

The neural network model to solve the pre-consolidation stress

Ran An^{1,2, a}, Ling-wei Kong^{2, b} and Cheng-sheng Li^{1,2, c}

¹State Key Laboratory of Geomechanics and Geotechnical Engineering, Institute of Rock and Soil Mechanics, Chinese Academy of Sciences, Wuhan 430071, China

²University of Chinese Academy of Sciences, Beijing, 100049, China

^aarwhrsm@163.com, ^blwkong@whrsm.ac.cn, ^clichengsheng@outlook.com

Keywords: pre-consolidation stress, BP artificial neural network, e - p curves, MATLAB

Abstract: Pre-consolidation stress is an indicator representing stress history of soil and an important parameter which reflects the deformation characteristics of soil. Finding an accurate and easy way to solve pre-consolidation stress is of great significant in the study of engineering construction. Based on the previous studies on solutions of the pre-consolidation stress and analysis of the artificial neural network theory, a new approach – the BP neural network model to solve pre-consolidation stress is proposed. Based on the platform of MATLAB, the BP neural network model is established and trained by random data samples from compression tests. By operating three different algorithms in computing analysis, the L-M algorithm is identified as the optimized to be applied in the model. Using the neural network model with training completion, the output port can export rapid and accurate inversions of the predicting pre-consolidation stress. The forecasting errors are greatly reduced compared to the Casagrande method and the numerical plate method according to error analysis of these three methods. Therefore, the BP artificial neural network model to solve the pre-consolidation stress is proved to have a good feasibility and a promotional value in real engineering.

1. Introduction

Pre-consolidation stress is the maximum effective stress that a soil has suffered throughout its history. It reflects the stress history of soils [1]. According to the value of pre-consolidation stress, the consolidation condition of soil can be divided into three types which are over-consolidation, normal consolidation and under-consolidation. Pre-consolidation stress is of great significant in predicting expected settling volume of foundations or embankments because underestimated pre-consolidation may cause overestimation of the magnitude of consolidation settlement and make more expensive and time-consuming geotechnical solutions [2]. Therefore, it is vital to find an accurate and easy solution to determine pre-consolidation stress.

Pre-consolidation stress is normally calculated by the data from a uniaxial confined compressive stress test. And the data from tests are plotted on logarithmic coordinates with the normal effective stress (p) against Cartesian coordinates with void ratio (e). Many researchers have proposed methods to estimate the pre-consolidation stress. The most widely used method is called ‘Casagrande method’ (Casagrande (1936)). [3] Other researchers, such as Van Zelst (1945), Schmertmann et al. (1955), Gregory et al. (2006) and Liu Yong-hai et al [4~7] have proposed their improving methods after Casagrande. The flaw of ‘Casagrande method’ is during solving process, the spot of the maximum curvature is hard to further exact, as a consequence to influence the acquisition of values of pre-consolidation stress. Also, the procedure of graphing method is complicated. Most of the existing solutions are fitting e - P and e - $\log p$ curves or using other mathematic models to locate the spot of the maximum curvature. Essentially, these measures are the refinements to ‘Casagrande method’ and they still cannot avoid errors from the process of fitting e - $\log p$ curves. As a result, the pre-consolidation stress cannot be acquired precisely via these conventional methods. Therefore, this paper presents a brand new thought—the BP neural network predicting model to calculate pre-consolidation stress which is based on a dominant function to reduce the initial error with fitting curves.

Artificial neural network (ANN) was proposed and developed on the basis of modern neuroscience [8]. It imitates the human brain to discover the interaction between inputs and outputs through

training. It has the ability of self-adjusting and self-organizing so as to complete the specific targeting tasks by 'supervised learning'. The most valuable function of the ANN is its ability of self-learning and acquiring the satisfactory forecasting results even though the data is not complete précised.

The network structure and working style of the ANN can be divided into two major categories, which are the feed-forward network and the feedback network [9]. One of the most widely-used network structures is the BP neural network (also known as back-propagation network), which is a kind of tutor-guiding feed-forward network. BP neural network is an iterative gradient algorithm to solve minimum mean square difference between actual outputs of forward network and expected outputs [10]. BP algorithm is a kind of δ algorithm, which is a supervised studying algorithm [11]. It has the ability to take all kinds of factors of extremely complex nonlinear problems into consideration and solve these synthetically via adjusting the weight of every signal node by its self-organizing system. Hence, the BP neural networks can be regarded as a highly nonlinear mapping between inputs and outputs.

As the Fig.1 shows, the system will use the interpolations between studying samples ($P_1, P_2, P_3 \dots$) and expected vector ($T_1, T_2, T_3 \dots$) to amend their weights, in order to narrow the gaps between the output vector (A_1, A_2, A_3, \dots) and expected vector ($T_1, T_2, T_3 \dots$), which means to make quadratic sum of error from the network output layer to minimize (Fig.1). In order to approaching the target, it calculates the weights of networks and changes of errors which relative to the slope direction of error function. The algorithm consists of two parts: the forward pass of information and the backward of errors. During the forward pass, input information pass to the output layer via the process from inputting to calculating by layer because the state of each previous neuron only influence its following one. If the desired outputs do not appear in the output layer, the variation values of errors will be calculated and then pass backward along the original linking path to modify the weights of each neuron until a desired goal appears.

