ICIC Express Letters
Volume **10**, Number **7**, July **2016**

ICIC International ©2016 ISSN 1881-803X
pp. 1627–1633

TIME SERIES PREDICTION OF BANK CASH FLOW BASED ON BP NEURAL NETWORK ALGORITHM

Jiesheng Wang[1,2,*] and Chenxu Ning[1]

[1]School of Electronic and Information Engineering
[2]National Financial Security and System Equipment Engineering Research Center
University of Science and Technology Liaoning
No. 185, Qianshan Road, Anshan 114051, P. R. China
[*]Corresponding author: wang_jiesheng@126.com; NCX710754167@163.com

Received December 2015; accepted March 2016

Abstract. *In order to improve the accuracy and real-time of all kinds of information in the cash business and enhance the linkage between cash inventory forecasting and cash management information in the commercial bank, the improved BP neural network algorithms are adopted to realize the time series prediction of bank cash flow, which are momentum back propagation (MOBP), variable leaning rate back propagation (VLBP), conjugate gradient back propagation (CGBP) and Levenberg-Marquardt BP neural network algorithms. The simulation experiments are carried out on the reality commercial bank's cash flow data and the predictive performance comparison results show the effectiveness of the proposed methods.*
Keywords: Time series prediction, Bank cash flow, BP neural network

1. **Introduction.** In recent year, the time series modeling and prediction are one of the most active research topics in the academic research and engineering practice [1]. The bank cash flow forecasting management information system is designed to create a system management platform for the prediction and analysis of the commercial bank cash flow. Its purpose is to provide effective data all levels of organization to analyze and assess cash business operation conditions. A new multivariate, time-series prediction model is set up by employing the past values of earnings, short-term accruals and cash-flows [2]. A new multivariate radial basis functions neural network model is proposed to predict the complex chaotic time series simultaneously with single-step prediction and multistep prediction [3]. A random data-time effective radial basis function neural network in determination of the output weights, the center vectors and the widths in the hidden layer of the network is proposed to predict the financial price series of crude oil, SSE, N225 and DAX [4]. The first moving average prediction method, the second moving average prediction method, the first exponential smoothing prediction and the second exponential smoothing prediction methods are adopted to realize the time series prediction of bank cash flow, respectively [5]. Aiming at the complexity and the dynamic progress of a construction project, a novel inference model, named as adaptive time dependent least squares support vector machine (LS-SVM$_{\mathrm{AT}}$) is proposed for forecasting cash flow demand throughout various phases of the project [6]. The time series and cross-sectional behavior of the accounting rate of return (ARR) and tests on whether ARR is a good predictor of the cash flow per share (CFPS) are investigated [7]. The evolutionary fuzzy support vector machine inference model for time series data (EFSIM$_{\mathrm{T}}$), an artificial intelligence hybrid system focusing on the management of time series data characteristics which fuses fuzzy logic (FL), weighted support vector machines (weighted SVMs) and a fast messy genetic algorithm (fmGA), represents a promising alternative approach to predicting cash flow [8].

BP neural network is a multi-layer forward propagation neural network which is one-way transmission with advantages of parallel processing, distributed storage and adaptive learning. An approach of back propagation neural network with rough set (RSBP) is presented for complicated STLF with dynamic and non-linear factors to develop the accuracy of predictions [9]. To predict the time delay induced in the networked control systems online, a BP feed-forward neural network is adopted and the training algorithm of the BP neural network is discussed in details [10]. The influencing factors of blasting vibration parameters were analyzed based on the test of an iron mine of Xinjiang province and the BP neural network model was established to predict the peak vibration velocity, dominant frequency, and the time of duration, respectively [11].

Aiming at the prediction problem of commercial bank cash flow, BP neural network algorithm is used in bank cash flow time series forecasting. The simulation results show the effectiveness of the proposed method. The paper is organized as follows. In Section 2, the basic principle of BP neural network is introduced. The time series prediction of bank cash flow based on BPNN is presented in Section 3. In Section 4, the simulation experiments and results analysis are introduced in details. Finally, the conclusion illustrates the last part.

2. **Principle of BP Neural Network.**

2.1. **Topology of BP neural network.** Back propagation neural network (BPNN) is a kind of supervised learning algorithm, whose major characteristic is the signal forward and error backing propagation by observed data. BP neural network includes input layer, hidden layer and output layer, whose topological structure is shown in Figure 1.

