

Research on Stock Price Prediction Based on BP Wavelet Neural Network with Mexico Hat Wavelet Basis

Wang pingan
XiJing College
Xi'an, Shaanxi, 710123, China
fyw2121@sohu.com

Lou Yuanwei
Air Traffic Control and Navigation College
Air Force Engineering University
Xi'an, Shaanxi, 710051, China
911yhk@163.com

Lei Lei
School of Business Administration
Henan University of Economics and Law
Zhengzhou, Henan, 050046, China
lyx521@yahoo.com.cn

Abstract—In order to improve the prediction ability of stock price, a prediction method based on Wavelet Neural Network (WNN) is proposed. First BP algorithm is used to optimize the parameters of WNN with Mexico Hat wavelet basis for the establishment of the stock price prediction model, and then the built model is applied to predict the stock price movement on the basis of 15 features. The simulations on daily closing price index of SSE Composite Index indicate that, the proposed method has the advantages of simple structure, strong implementation and good prediction accuracy, and gets better stock price prediction in contrast with single neural network and genetic neural network. This verifies the feasibility and effectiveness of the method in the application of stock price prediction.

Keywords—Wavelet Neural Network; BP algorithm; Stock Price; Prediction

I. INTRODUCTION

Stock is the product of market economy. The stock investment has become an important part of people's daily life. The soaring and plunging in stock market can cause the oscillation of the financial market, which has a direct impact on the stability of financial market and the healthy development of economy. If the climbing and falling of stock price can be predict accurately, and accordingly the stock market is timely regulated and guided, it can undoubtedly provide the strong confidence and guarantee to sustainable economy development [1][2].

Since the beginning of the last century, after a long-term development, the research on neural network has made great progress. Many scholars have tried to use neural network in stock prediction research. In 1988, neural network was first used to predict the daily return rate of IBM ordinary shares by White et al. The Neural Network (NN) has excellent self-organizing, self-adapting, self-learning, distributed processing, strong fault tolerance and so on, which provides a

new mode for stock price prediction and changes the existing mode completely with many achievements, however their achieved results are unsatisfactory [3~9]. Wavelet Neural Network (WNN) is a new neural network developed on the basis of wavelet analysis. Compared with the traditional neural network, WNN has a great performance improvement, suitable for all areas of neural network application [10][11].

On the basis of this, a prediction method of stock price based on WNN is proposed to meet the requirement on accuracy and practicability.

II. WAVELET NEURAL NETWORK

Wavelet neural network combines the time- frequency localization characteristics of wavelet transform and the self-learning function of neural network, improving greatly the performance. WNN is a special form of neural network. Because of the outstanding performance in solving nonlinear problem, it is fit for all areas in which the neural network used. In the learning algorithm of WNN, BP algorithms are most widely used.

For the discrete wavelet transform, through the wavelet function

$$\psi_{m,n}(t) = 2^{-\frac{m}{2}} \psi(2^{-m}t - n), \quad m, n \in Z \quad (1)$$

A set of orthogonal basis of the space can be constructed, $L^2(R) = \bigoplus_m W_m$. Using the multi-resolution analysis of the space $L^2(R)$, a series of closed subspaces $\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \cdots$ can be given, $V_m = \text{span}\{2^{-\frac{m}{2}} \times \varphi(2^{-m}t - k)\}$, $V_{m+1} = V_m \oplus W_m$, $\varphi(t)$ is the scaling function corresponding

to the wavelet. Then the function in $L^2(R)$ can be decomposed into

$$\hat{f} = \sum_{m,n} w_{m,n} \psi_{m,n}(t) = \sum_{m,n} w_{m,n} 2^{-\frac{m}{2}} \psi(2^{-m}t - n) \quad (2)$$

The above formula can directly be realized through a feed forward WNN with single hidden layer, as shown in Figure 1, where $x_i (i=1, \dots, I)$ is the network input, $y_k (k=1, \dots, K)$ is the network output, w_{ji} is the connection weight from input layer to hidden layer, w_{kj} is the connection weight from hidden layer to the output layer, and I, J, K is the number of input layer nodes, hidden layer nodes and output layer nodes respectively. The activation function of hidden layer nodes is the orthogonal wavelet function $\psi_{m,n}(t)$.

