# Nonlinear time series of deformation forecasting using improved BP neural networks

**Cai-yun Gao[1, 2*], Xi-min Cui[1]**

[1]*College Of Geoscience and Surveying Engineering, China University of Mining and Technology, Beijing 100083, China*

[2]*College of Surveying and Mapping Engineering, Henan University of Urban Construction, Pingdingshan 467044, China*

*Received 6 July 2014, www.tsi.lv*

**Abstract**

Although the back propagation neural network has been successfully employed in various fields and demonstrated promising results, literatures show its performance still could be improved. Therefore, we present a comprehensive comparison study on the application of different BP algorithm in time series of deformation forecasting. Four types of typical improved BP algorithm, namely, momentum, conjugate gradient, Quasi-Newton and Levenberg-marquardt algorithms, are investigated. An illustrative example of high-rise building settlement deformation is adopted for demonstration. Results show that the improved BP algorithms can increase the prediction accuracy and have faster convergence speed.

*Keywords:* artificial neural networks, back propagation, deformation forecasting, learning rate, convergence, improved

## 1 Introduction

Artificial neural networks are often used to model and forecast complicated nonlinear time series, especially deformation systems (such as dam deformation, land subsidence in coal mine, settlement of subway tunnel etc.). Theoretically, multilayer feed-forward neural networks can accurately approximate almost any nonlinear function and thus error back propagation algorithm (BP) have been widely researched and applied in time series of deformation forecasting [1-5].

Although there are many successful applications of standard BP algorithm, it has many drawbacks, such as:

1) require a long time to train the networks,

2) depending on the choice of the initial weight and number of hidden neurons,

3) being sensitive to the learning rate,

4) having poor generalization for complicated nonlinear functions [6-14]. To overcome these drawbacks, a large number of researchers concentrate upon improvements of BP in two aspects. On the one hand, a number of researchers focus mainly on improvements of based on standard gradient descent, including automatically adjusting the learning rate algorithm as training, additional momentum factor, and resilient of BP algorithm etc. On the other hand, a number of researchers focus mainly on based on the standard numerical optimization, including Quasi-Newton algorithm, LM (Levenberg-marquardt) algorithm. For example, the Quasi –Newton training algorithm improved the convergence rate of the standard BP algorithm but requires computing the Hessian matrix, so this leads to a large computational burden and storage expense [1-14]. In this paper we present a comprehensive comparison study on the application of different BP algorithm in time series of deformation forecasting, four types of typical improved BP algorithm, namely, momentum, conjugate gradient, Quasi-Newton and Levenberg-Marquardt algorithms, are investigated.

The remaining paper is organized as follows. In Section 2, the fundamentals of standard back propagation are introduced and discussed. Based on the concepts in Section 2, the improved BP algorithms are presented in Section 3. Then, an illustrative example of high-rise building settlement deformation is adopted to demonstrate the adaptability and effectiveness of the improved BP algorithms in Section 4. Finally, Section 5 concludes the paper.

## 2 Standard back propagation algorithm (SDBP)

Multilayer feed forward BP network is one of the most popular techniques in the field of ANN. Standard back propagation is the generalization of the Widrow-Hoff learning rule to multilayer networks and nonlinear differentiable transfer functions. The common topology of a BP neural network model is illustrated in Figure 1. Input vectors and the corresponding target vectors are used to train a network until it can approximate a function, associate input vectors with specific output vectors, or classify input vectors in an appropriate way as defined by you [1-5].

In general, the BP algorithm includes the forward and the backward process. In the forward process, a vector is added to the input layer, which is then spread along the network, finally, an output vector is

* ***Corresponding author*** e-mail: gaocaiyungao@163.com

obtained as a response of the input vector, in which the synaptic weights cannot be changed.