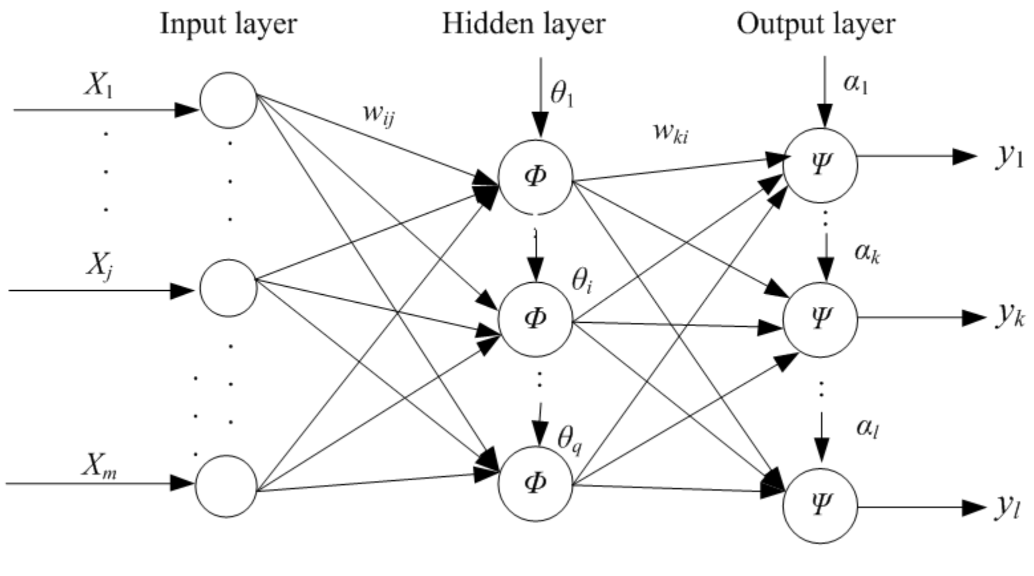

FIGURE 1. Topology of BP neural network

In Figure 1, X_j represents the input of the input layer at node j, $j = 1, \cdots, m$; w_{ij} is the weight between the node i in hidden layer and the node j in input layer; θ_i is a threshold value of the i-th hidden layer node; $\phi(x)$ is the excitation function of hidden layer; w_{ki} is a weight between the node k in output layer and the node i in hidden layer, $i = 1, \cdots, q$; a_k is a threshold value of the k-th output layer node, $k = 1, \cdots, l$; $\psi(x)$ is the excitation function of output layer; y_k is the output of the output layer at node k, $k = 1, \cdots, l$.

2.2. **Learning process of BPNN.** The learning process of BPNN is to realize the relationship between signaling forward and error backing propagation. The general learning steps of BP neural network are described as follows.

Step 1: Initialize each weight value to a small random number with distributed uniformly random numbers as the initial connection weights and the threshold values of the nodes.

Step 2: Calculate the actual output of BPNN:

① For the input layer nodes, their output O_j^I is equal to the input data X_j, that is to say $O_j^I = X_j$, $j = 1, \cdots, m$.

② For the hidden layer nodes, their input is described as follows:

$$net_i^H = \sum_{j=1}^{m} w_{ij}^{HI} O_j^I, \quad i = 1, \cdots, q \tag{1}$$

The output is

$$O_i^H = f\left(net_i^H - \theta_i^H\right) \tag{2}$$

where w_{ij}^{HI} is the connection weights between the node i in hidden layer and the node j in input layer; θ_i is a threshold value of hidden layer node i; q is the number of hidden layer nodes; O_j^I is the output of the input layer at node j, that is X_j; f is Sigmoid function.

③ Input of the output layer nodes is described as follows:

$$net_k^O = \sum_{i=1}^{q} w_{ki}^{OH} O_i^H, \quad k = 1, \cdots, l \tag{3}$$

The output of the output layer nodes is

$$y_k = f\left(net_k^O - \theta_k^O\right) \tag{4}$$

where w_{ki}^{OH} is the connection weights between the output layer node k and the hidden layer node i; θ_k^O is a threshold value of the output layer node k.

