Journal of Southeast University (English Edition) Vol. 24, No. 1, pp. 90 – 94 Mar. 2008 ISSN 1003—7985

River channel flood forecasting method of coupling wavelet neural network with autoregressive model

Li Zhijia Zhou Yi Ma Zhenkun

(College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China)

Abstract: Based on analyzing the limitations of the commonly used back-propagation neural network (BPNN), a wavelet neural network (WNN) is adopted as the nonlinear river channel flood forecasting method replacing the BPNN. The WNN has the characteristics of fast convergence and improved capability of nonlinear approximation. For the purpose of adapting the time-varying characteristics of flood routing, the WNN is coupled with an AR real-time correction model. The AR model is utilized to calculate the forecast error. The coefficients of the AR real-time correction model are dynamically updated by an adaptive fading factor recursive least square(RLS) method. The application of the flood forecasting method in the cross section of Xijiang River at Gaoyao shows its effectiveness.

Key words: river channel flood forecasting; wavelet neural network; autoregressive model; recursive least square (RLS); adaptive fading factor

River flood forecasting is a complicated nonlinear problem. Traditionally, river flood forecasting has been mainly handled by hydrodynamic methods. These methods utilize Saint-Venant equations to describe the movement characteristics of flood waves. However, the difficulty in obtaining accurate channel topography data restricts the pratical application of the methods[1].

In recent years, artificial neural networks (ANN) are increasingly utilized in the field of river flood forecasting, and the results are satisfactory. The ANN is a kind of input-output or "black-box" model. The attention of the ANN is concentrated on identifying the relationship between the input and the output of a system without trying to describe the internal mechanism of the system. For river channel flood forecasting, a back-propagation neural network (BPNN) is the most widely used ANN model. BPNN is a sort of feed-forward neural network(FNN). BP is the learning algorithm of a BPNN. So far, most BPNNs used in the field of flood forecasting are of a three-layer structure: one input layer, one hidden layer and one output layer. It has been proved that a BPNN with just one hidden layer can approximate an arbitrary nonlinearity[2]. Nevertheless, the learning problems of slowness in convergence and a tendency to fall into local minimums hinder the practical operation of BPNNs[3-4]. For the purpose of overcoming those defects of the BPNN, a wavelet neural network(WNN) is adopted in this paper. The WNN uses the nonlinear wavelet basis function as the transfer function of hidden nodes instead of common sigmoid functions. The wavelet transform has the advantage of time-frequency localization capacity, and the BPNN has the powerful ability of self-learning. The combination of the two mathematical methods forms an outstanding tool for nonlinear approximation[3,5]. The contrast between the WNN and the BPNN in this paper on learning progress demonstrates the superiority of WNNs in convergence rate and learning accuracy.

Despite the improvements of the WNN to the BPNN, the WNN only implements a static mapping. That is, it is still restricted to time-invariant problems. However, river channel flood routing is a highly complicated time-varying problem influenced by many random factors. In order to improve the practical forecast effects, a combination of the AR model and the adaptive fading factor RLS method is utilized to predict the forecast errors of the WNN. The revised forecast value is the sum of the output of the WNN and the forecast error. The proposed method is applied in the cross section of Xijiang River at Gaoyao. In comparison with the BPNN and the single WNN, this method shows the improved forecast accuracy, and it is effective in practice.

1 WNN Used in River Flood Forecasting

1.1 Basic theory of WNN

A square integrable function $\psi(t) \in L^2(\mathbf{R})$ is called a "mother wavelet" if it satisfies the following condition:

$$\int |\psi(t)|^2 \partial t < \infty \tag{1}$$

which ensures that $\psi(t)$ has finite energy, and

$$C_\psi = \int_{\mathbf{R}^*} |\psi(\omega)|^2 |\omega|^{-1} \partial\omega < \infty \tag{2}$$

where $\mathbf{R}^*$ denotes the set of nonzero real numbers and $\psi(\omega)$ is the Fourier transform of $\psi(t)$. Eq. (2) is referred to as the admissibility condition.