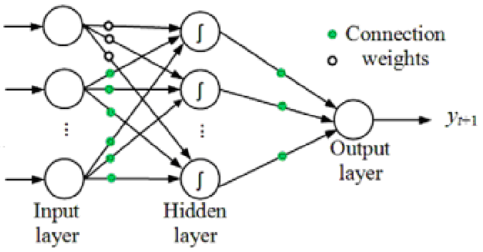


FIGURE 1 The structure of BP neural networks

Then the backward process, an error signal will be obtained by comparing the output signal with the defined output; the error signal is then forward-spread to modify the weight from one output layer to another. The forward and the backward process alternate and constantly circulate, then will be convergent with the defined output in some states. The BP algorithm is carried out as follows:

$$\hat{y}_k(t) = \sum_{j=1}^{p} v_{jt} \cdot f[\sum_{i=1}^{m} w_{ij} \cdot y_i(t) + \theta_j] + r_t \,, \tag{1}$$

where, $f$ is activation function, $i$ and $j$ are the number of neurons of hidden layer and output layer, respectively, $y_i$ is input vector, $w_{ij}$ and $v_{jt}$ are weights between the input/hidden layers and hidden/output layers, respectively, $\theta_j$ and $r_t$ are the bias of neurons. Let $y_k(t)$ be desired output of neural network. There is an error between actual output and desired output, this error, named mean square error, can be expressed as follows:

$$E = \frac{1}{2}\sum_{k=1}^{N_1}\sum_{t=1}^{n}[y_k(t) - \hat{y}_k(t)]^2 \,. \tag{2}$$

Standard back propagation is a gradient descent algorithm, as is the Widrow-Hoff learning rule, in which the network weights moved along the negative of the gradient of the performance function. The updating rules are as follows:

$$x(k+1) = x(k) - ag(k) \,, \tag{3}$$

$$g(k) = \frac{\partial E(k)}{\partial x(k)} \,. \tag{4}$$

where, $\Delta x(k)$ is weight and bias matrix from the input layer to the hidden or the hidden layer to the output layer at the *k-th* learning step, $a$ is the learning rate, $g(k)$ is the grads of E at the *k-th* learning step, $E(k)$ is the error function of the network at the *k-th* learning step, $k$ is the step of training iteration.

## 3 Improved back propagation algorithm

There are a number of variations on the standard back propagation algorithm that are based on other optimization techniques, such as momentum back propagation, conjugate gradient and Newton methods etc. In this section, describes the learning procedures of several improved back propagation algorithm [7-12].

### 3.1 MOMENTUM BACK ALGORITHM (MOBP)

Gradient descent with momentum, allows a network to respond not only to the local gradient, but also to recent trends in the error surface. MOBP depends on two training parameters: the learning rate and the amount of momentum, momentum is set between 0 (no momentum) and values close to 1 (lots of momentum).In SDBP, the learning rate $a$ has a small value because it can decrease with a change of weighted value at the learning step. Consequently, the learning becomes very slow. The MOBP can accelerate the learning step of BP. The learning method of the MOBP is the same to BP-ANN, but it introduces additional momentum factor $\eta$, $0<\eta<1$.

The change of weighted value at the *k+1-th* learning step of MOBP can be expressed as follows:

$$\Delta x(k+1) = \eta \Delta x(k) + \alpha(1-\eta)\frac{\partial E(k)}{\partial x(k)} \,, \tag{5}$$

where, $\Delta x(k)$ is the change of weighted values at the *k-th* learning step. Therefore, the relevant weighted value $x(k+1)$ is given as:

$$x(k+1) = x(k) + \Delta x(k+1) \,. \tag{6}$$

### 3.2 CONJUGATE GRADIENT BACKALGORITHM (CGBP)

The SDBP algorithm adjusts the weights in the steepest descent direction. This is the direction in which the performance function is decreasing most rapidly, however, this does not necessarily produce the fastest convergence. To produces generally faster convergence than steepest descent directions, the search of MOBP algorithm is performed along conjugate directions. All of the conjugate gradient algorithms start out by searching in the steepest descent direction on the first iteration

$$p(0) = -g(0) \,. \tag{7}$$

A line search is then performed to determine the optimal distance to move along the current search direction:

$$\mathrm{x}(k+1) = x(k) + ap(k) \,, \tag{8}$$

$$\mathrm{p(k) = -g(k)} + \beta\mathrm{(k)p(k\text{-}1)} \tag{9}$$

where, $p(k)$ is the search direction at *k+1-th* iteration, the constant $\beta$(k) is computed by various versions of conjugate gradient such as Fletcher-Reeves update, Polak-Ribiére update, Powell-Beale Restarts, Scaled Conjugate

Gradient. Take Fletcher-Reeves of conjugate gradient for example, the formula is follows:

$$\beta(k) = \frac{g^T(k)g(k)}{g^T(k-1)g(k-1)} . \tag{10}$$

3.3 QUASI-NEWTON ALGORITHM (QNBP)

Newton's method is an alternative to the conjugate gradient methods for fast optimization. The basic step of Newton's method can be expressed as follows:

$$x(k+1) = x(k) - A^{-1}(\mathrm{k})g(k) , \tag{11}$$

where, $A(k)$ is the Hessian matrix (second derivatives) of the performance index at the current values of the weights and biases but is often complex to compute. There is a class of algorithms are called quasi-Newton (or secant) methods which is based on Newton's method, but doesn't require calculation of second derivatives. The most successful in studies is the Broyden, Fletcher, Goldfarb, and Shanno update.