Step 3: The error of the output node is calculated by the following equation.

$$e_k = d_k - y_k \tag{5}$$

Then calculate the error squared sum of all output nodes and obtain the energy function.

$$E = \frac{1}{2} \sum_{k=1}^{l} (d_k - y_k)^2 \tag{6}$$

If E is less than predetermined value, turn to step (5); otherwise, continue to step (4).

Step 4: Adjust the weights of BPNN.

① The weights between the output layer nodes and the hidden layer nodes are adjusted as follows.

$$\begin{aligned} \bar{w}_{ki}^{OH} &= w_{ki}^{OH} + \Delta w_{ki}^{OH} \\ \Delta w_{ki}^{OH} &= \eta \sigma_k^O . O_i^H \\ \sigma_k^O &= (d_k - y_k) . y_k (1 - y_k) \end{aligned} \tag{7}$$

where η is the training rate, general $\eta = 0.01 \sim 1$.

② The weights w_{ij}^{HI} between the hidden layer nodes and the input layer nodes are adjusted as follows.

$$\begin{aligned} \bar{w}_{ij}^{HI} &= w_{ij}^{HI} + \Delta w_{ij}^{HI} \\ \Delta w_{ij}^{HI} &= \eta \sigma_i^H . O_j^I \\ \sigma_i^H &= O_i^H \left(1 - O_i^H\right) \sum_{k=1}^{l} \sigma_k^O w_{ki}^{OH} \end{aligned} \tag{8}$$

Step 5: Carry on the next training samples. The learning process of BPNN is complete until each training sample satisfies the target.

3. Time Series Prediction of Bank Cash Flow Based on BPNN.

3.1. Determination of BP neural network structure. In order to predict the bank cash flow with the BP neural network model, firstly the network structure is determined. In this paper, the multiple input and single output three layers BP neural network is used. The neuron transfer function in hidden layer is used bipolar S type tangent function (tansig):

$$f(x) = \frac{1 - e^{-x}}{1 + e^{-x}}, \quad (-1, 1) \tag{9}$$

The neuron transfer function in output layer uses the linear transfer function (purelin):

$$f(x) = x \tag{10}$$

BP neural network is a multi-layer forward propagation neural network which is one-way transmission with advantages of parallel processing, distributed storage and adaptive learning. However, it has the problems of slow convergence speed and local extreme. Therefore many improved algorithms are proposed, which can be divided into two categories roughly:

(1) The improved methods based on standard gradient descent, including additional momentum back propagation (MOBP) algorithm and variable learning rate back propagation (VLBP) algorithm;

(2) The improved methods based on standard numerical optimization, which include the conjugate gradient back propagation (CGBP) algorithm and Levenberg-Marquardt back propagation (LMBP) algorithm.

3.2. Performance index of prediction accuracy. The prediction error is the deviation between the predicted results and the actual results, which determines the prediction accuracy. Suppose $y_1, y_2, \cdots, y_n$ are the actual observed values of the predicted target and $\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n$ are the predicted values.

(1) Absolute error

$$a_t = y_t - \hat{y}_t, \quad t = 1, 2, \cdots, n \tag{11}$$

where a_t is the absolute error at the point t. Obviously, a_t is the most direct measure index of the prediction error, but it is affected by the measurement unit of predicted object. So it is unsuitable as the final measure indicator of prediction accuracy.

(2) Relative error

$$\hat{a}_t = \frac{a_t}{y_t} = \frac{y_t - \hat{y}_t}{y_t}, \quad t = 1, 2, \cdots, n \tag{12}$$

where $\widehat{a}_t$ is the relative error at the point t, which is usually expressed as a percentage and measures the accuracy of the predicted values relative to the observed value at the predicted point t.

(3) Prediction accuracy

$$A_t = 1 - |y_t - \hat{y}_t|/y_t \quad 0 \le |y_t - \hat{y}_t|/y_t \le 1 \tag{13}$$

$$A_t = 0 \quad |y_t - \hat{y}_t|/y_t > 1 \tag{14}$$

where A_t is the prediction accuracy at the prediction point t.

(4) Mean square error (MSE)

Mean square error (MSE) is a kind of convenient method to measure the average error to evaluate the degree of data change, which is described as follows.

$$MSE = \frac{1}{n}\sum_{t=1}^{n} (y_t - \hat{y}_t)^2 \tag{15}$$

where MSE is mean square error (MSE) at the point t.