In fact, the wavelet in formula (3) is also called the binary wavelet that has the zoom effect. In general, the selected magnification 2^{-m} corresponds to a part of the observed signals. To understand the details of the signal in depth, the magnification can be increased (i.e. decrease m), on the contrary, to roughly observe the signal content, the magnification can be decreased (i.e. increase m).

For the network in Fig.1, the wavelet basis function selected is different; the learning algorithm constituted is also different. In the paper, the Mexico hat basis function $\psi(t) = (1-t^2)e^{-t^2/2}$ is used to derive BP learning algorithm based on the orthogonal wavelet.

Suppose that the learning sample set $\{X_i, \hat{Y}_i\}_i^N$, where X is the input vector (x_1, x_2, \dots, x_I) , \hat{Y} is the expected output vector $(\hat{y}_1, \hat{y}_2, \dots, \hat{y}_K)$, N is the number of learning sample. Then the input of j -th node in hidden layer is $s_j = \sum_{i=1}^I w_{ji} x_i (j=1, 2, \dots, J)$, the output of j -th node in hidden layer is

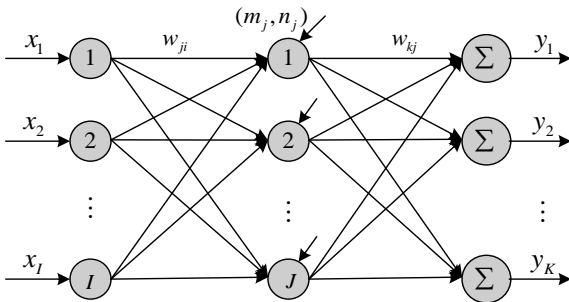


Fig. 1. WNN based on orthogonal wavelet

$$h_j = \psi_{m_j, n_j}(s_j) = 2^{-\frac{m_j}{2}} e^{-\frac{(2^{-m_j} s_j - n_j)^2}{2}} \times$$

$$\cos(5 * 2^{-m_j} s_j - 5n_j) \quad (j=1, 2, \dots, J) \quad (3)$$

The output layer is the linear combination of the wavelet frame. The output of k -th neuron is

$$y_k = \sum_{j=1}^J w_{kj} h_j, (k=1, 2, \dots, K) \quad (4)$$

If the error function is defined as $E = 0.5 \sum_{k=1}^K (\hat{y}_k - y_k)^2$, the instantaneous gradient vector of each parameter can be calculated by

$$\delta m_j = \frac{\partial E}{\partial m_j} = \sum_{k=1}^K (\hat{y}_k - y_k) w_{kj} \frac{\partial h_j}{\partial m_j} \quad (5)$$

$$\delta n_j = \frac{\partial E}{\partial n_j} = \sum_{k=1}^K (\hat{y}_k - y_k) w_{kj} \frac{\partial h_j}{\partial n_j} \quad (6)$$

$$\delta w_{ji} = \frac{\partial E}{\partial w_{ji}} = \sum_{k=1}^K (\hat{y}_k - y_k) \cdot w_{kj} \cdot \frac{\partial h_j}{\partial s_j} \cdot \frac{\partial s_j}{\partial w_{ji}} \quad (7)$$

$$\delta w_{kj} = \frac{\partial E}{\partial w_{kj}} = (\hat{y}_k - y_k) \cdot h_j \quad (8)$$

where

$$\begin{aligned} \frac{\partial h_j}{\partial m_j} &= 2^{-\frac{m_j}{2}} e^{-\frac{(2^{-m_j} s_j - n_j)^2}{2}} [2 \ln 2 \cdot s_j (2^{-m_j} s_j - n_j) \times \\ &\quad \cos(5 * 2^{-m_j} s_j - 5n_j) - \frac{1}{2} \ln 2 \cdot \cos(5 * 2^{-m_j} s_j \\ &\quad - 5n_j) + 10 \ln 2 \cdot s_j \cdot \sin(5 * 2^{-m_j} s_j - 5n_j)] \quad (9) \end{aligned}$$

