any other water resource time series.

Elementary attempt at developing the hybrid model is success. Future studies will be opened up from the manner of recombined with wavelet analysis and ANN to applications of wavelet network model.

We would like to thank the National Science Foundation Committee of China to support this research project (No: 50279023).

Correspondence to:
Wensheng Wang
Hydraulic School of Sichuan University
Chengdu, Sichuan 610065, China
Telephone: 001186-28-8540-3347
E-mail: **wangws70@sina.com**

References

Aussum A, Campbell J, Murfagh F. Wavelet-based feature extraction and decomposition strategies for financial forecasting. Journal of Computational Intelligence in France 1997:1-17.

Chui CK, ed. Wavelet—A Tutorial in Theory and Applications. Xian Jiaotong University Press, Xian, China, 1995:66-287 (in Chinese).

Deng J, ed. Grey Prediction and Grey Decision. Huazhong Ligong University Press, Wuhan, China, 1992:74-192 (in Chinese).

Ding J, Deng Y, ed. Stochastic Hydrology. Chengdu University of science and technology Press, Chengdu, China, 1988:190-230 (in Chinese).

French MN, Krajewski WF, Cuykendall RR. Rainfall forecasting in space and time using a neural network. J Hydrol 1992;137:1-31.

Half AH, Halff HM, Azmoodeh M. Predicting runoff from rainfall using artificial neural networks. Proc Engng Hydrol 1993; ASCE:761-5.

Hu T, Lam KC, Ng ST. River flow time series prediction with a range dependent neural network. Hydrol Sci J 2001;46(5):729-45.

Jing J, Ding J, ed. Genetic Algorithm and Its Application to Water Science. Sichuan University Press, Chengdu, China, 2000:25-160 (in Chinese).

Li X, Ding J, Li H. Combing neural network models based on Wavelet transform. Journal of Hydraulic 1999; 2:1-4 (in Chinese).

Li X, Ding J, Li H. Wavelet analysis and its potential application to hydrology and water resources. Journal of Sichuan Union University (Engineering Science) 1997;1(4): 49-52 (in Chinese).

Liu G, Ding J. Groundwater level forecast with Bp network model. Journal of Xian Geology College, 1997;2:45-50 (in Chinese).

Lu H. Forecast of shallow groundwater level in Beijing. Geotechnical Investigation & Surveying, 1997;1:67-71.

Mallat S G. Multifrequency channel decomposition of images and wavelet models. IEEE Trans on ASSp 1989;37(12):2091-110.

Mallat SG. A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans on PAMI 1989;11(7):674-93.

Qing G, Ding J, Liu G. Self-adapted BP algorithm and its application to flood forecast of river. Advance in Water Science 2002;13(1):37-41(in Chinese).

Raman H, Sunilkumar N. Multivariate modeling of water resource time series using artificial neural network [J]. Hydrol Sci J 1995;40(21):145-63.

Shensa MJ. Discrete wavelet transform: wedding the A Trous and Mallat algorithm [J]. IEEE Transactions on Signal Processing 1992;40:2464-82.

Tokar AS, Johnson PA. Rain-runoff modeling using artificial neural network. J Hydrol Engng ASCE 1999;4(3):232-9.

Wang W, Ding J, Xiang H. The multi-time scale analysis of hydrological time series with wavelet transform. Journal of Sichuan University 2002;35(4):14-7 (in Chinese).

Wang W, Yuan P, Ding J. Wavelet analysis and its application to the stochastic simulation of daily discharge process. Journal of Hydraulics 2000;11:43-8 (in Chinese).

Zhang L, ed. Artificial neural network models and its application. Fudan University Press, Shanghai, China. 1994:10-65 (in Chinese).