4. **Simulation Results and Analysis.** Based on the collected data in a commercial bank from January to April in 2012, the simulation experiments under the above mentioned four improved BPNN predictive methods are carried out. It can be seen clearly that which time series predictive method is more suitable for bank cash flow by comparing the experimental results. Five days are selected as the predictive analysis cycle, that is to say the historical inventory limit data for five days are used as the predictive basis, then order five days as input data to network in turn, and the followed day as the network output or target data. With the arrangement of rolling in this manner, the training samples of BP neural network are formed. The iteration times are set 500, the error index is 0.001, the number of the hidden layer neurons is 14 and the number of output neurons is 1. The four mentioned methods are used to carry out the training process and the comparison results are shown in Figure 2. The compared results of the iteration numbers in the field of convergence to the expected error are shown in Table 1.

It can be seen from Figure 2 that these algorithms can converge to the specified error after 500 epochs in addition to the MOBP algorithm and VLBP algorithm for the provided training samples. Seen from the comparison results in Table 1, the LMBP algorithm has a great advantage in the training speed than other three algorithms. In order to facilitate the observation, the inventory limits from January to March in 2012 are adopted (90 sample points are chosen), the iteration number is 500, and the error index is 0.001. The comparison results of cash flow predictive are shown in Figure 3, and the simulation comparison results of errors are shown in Figure 4. The four mentioned methods are used to carry out the simulation on the same pretreated experiments. The performance comparison results are shown in Table 2.

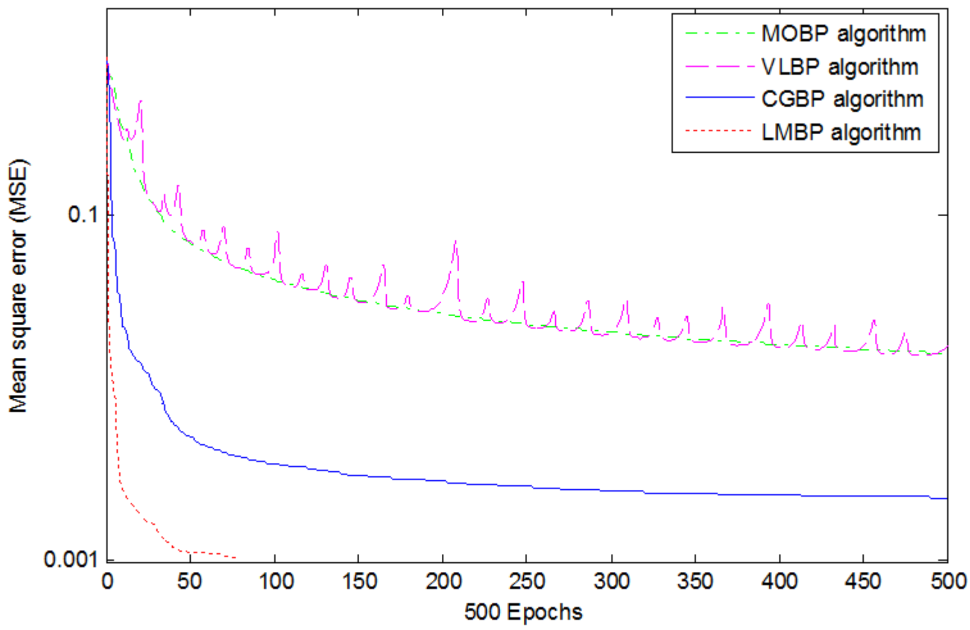

FIGURE 2. Comparison results of training error (2012)

TABLE 1. Comparison of iteration number

Improved BP algorithms	Epoch	The training error
MOBP algorithm	Failed to reach the specified error in specified steps	
VLBP algorithm	Failed to reach the specified error in specified steps	
CGBP algorithm	142	0.00089596
LMBP algorithm	10	0.00099187