$$\begin{aligned} \frac{\partial h_j}{\partial n_j} &= (2^{-m_j} s_j - n_j) 2^{-\frac{m_j}{2}} e^{-\frac{(2^{-m_j} s_j - n_j)^2}{2}} \times \\ &\quad \cos(5 * 2^{-m_j} s_j - 5n_j) + 5 \cdot 2^{-\frac{m_j}{2}} \times \\ &\quad e^{-\frac{(2^{-m_j} s_j - n_j)^2}{2}} \sin(5 * 2^{-m_j} s_j - 5n_j) \quad (10) \end{aligned}$$

$$\begin{aligned} \frac{\partial h_j}{\partial s_j} &= -2^{-\frac{m_j}{2}} [2^{-m_j} (2^{-m_j} s_j - n_j) \cos(5 * 2^{-m_j} s_j \\ &\quad - 5n_j) + 5 \cdot 2^{-m_j} \sin(5 * 2^{-m_j} s_j - 5n_j)] \quad (11) \end{aligned}$$

$$\frac{\partial s_j}{\partial w_{ji}} = x_i \quad (12)$$

For the conventional BP algorithm, the network parameters can be adjusted by

$$\theta(k+1) = \theta(k) + \lambda \Delta \theta(k) + (1-\lambda) \eta(k) \frac{\partial E_k}{\partial \theta(k)} \quad (13)$$

where θ represents the parameter vector $\{w_{ji}, w_{kj}, m_j, n_j\}$, $\Delta\theta(k)$ is the parameter difference after last adjustment. λ is the momentum factor that can reduce the sensitivity of network on the error surface detail to effectively protect from falling into local minimum. $\eta(k)$ is the learning rate. If $\eta(k)$ is taken too small, the algorithm has slow convergence and easy to fall into local minimum; if too larger, the algorithm has fast convergence, but may oscillate. Thus, the adaptive learning rate can be introduced to dynamically adjust $\eta(k)$. The specific adjustment formula is as follows

$$\eta(k) = \begin{cases} 1.05\eta(k-1) & E(k) < E(k-1) \\ 0.7\eta(k-1) & E(k) > 1.04E(k-1) \\ \eta(k-1) & \text{otherwise} \end{cases} \quad (14)$$

where 1.05, 0.7, 1.04 are all empirical values.

WNN derived by orthogonal wavelet transform is actually that, on the basis of orthogonal compactly supported wavelets constructed by Daubechies and multiresolution analysis, select the most appropriate scaling parameters and the translation parameters by the scaling function, adjusting the wavelet coefficient to achieve the best approximation to the function.

III. EXPERIMENT SIMULATION

Since China Shanghai stock market started early and has accumulated a huge mass of data, and the market index combines the influence of various aspects with a strong representation. Therefore, SSE Composite Index daily closing index can be taken as the research object. Research goal is to predict the following day variation movement T of SSE Composite Index. According to the size of the percent yield, 5 variation movements can be defined as

Plunging	($T = 1$)	$R \in (-\infty, -2.0\%]$
Falling	($T = 2$)	$R \in (-2.0\%, -0.5\%]$
Invariability	($T = 3$)	$R \in (-0.5\%, 0.5\%)$
Climbing	($T = 4$)	$R \in [0.5\%, 2.0\%)$
Soaring	($T = 5$)	$R \in [2.0\%, \infty)$

where R denotes the percent yield, and can be defined by the following formula

$$R_t = (P_t - P_{t-1}) / P_{t-1} \times 100\% \quad (15)$$

where P_t is the price series.

For SSE Composite Index closing index, the data on its fluctuation trend and corresponding feature from 1/2003 to 11/2009 is separate into 7 sample sets on a yearly basis. Each sample set is divided into the training sample set and test sample set with 4:1 ratio, as shown in Table 1. Parameter settings: w_{ji}, w_{kj}, m_j, n_j can be initialized as the random number in $[-0.5, 0.5]$, the maximum training number is 200.

The capability comparison of is given in Table 2 where are the average training results for 50 times. The prediction error plots of WNN model with Mexico Hat wavelet basis is shown in Fig. 2 at 2003. The comparison of convergence curves of 3 modeling methods is shown in Fig. 3 (BP Neural Network (BPNN), GA Neural Network (GANN) and WNN with Mexico Hat wavelet basis). The comparison of average prediction result between 3 models is shown in Fig. 4.