From Eq. (1) and Eq. (2) we can see that $\psi(t)$ is an oscillating, damped function. The mother wavelet $\psi(t)$ is a prototype for generating the other wavelet basis functions. They can be expressed as

$$\left\{\psi_{a,b}(t) \mid \psi_{a,b}(t) = |a|^{-\frac{1}{2}}\psi\left(\frac{t-b}{a}\right)\right\} \qquad a\in\mathbf{R}^*, b\in\mathbf{R} \tag{3}$$

where a and b are the scale and translation parameters, respectively; $|a|^{-\frac{1}{2}}$ is the normalization factor.

The continuous wavelet transform of a time function $f(t)$

Received 2007-08-19.

Biography: Li Zhijia(1962—), male, doctor, professor, zhijia-li@vip.sina.com.

Foundation item: The National Natural Science Foundation of China(No. 50479017).

Citation: Li Zhijia, Zhou Yi, Ma Zhenkun. River channel flood forecasting method of coupling wavelet neural network with autoregressive model[J]. Journal of Southeast University (English Edition), 2008, 24(1): 90 – 94.

$\in L^2(\mathbf{R})$ is defined as

$$W_{\mathrm{f}}(a, b) = \langle f, \psi_{a,b} \rangle = |a|^{-\frac{1}{2}} \int_{\mathbf{R}} f(t) \bar{\psi}\left(\frac{t-b}{a}\right) \partial t \quad (4)$$

where $W_{\mathrm{f}}(a, b)$ denotes the wavelet coefficient; $\langle\ \rangle$ is the inner product; $\bar{\psi}\left(\frac{t-b}{a}\right)$ is the complex conjugation of $\psi\left(\frac{t-b}{a}\right)$.

According to the Parseval identity, the inversion formula of Eq. (4) is

$$f(t) = \frac{1}{C_{\psi}} \int_{\mathbf{R}^*} \int_{\mathbf{R}} W_{\mathrm{f}}(a, b) \psi\left(\frac{t-b}{a}\right) \frac{\partial a \partial b}{a^2} \quad (5)$$

which means that the function $f(t)$ can be reconstructed.

On the basis of Eq. (5), a certain function $f(x_1, x_2, \ldots, x_M) \in L^2(\mathbf{R}^M)$ can be approximated by a family of locally supported wavelet basis functions:

$$\hat{f}(x_1, x_2, \ldots, x_M) = \sum_{j=1}^{N} w_j \psi\left[\frac{\sum_{i=1}^{M} w_{ij} x_i - b_j}{a_j}\right] + \bar{f} \quad (6)$$

where $\hat{f}(x_1, x_2, \ldots, x_M)$ is the approximated function, w_j and w_{ij} are the weight coefficients, N denotes the number of wavelet basis functions, and $\bar{f}$ is introduced to treat the nonzero mean function since the wavelet function has zero mean value[6].

In fact, Eq. (6) represents the decomposition form of $f(x_1, x_2, \ldots, x_M)$; namely, $f(x_1, x_2, \ldots, x_M)$ can be approached by the weighted sum of a series of wavelet basis functions. This equation can be realized by a special three-layer FNN: the WNN. Fig. 1 gives the structure of the WNN, where C_0 is the threshold of the output layer's node.

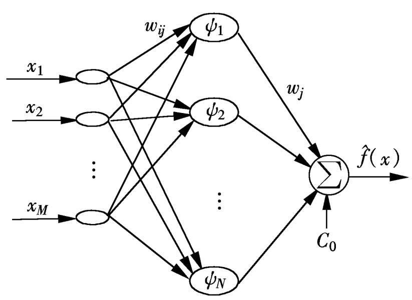

Fig. 1 Structure of the WNN

The structure of the WNN is similar to that of the radical basis function (RBF) network. It is better than the RBF because the newly introduced parameters, translation and scale factors, are adaptively rectified according to the local characteristics of the input signal in the training process of the WNN while the center and width parameters of the RBF are mainly selected by experience. So the WNN have more degrees of freedom and the enhanced ability of function approaching[4].