3.4 LEVENBERG-MARQUARDT ALGORITHM (LMBP)

Levenberg-marquardt is the fastest method for training moderate-sized feed forward neural networks. It was designed to approach second-order training speed without having to compute the Hessian matrix. The LM algorithm's update formula is given as:

$$x(k+1) = x(k) - (J^{\mathrm{T}}J + \mu J)^{-1}J^{\mathrm{T}}\mathrm{e} , \tag{12}$$

where, $J$ is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases, and is a vector of network errors.

## 4 Experimental results

To demonstrate the effectiveness of the SDBP and improved BP models, we use deformation time series of a high-rise building as an illustrating example, the data in Figure 2 quoted from reference literature [15]. In this study, we use the historical deformation from 1 to 36 as our research data. There are 36 observations, where 1-30 are used for model fitting and 31–36 are reserved for testing.

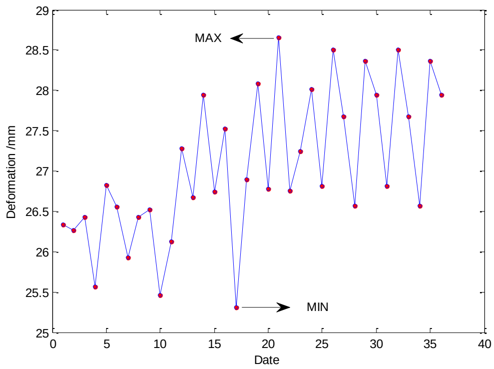


FIGURE 2 The deformation data set

4.1 COMPARISON BETWEEN CONVERGENCE OF FIVE BP MODELS

It has showed in Section 2 and 3 that the BP network has great differences in learning rate, convergence speed and iteration times under the different learning algorithm. For prediction of deformation time series, we can't make sure that in which learning algorithm, the forecast effect of BP network could reach the optimum, and therefore, it is necessary to discuss how the BP networks influence prediction accuracy in different learning algorithm. In calculation, set the preconditions as follows:

The neural network consists of three layers, i.e., input layer, hidden layer, and output layer. The neuron number of the input layer is selected to be 5, and the input to the input layer is a vector of the known historical deformation values, which are upgraded constantly by a fixed-size sliding data window according to the delay deformation time obtained newly. Through the optimization of one-dimensional region search algorithm of BP network [5], the neuron number of the hidden layer is selected to be 13, the output layer has only one neuron, and its output is just the predicted deformation time delay. The activation function of a node is a sigmoid function. Figure 3-7 shows the convergence curve simulated by matlab.