TABLE I. TRAINING AND TEST SAMPLE SET YEARLY

Year	Training sample set	Test sample set	Total
2003	200	50	250
2004	198	49	247
2005	195	48	243
2006	198	48	246
2007	192	47	239
2008	156	40	196
2009	197	45	232
Total	1139	282	1667

TABLE II. MODEL AND CORRESPONDING TRAINING AND TEST ACCURACY FROM 2003 TO 2009

Year	Training accuracy	Test accuracy
2003	65.5	59.2
2004	69.2	63.3
2005	66.7	62.5
2006	72.2	68.8
2007	70.3	68.1
2008	67.9	64.5
2009	64.5	61.7

As can be seen from results, WNN is significantly better than BPNN and GANN in prediction effects. This shows that WNN technique based on BP algorithm with Mexico Hat wavelet basis can effectively reduce the prediction error of stock price caused by nonlinear characteristics of the system for a certainty. Compared with BPNN and GANN, WNN has the better ability with 15 features of training error, testing error, convergence speed and stability. This fully shows that WNN can effectively improve the prediction accuracy for stock price.

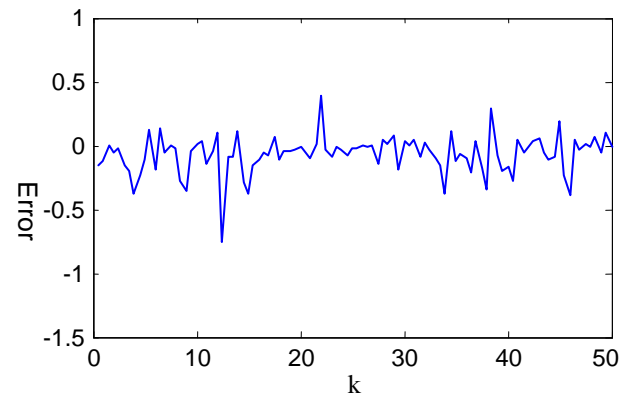


Fig. 2. Prediction error of WNN model with Mexico Hat wavelet basis

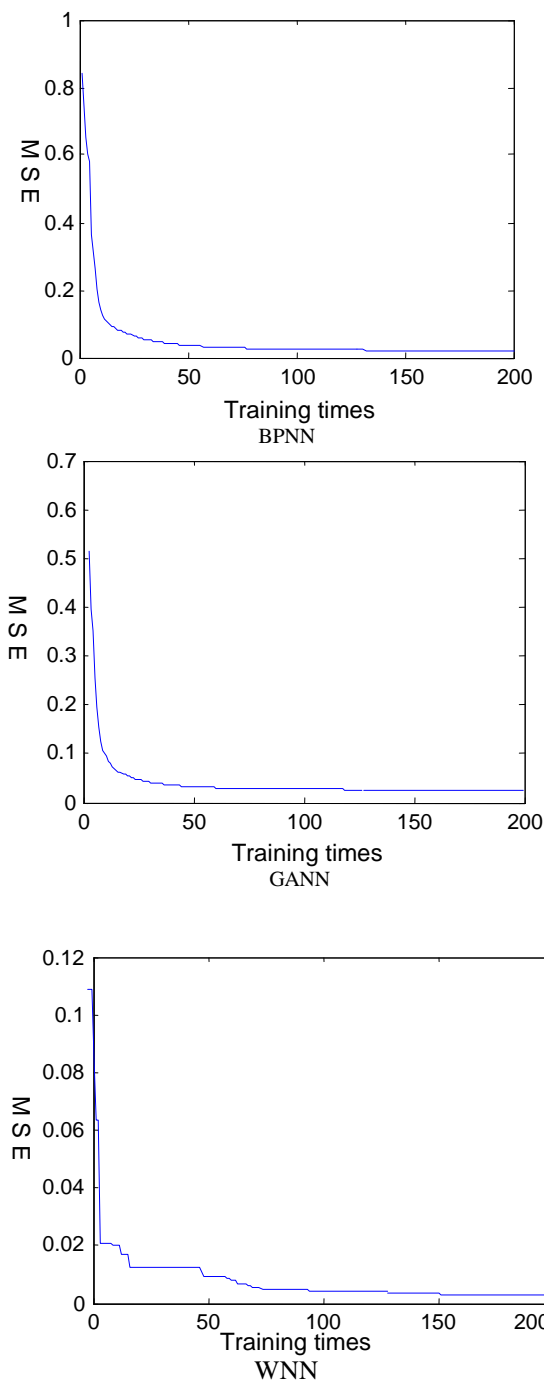


Fig. 3. Comparison on training convergence curves for various modeling method

IV. CONCLUSION

For the stock price prediction, a prediction method based on BP algorithm with Mexico Hat wavelet basis is proposed. Simulation shows that, the proposed WNN method not only inherits the nonlinear mapping ability of neural network and the local property of wavelet transform, but also has the fast optimum capability. Compared with BPNN and GANN in the prediction of stock price, the proposed WNN method has obvious improvement in training, generalization and convergence with great superiority. Thereby, it provides an

effective solution for the prediction of stock price. It is believed that, with further research, by selecting more reasonable stock price volatility indicators, the prediction accuracy can still be improved.

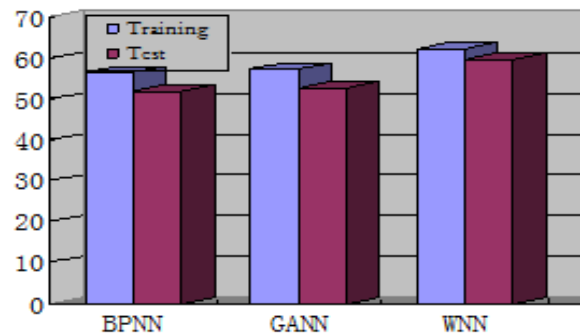