1.2 Practical structure and learning algorithm of the WNN

The water level of a cross section can be considered as the output of a nonlinear multi-input, single-output system[7]:

$$h(t) = F[h(t-1), \ldots, h(t-n_1), h_{\mathrm{u}}(t), \ldots, h_{\mathrm{u}}(t-n_2), h_{\mathrm{d}}(t), \ldots, h_{\mathrm{d}}(t-n_3)] \quad (7)$$

where $h(t)$ is the water level of a certain cross section; $h_{\mathrm{u}}(t-i)(i=0, 1, \ldots, n_2)$ and $h_{\mathrm{d}}(t-j)(j=0, 1, \ldots, n_3)$ are the water level of the upstream cross section and downstream cross section; n_1, n_2 and n_3 are the AR orders of h, h_{u} and h_{d}, respectively; $F(\cdot)$ denotes an unknown nonlinear function.

The nonlinear function $F(\cdot)$ can be effectively identified by utilizing the WNN illustrated in Fig. 1. The inputs of the WNN are $h(t-i)(i=1, 2, \ldots, n_1)$, $h_{\mathrm{u}}(t-i)(i=0, 1, \ldots, n_2)$ and $h_{\mathrm{d}}(t-j)(j=0, 1, \ldots, n_3)$. The output is the approximated water level $\hat{h}(t)$.

The widely used Morlet mother wavelet is selected as the transfer function of the WNN's hidden nodes, which has the following expression:

$$\psi(t) = \cos(1.75t) \exp\left(-\frac{t^2}{2}\right) \quad (8)$$

The parameters of the WNN need to be initialized before learning. The connection weights and thresholds are initialized by random values, while the translation and scale factors should be initialized according to wavelet theory[8]. Based on wavelet theory, when the input nodes are greater than 1, the concentration region in the time domain of the series of a mother wavelet's expansion is

$$[b_j + a_j t^* - a_j \Delta\Psi, b_j + a_j t^* + a_j \Delta\Psi] \quad (9)$$

where t^* is the center of the mother wavelet, and $\Delta\Psi$ is the radius.

For the purpose of covering the input vector by the series of a mother wavelet's expansion, the initial values of b_j and a_j must meet the following relationships:

$$\left.\begin{aligned} b_j + a_j t^* - a_j \Delta\Psi &\leqslant \sum_{i=1}^{M} w_{ij} x_{i\min} \\ b_j + a_j t^* + a_j \Delta\Psi &\geqslant \sum_{i=1}^{M} w_{ij} x_{i\max} \end{aligned}\right\} \quad (10)$$

where $x_{i\min}$ and $x_{i\max}$ denote the minimum and maximum input values of the i-th node in the input layer.

On the basis of Eq. (10), the recommended initial values of b_j and a_j are

$$\left.\begin{aligned} a_j &= \frac{\sum_{i=1}^{M} w_{ij} x_{i\max} - \sum_{i=1}^{M} w_{ij} x_{i\min}}{2\Delta\Psi} \\ b_j &= \frac{\sum_{i=1}^{M} w_{ij} x_{i\min}(t^* + \Delta\Psi) - \sum_{i=1}^{M} w_{ij} x_{i\max}(t^* - \Delta\Psi)}{2\Delta\Psi} \end{aligned}\right\} \quad (11)$$

2 AR Real-Time Correction Model

2.1 Basic theory

The AR model is the most widely utilized real-time correction method due to its simplicity in theory and its favourable effects. An AR real-time correction model is also selected in this paper based on the following two considerations: On the one hand, the complex input-output relationship of the WNN makes it difficult to correct the parameters directly; on the other hand, updating the WNN output can directly take both the impacts of the inaccurate parameters and the errors existing in the input-output data into account.