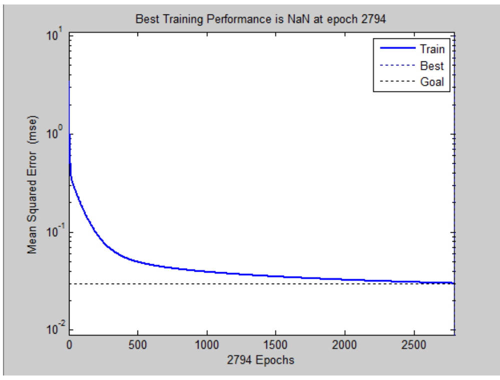


FIGURE 3 Effect of training speed according to SDBP

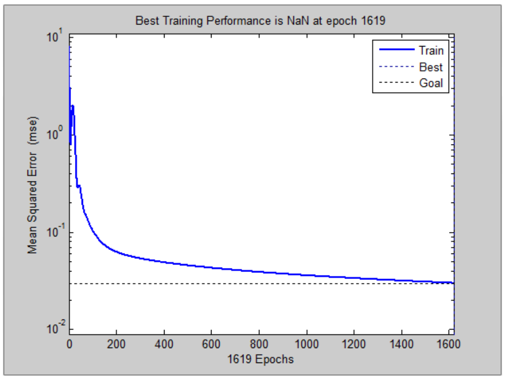


FIGURE 4 Effect of training speed according to MOBP

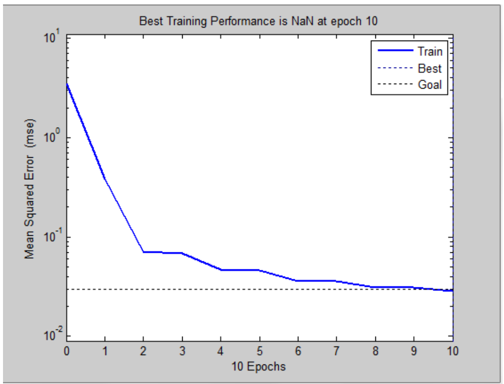


FIGURE 5 Effect of training speed according to CGBP

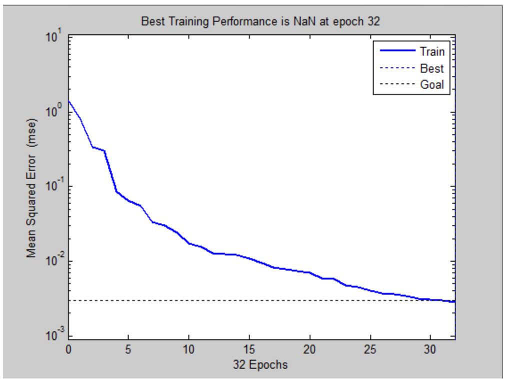


FIGURE 6 Effect of training speed according to QNBP

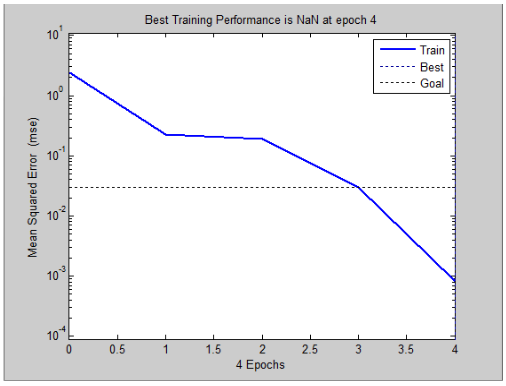


FIGURE 7 Effect of training speed according to LMBP

From the results, it is seen that the training convergence speed of the MOBP, CGBP, QNBP and LMBP are obviously faster than that of the SDBP. It is obvious that the convergence speed is greatly improved.

### 4.2 COMPARISON BETWEEN THE PREDICTION OF FIVE BP MODELS

Data from 1 to 30 is simulated and data of 31and 36 is the prediction. Then, error of simulation and prediction can be calculated, and the results are in Tables 1 and 2. The simulation and prediction results show that improvement BP models are higher than SDBP.

TABLE 1 Simulation error of five BP models (units: mm)

| | SDBP | MOBP | CGBP | QNBP | LMBP |
|---|---|---|---|---|---|
| AAE | 0.6009 | 0.5177 | 0.5273 | 0.5384 | 0.3049 |
| ARE | 0.0224 | 0.0193 | 0.0195 | 0.0199 | 0.0114 |

(AAE: Average absolute error; ARE: Average relative error)

TABLE 2 Prediction error of five BP models (units: mm)

| Date | SDBP | MOBP | CGBP | QNBP | LMBP |
|---|---|---|---|---|---|
| 31 | 0.0691 | 0.4867 | 0.4220 | 0.4324 | 0.1359 |
| 32 | 0.1605 | 0.0299 | 0.3491 | 0.4808 | 0.3398 |
| 33 | -0.2013 | 0.1906 | 0.1445 | 0.1548 | -0.008 |
| 34 | -0.5593 | 0.2535 | 0.0449 | 0.1042 | 0.5342 |
| 35 | -0.2878 | 0.2827 | 0.1984 | 0.5736 | 0.2093 |
| 36 | 0.7427 | 0.39240 | 0.4916 | 0.2803 | 0.1029 |
| AAE | 0.3368 | 0.2727 | 0.2751 | 0.3377 | 0.2216 |
| ARE | 0.0122 | 0.0099 | 0.0099 | 0.0121 | 0.0081 |

(Note: AAE, Average Absolute Error; ARE: Average Relative Error)

## 5 Conclusions

From the results in this study, the following three conclusions can be drawn:

1) ANN is a jumped-up interceptive subject, using it to predict deformation time series, it is feasible and practically.