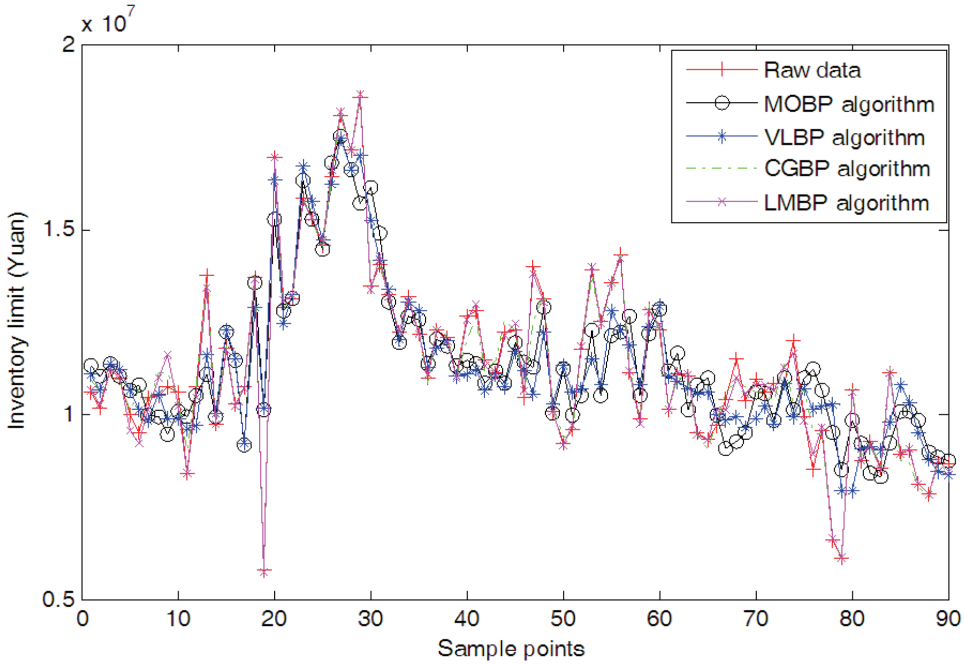

FIGURE 3. Comparison of predictive results (2012)

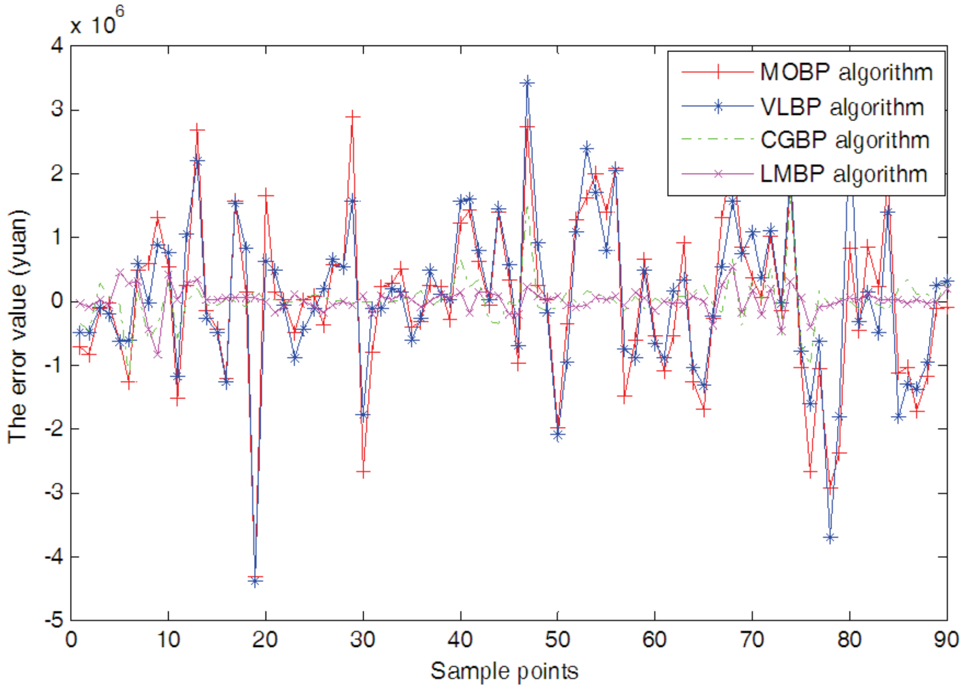

FIGURE 4. Comparison of predictive error (2012)