The forecast error sequence is represented by a p-order AR model:

$$e(t) = \theta_1 e(t-1) + \theta_2 e(t-2) + \ldots + \theta_p e(t-p) + \xi(t) \tag{12}$$

where $\xi(t)$ is the model error at time t, which is regarded as a white noise; $e(t)$ is the forecast error at time t; $\theta_i (i=1,2, \ldots, p)$ denotes the model coefficient.

According to Eq. (12), the predicted forecast error at time step t is

$$\hat{e}(t) = \theta_1 e(t-1) + \theta_2 e(t-2) + \ldots + \theta_p e(t-p) \tag{13}$$

where $\hat{e}(t)$ is the predicted error.

Suppose that $\hat{h}(t)$ is the calculated output of the WNN at time step t, then the updated value $\hat{h}_c(t)$ can be expressed as

$$\hat{h}_c(t) = \hat{h}(t) + \hat{e}(t) \tag{14}$$

The coefficients of the AR model can be estimated by online recursive estimation methods such as the RLS method.

2.2 Adaptive fading factor RLS method

Rewrite Eq. (14) in a vector form:

$$\hat{h}_c(t) = \hat{h}(t) + \boldsymbol{X}^{\mathrm{T}}(t)\boldsymbol{\theta}(t) \tag{15}$$

where $\boldsymbol{X}(t) = \{e(t-1), \ldots, e(t-p)\}^{\mathrm{T}}$, $\boldsymbol{\theta}(t) = \{\theta_1(t), \ldots, \theta_p(t)\}^{\mathrm{T}}$.

The performance index of the fading factor RLS is

$$J(\boldsymbol{\theta}) = \sum_{t=1}^{T} \mu^{(T-t)} [\xi(t)]^2 = \sum_{t=1}^{T} \mu^{(T-t)} [e(t) - \boldsymbol{X}^{\mathrm{T}}(t)\boldsymbol{\theta}(t)]^2 \tag{16}$$

where $\mu \in (0, 1]$ is the fading factor, and $e(t)$ is the actual forecast error.

The idea of the fading factor RLS is to revise the former estimated coefficient vector by the new data $e(t)$ on the basis of minimizing the performance index $J(\boldsymbol{\theta})$. In order to overcome the data saturation phenomenon, the weight $\mu^{(T-t)}$ is adopted in the exponential weighting form to strengthen the effects of new data artificially.

A fixed fading factor is adequate for a stationary changing time-variant system. But it cannot track the dynamic changing characteristics of flood movements efficiently. The fading factor should be adaptively adjusted for such problems.

According to these ideas, the recursive formula of the adaptive fading factor RLS[9] is expressed as

Error forecasting

$$\hat{e}(t) = \boldsymbol{X}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1) \tag{17}$$

Calculation of gain

$$\boldsymbol{K}(t) = \boldsymbol{P}(t-1)\boldsymbol{X}(t)[\mu(t-1) + \boldsymbol{X}^{\mathrm{T}}(t)\boldsymbol{P}(t-1)\boldsymbol{X}(t)]^{-1} \tag{18}$$

Coefficient vector estimation

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \boldsymbol{K}(t)[e(t) - \boldsymbol{X}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)] \tag{19}$$

Calculation of fading factor

$$\mu(t) = 1 - \frac{[1 - \boldsymbol{X}^{\mathrm{T}}(t)\boldsymbol{K}(t)][e(t) - \hat{e}(t)]^2}{\sum_0} \tag{20}$$

Calculation of covariance matrix

$$\boldsymbol{P}(t) = \frac{[1 - \boldsymbol{K}(t)\boldsymbol{X}^{\mathrm{T}}(t)]\boldsymbol{P}(t-1)}{\mu(t)} \tag{21}$$