Fig. 4. Comparison of average prediction accuracy for 3 models

REFERENCES

- [1] T. White, "Economic prediction using neural networks: the case of IBM daily stock returns," *IEEE International Conference on Neural Networks*, vol. 2, No. 6, 1988, pp. 451-458.
- [2] L. Zhu, J. M. He, S. Chang, "Rough set integrated neural network model for forecasting stock price," *Chinese Journal of Management Science*, vol. 10, No. 4, 2002, pp. 7-12.
- [3] O. Dennis, M. Charles, "Neural network forecasts of Canadian stock returns using accounting ratios," *International Journal of Forecasting*, vol. 19, No. 3, 2003, pp. 453-465.
- [4] T. Kimoto, K. Asakawa, M. Yoda, et al, "Stock market prediction system with modular neural networks," *1990 IJCNN, International Joint Conference on Neural Networks*, vol. 1, 1990, pp. 1-6.
- [5] N. Baba, M. Kozaki, "An intelligent forecasting system of stock price using neural networks," *1992IJCNN, International Joint Conference on Neural Networks*, vol. 1, 1992, pp. 371- 377.
- [6] M. H. Pesaran, M. A. Timermann, "Recursive modeling approach to predicting stock returns," *Cambridge: University of Cambridge Press*, 1999.
- [7] K. Kwong, "Financial forecasting using neural network or machine learning techniques," *Thesis of Electrical Engineering, University of Queensland*, 2001.
- [8] M. T. Philp, K. Paul, S. O. Choy, "Design and implementation of NNS for Hong Kong stock price forecasting," *Engineering Applications of Artificial Intelligence*, vol. 20, No. 4, 2007, pp. 453-461.
- [9] A. Leonel Laboissiere, A. S. R. Fernandesb, G. L. Guilherme, "Maximum and minimum stock price forecasting of Brazilian power distribution companies based on artificial neural networks," *Applied Soft Computing*, vol. 35, 2015, pp. 66-74.
- [10] Q. H. Zhang, A. Benveniste, "Wavelet network," *IEEE Transactions on Neural Networks*, vol. 3, No. 6, 1992, pp. 889- 898.
- [11] Q. H. Zhang, A. Benveniste, "Using wavelet network in nonparametric estimation," *IEEE Transactions on Neural Networks*, vol. 8, No. 2, 1997, pp. 227-236.