TABLE 2. Performance comparison

Performance index	MOBP	VLBP	CGBP	LMBP
MSE	8.62×10^{-5}	8.76×10^{-5}	2.35×10^{-6}	8.86×10^{-7}
Absolute error	7.09×10^{5}	5.12×10^{5}	1.27×10^{5}	7.60×10^{4}
Relative error (%)	4.67%	4.32%	1.29%	0.46%
Prediction accuracy (%)	95.33%	95.68%	98.71%	99.64%

It can be seen clearly from the above simulation comparison results that the four algorithms are ideal, but Levenberg-Marquardt back propagation (LMBP) neural network algorithm is better than other three algorithms for predicting the bank cash flow time series. The extent of the predicted values fitting with the original time series is very good, which reflects the variation trends of the original time series in different time periods. The performance compared results in Table 2 show that the LMBP neural network method obtains the higher prediction precision for the cash flow time series.

5. **Conclusions and Future Work.** Four time series predictive methods (MOBP algorithm, VLBP algorithm, CGBP algorithm and LMBP algorithm) are adopted to realize the real-time prediction of cash flow in the commercial bank. By comparing the time series predictive performance under four algorithms and analyzing the simulation results in-depth, the LMBP algorithm prediction method is the optimal prediction method. In future, other neural networks and swarm intelligent algorithms will be combined to realize the real-time prediction of cash flow in the commercial bank.

Acknowledgment. This work is partially supported by the National Key Technologies R & D Program of China (Grant No. 2014BAF05B01), the Project by National Natural Science Foundation of China (Grant No. 21576127), the Program for Liaoning Excellent Talents in University (Grant No. LR2014008), the Project by Liaoning Provincial Natural Science Foundation of China (Grant No. 2014020177), the Program for Research Special Foundation of University of Science and Technology of Liaoning (Grant No. 2015TD04) and the Opening Project of National Financial Security and System Equipment Engineering Research Center (Grant No. USTLKEC201401).

REFERENCES

[1] N. K. Kasabov and Q. Song, DENFIS: Dynamic evolving neural-fuzzy inference system and its application for time-series prediction, *IEEE Trans. Fuzzy Systems*, vol.10, no.2, pp.144-154, 2002.

[2] K. S. Lorek and G. L. Willinger, A multivariate time-series prediction model for cash-flow data, *Accounting Review*, vol.71, no.1, pp.81-102, 1996.

[3] D. Chen and W. Han, Prediction of multivariate chaotic time series via radial basis function neural network, *Complexity*, vol.18, no.4, pp.55-66, 2013.

[4] H. Niu and J. Wang, Financial time series prediction by a random data-time effective RBF neural network, *Soft Computing*, vol.18, no.3, pp.497-508, 2014.

[5] W. Cui, J. Wang and C. Ning, Time series prediction method of bank cash flow and simulation comparison, *Algorithms*, vol.7, no.4, pp.650-662, 2014.

[6] M. Y. Cheng, N. D. Hoang and Y. W. Wu, Cash flow prediction for construction project using a novel adaptive time-dependent least squares support vector machine inference model, *Journal of Civil Engineering & Management*, pp.679-688, 2015.

[7] G. Kelly and M. I. Tahir, Time series and cross-sectional dynamics of accounting rate of return and the prediction of cash flow per ordinary share, *Accounting Accountability & Performance*, vol.15, no.2, pp.1-39, 2009.

[8] M. Y. Cheng, Y. W. Wu, L. T. Dan et al., Enhanced time-dependent evolutionary fuzzy support vector machines inference model for cash flow prediction and estimate at completion, *International Journal of Information Technology & Decision Making*, vol.12, no.4, pp.679-710, 2013.

[9] Z. Xiao, S. J. Ye, B. Zhong et al., BP neural network with rough set for short term load forecasting, *Expert Systems with Applications*, vol.36, no.1, pp.273-279, 2009.

[10] J. Yi, Q. Wang, D. Zhao et al., BP neural network prediction-based variable-period sampling approach for networked control systems, *Applied Mathematics and Computation*, vol.185, no.2, pp.976-988, 2007.

[11] X. Luo, X. Fan, Z. Dai et al., BP neural network application and model modification in prediction of open-pit mine blasting vibration parameters, *Journal of Central South University*, vol.44, no.12, pp.5019-5024, 2013.