3 Example Application and Analysis

The method presented in this paper is applied in the flood forecasting in the cross section of Xijiang River at Gaoyao in order to verify its effectiveness. Xijiang River is the largest river in the Pearl River Basin. The Gaoyao hydrological station is an important control station in the lower reaches of Xijiang River. The hydrological data obtained from it are significant for the flood prevention of the lower reaches of the Pearl River Basin. The water level of Gaoyao has a close connection with the water levels of Wuzhou in the upstream and Makou in the downstream. This connection can be established by the WNN constructed in this paper. According to Eq. (7), the input vector of the WNN is

$$\boldsymbol{h} = \{h_{\mathrm{G}}(t-1), \ldots, h_{\mathrm{G}}(t-n_1), h_{\mathrm{W}}(t), \ldots, h_{\mathrm{W}}(t-n_2), h_{\mathrm{M}}(t), \ldots, h_{\mathrm{M}}(t-n_3)\}^{\mathrm{T}} \tag{22}$$

where h_{G}, h_{W} and h_{M} denote the water levels of Gaoyao, Wuzhou and Makou, respectively; n_1, n_2 and n_3 are the AR orders of h_{G}, h_{W} and h_{M}, respectively.

By using the AIC criteria[10], n_1, n_2 and n_3 are determined to be 3, 1 and 1.

The water level data of ten representative floods from 1968 to 1997 are selected. The anterior eight floods are used for the learning of the WNN and the remaining two floods are for verification. A normal BPNN of the same input-output structure is generated for comparison.

Three evaluation indices are chosen. They are the relative error of the flood peak level, the absolute error of the flood peak level and the mean absolute error of the water level during the flood process.

The comparisons of the results are shown in Tab. 1 and Figs. 2 to 4.

Tab. 1 Comparison of the calculated characteristic values of the three methods

Serial number of the verifying floods	River flood forecasting method	Evaluation index		
		Relative error of the flood peak level/%	Absolute error of the flood peak level/m	Mean absolute error of the water level during the flood process/m
1997062210	BPNN	−6.10	0.36	0.13
	WNN	−5.25	0.31	0.11
	WNN plus AR model	0.1	0.01	0.01
1997080720	BPNN	−5.69	0.55	0.17
	WNN	−4.86	0.47	0.14
	WNN plus AR model	0	0	0.01

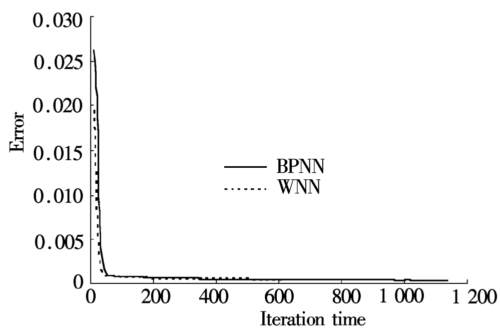

Fig. 2 Comparison of the learning processes of BPNN and WNN

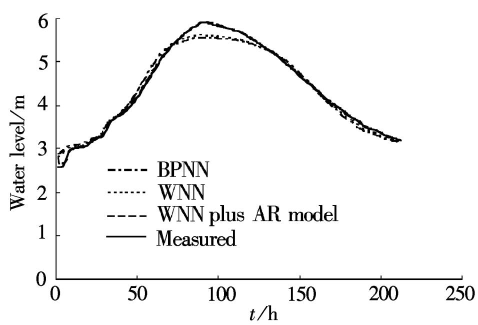

Fig. 3 Comparison of the 1997062210th flood's calculated and measured water level processes

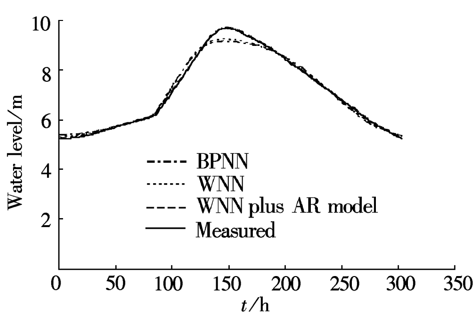

Fig. 4 Comparison of the 1997080720th flood's calculated and measured water level processes

The comparison of the learning process in Fig. 2 demonstrates the evidently improved learning ability of the WNN. The calculated characteristic values in Tab. 1 show that the river flood forecast accuracies of the WNN and the WNN with the AR real-time correction model are improved to different degrees compared with the BPNN. Thereinto, the WNN with the AR real-time correction model obtains the best forecast results. Figs. 3 and 4 also show that the forecast water level processes by using the WNN with the AR real-time correction model are closer to the measured water level processes than the forecast water level processes by using the WNN and the BPNN, especially in the periods of flood peaks.