2) The improvement BP learning algorithm is much better than standard BP algorithm for deformation time series prediction. The LM algorithm's convergence rate is the quickest one and has higher prediction precision.

3) The BP neural network under the different learning algorithm has different characteristics in network training, when we use BP neural network to do the modelling prediction, we should select the best learning algorithm to set up the network according to the actual situation of prediction problem.

## Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (41071328)and Key Laboratory of Mine Spatial Information Technologies of SBSM (State Bureau of Surveying and Mapping) (KLM201306).

## References


[1] Xu H, Li G 2005 Application of BP neural network to observed data processing for dam based on matlab *Engineering Journal of Wuhan University* **38**(3) 50-4

[2] Wu Y P, Teng W F, Li Y W 2007 Application of grey neural network model to landslide deformation prediction *Journal of Rock Mechanics and Engineering* **26**(3) 632-6

[3] Peng X G, Wang J, Sun L H 2009 Application of optimal dimension modeling for BP network in prediction of times of deformation *Journal of Geodesy and Geodynamics* **29**(6) 113-6

[4] Cao Y B, Yang E C, Xie L F 2012 Study of landslide deformation prediction based on gray model-evolutionary neural network model considering function of environmental variables *Rock and Soil Mechanics* **33**(3) 847-52

[5] Gao C Y, Gao N 2012 Comparison with RBF and BP neural network used in deformation prediction *Yellow River* **34**(7)131-4

[6] Gao N, Cui X M, Gao C Y 2012 An effective hybrid approach for processing deformation monitoring data *Advanced Materials Research* 446-9, 3247-51

[7] Zhao X Y, Lai K S, Dai D M 2007 An improved BP algorithm and its application in classification of curface defects of steel plate *Journal of iron and steel research* **14**(2) 52-5

[8] Hu Q C, Hu B, Jiang H F 2013 Application of BP artifical neural network to the displacement prediction of deep foundation pile *Safety and Environment Engineering* **20**(3) 154-8

[9] Yi Q L, Zeng H E, Huang H F 2013 Reservoir landslide deformation forecast using BP neural network *Hydrogeology and Engineering Geology* **40**(1) 124-7

[10] Miao X Y, Chu J K, Du X W 2011 Application of LM-BP neural network in predicting dam deformation *Computer Engineering and Applications* **47**(1) 220-2

[11] Lera G,Pinzolas M *IEEE Transactions on Neural Networks* **13**(5) 1200-3

[12] He Z Y, Zheng W 2008 Deformation prediction of deep foundation pit based on bp neural network *Journal of South China University of Technology* **36**(10) 92-6

[13] Gao N, Wu L C, Gao C Y 2006 The problems of GPS height conversion by artificial neural network method *Engineering of Surveying and Mapping* **15**(3) 67-9

[14] Hu W S, Sha Y J 2003 Converting GPS height by a new method based on neural network *Journal of Southeast University(English Edition)* **19**(1) 53-7

[15] Gao C Y, Pan C J, Gao N 2012 An effective combined approach based on GM and AR for deformation analysis of high-rise buildings *Advanced Materials Research* 368-73 2123-7


## Authors


**Cai-Yun Gao, born in October, 1980, Jinzhou County, Hebei Province, China**

**Current position, grades:** Ph. D. student of college of geoscience and surveying engineering, China University of Mining and Technology, Beijing, China.
**University studies:** Master of Engineering in Geodesy and Surveying Engineering from East China Institute Of Technology in China. Bachelor of Engineering in Surveying and mapping engineering from Institutes Of Technology Of Hebei in China.
**Scientific interest:** intelligence algorithm in the application of geodesy, surveying engineering, disaster prediction.
**Publications:** 15 papers.
**Experience:** teaching experience of 8 years, 6 scientific research projects.



**Xi-Min Cui, born in June, 1967, KuanDian County, LiaoNing Province, China**

**Current position, grades:** the professor of College Of Geoscience and Surveying Engineering, China University of Mining and Technology, Beijing.
**University studies:** Doctor of Engineering from China University of Mining and Technology (Beijing) in China.
**Scientific interest:** deformation monitoring and surveying and mapping data processing.
**Publications:** 50 papers.
**Experience:** teaching experience of 30 years, 15 scientific research projects.