4 Conclusion

In this paper, the river flood forecasting method of coupling WNN with AR real-time correction is proposed and used in the flood forecasting in the cross section of Xijiang River at Gaoyao. In comparison with the normal BPNN, the WNN shows the ability of accelerated convergence speed in the learning period and enhanced learning accuracy. In order to adapt the time-varying flood routing problem, the WNN is coupled with an AR real-time correction model in practical application. An adaptive fading factor RLS method is integrated with the AR model to calculate the forecast errors of the WNN continuously. The WNN forecast value is corrected by adding the forecast error to it. This improvement made on the WNN leads to increased forecast accuracy. The proposed method in this paper has considerable application values.

References

[1] Zhu Xingming, Lu Changna, Wang Ruyun, et al. Artificial neural network model for flood water level forecasting[J]. *Journal of Hydraulic Engineering*, 2005, **36**(7): 806 − 811. (in Chinese)

[2] Narendra K S, Parthasarathy K. Identification and control of dynamic systems using neural networks [J]. *IEEE Transactions on Neural Network*, 1990, **1**(1): 4 − 27.

[3] Huang Dexian, Wang Jingchun, Jin Yihui. Application research of wavelet neural networks in process control[J]. *Tsinghua Science and Technology*, 1999, **39**(1): 91 − 94.

[4] Li Huanqin, Wan Baiwu. Using high dimension wavelet neural network model product quality for hot rolling mill [J]. *Chinese Journal of Engineering Mathematics*, 2006, **23**(4): 614 − 618.

[5] Zhang Q H, Benveniste A. Wavelet network[J]. *IEEE Transactions on Neural Network*, 1992, **3**(6): 889 − 898.

[6] Stéphane M. *A wavelet tour of signal processing*[M]. Beijing: China Machine Press, 2003. (in Chinese)

[7] Liu Dongying, Song Xingyuan. BP neural network model of river flood forecast and its application[J]. *China Rural Water and Hydropower*, 2005(2): 47 − 50. (in Chinese)

[8] Zhao Xuezhi, Zou Chunhua, Chen Tongjian et al. A research on the initialization of parameters of wavelet neural networks [J]. *Journal of South China University of Technology: Natural Science Edition*, 2003, **31**(2): 77 − 79, 84. (in Chinese)

[9] Zhou Yi, Li Zhijia. Research on the application of three improved RLS procedures in real-time flood forecasting[J]. *Water Power*, 2006, **32**(8): 14 - 16. (in Chinese)

[10] Akaike H. A new look at the statistical model identification [J]. *IEEE Transactions on Automatic Control*, 1974, **19**(6): 716 - 723.

基于小波神经网络和自回归模型耦合的河道洪水预测方法

李致家　周　轶　马振坤

(河海大学水文水资源学院,南京 210098)

摘要:在分析非线性河道洪水预报方法中常用 BP 神经网络不足的基础上,采用具有快速收敛和更有效非线性逼近能力特性的小波神经网络. 为适应洪水演进的时变特性,将所建立的用于河道洪水预报的小波神经网络与自回归实时校正模型耦合,校正值为小波神经网络预报值与自回归模型预报误差之和. 自回归实时校正模型的参数通过自适应衰减因子递推最小二乘动态更新以提高校正效果. 将该方法应用于西江高要断面洪水预报,计算结果验证了其有效性.

关键词:河道洪水预测;小波神经网络;自回归模型;递推最小二乘;自适应衰减因子

中图分类号:P338