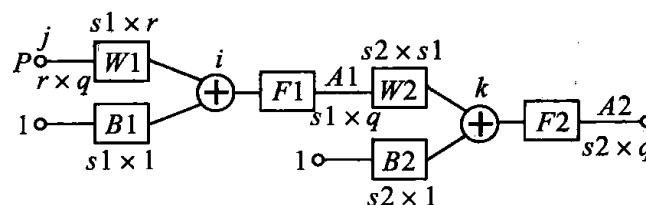


Fig.1 The structure of BP artificial neural network

2 Establishment of the Model

2.1 The determination of input and output items

Previous researchers commonly regarded the influential relative parameters as input signals and evaluation targets as output results. Directly using the internal relation curves in artificial neural networks to analysis and forecast is unusual. This paper do not use parameters of soil, such as: natural weights, moisture content, natural void ratio, coefficient of consolidation, coefficient of compressibility, as input signals. Therefore it uses the trend of $e-p$ curves in the artificial neural networks to analyze directly for reducing the information loss in the process of analysis.

Every different stress of soil has a corresponding void ratio during the compression tests. With the increase of consolidation stress P , the decline of the void ratio e is similar to the negative exponential type (or inverse proportional function), and the $e-p$ curve is leveling off (Fig.2). If taking the density of distribution as ordinate, the void ratio frequency distribution curve will be obtained (Fig.3).

Bring $e-p$ curves into the neural network model. Use the weights in the model to complete the recognition processing for void ratios e with different densities of distribution. Then continually improve the recognition ability by training of objective supervision to provide feedback void ratios e with target characteristic. Therefore, this model takes void ratios under different stresses as input

items, and void ratios under experimental pre-consolidation stress as target output items. Finally, use linear interpolation invert the relative pre-consolidation stress P according to forecasting void ratios e .

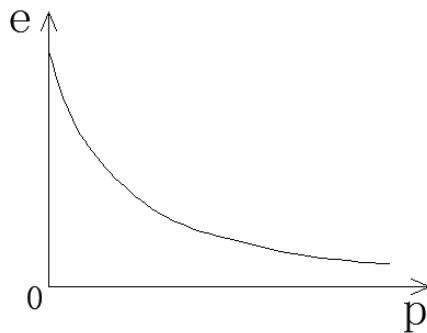


Fig.2 e - p curve

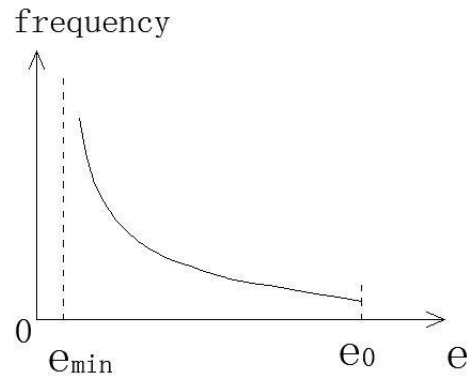


Fig.3 The frequency distribution curve of void ratio

2.2 Determination of the Number of Hidden Layers

Choosing the number of hidden layers and the number of hidden layer unit is one of the difficulties of model designs. If the nodes with hidden layers are not enough, the generalization ability of a network may be weak; if the nodes with hidden layers are overmuch, the time may need to add and the training error also may not be the smallest meanwhile. It is commonly to use empirical formula or trials to determine the number of hidden layers l .

$$l = \sqrt{m + n} + a \quad (1)$$

$$\text{or } l = (0.43mn + 0.12n^2 + 2.54m + 0.77n + 0.86)^{\frac{1}{2}} \quad (2)$$

Where n is the number of output nodes, m the number of input nodes, a is constant from 1 to 10.

2.3 The Determination of Training Function

In the BP neural network training process, the conjugate gradient algorithm, the Quasi-Newton algorithm and the Levenberg - Marquardt algorithm (L - M algorithm) are often being applied. The following model analysis will prove the L - M algorithm has the best effect to calculate.

The L - M algorithm is based on the algorithm of standard numerical technique which replaces mean square errors to square errors. Its iterative arithmetic needs to calculate the inverses of matrixes. And it is the fastest algorithm of all the training neural networks so it make the model has the advantages of a fast rate of convergence and a high robustness .

Using the L - M algorithm to solve nonlinear least squares problems is quite efficient, its square error formula includes the values of weights and thresholds under each layers in BP networks. The formula is shown as Eq.3 and Eq.4.

$$E = \frac{1}{2} \sum \sum (d_{ki} - O_{ki})^2 = \frac{1}{2} \sum_k (\varepsilon^k)^2 = \frac{1}{2} \|\varepsilon\|^2 \quad (3)$$

In Eq.3 and Eq.4, E is sum of squares of errors, d_{ki} is the expected output value in the system, O_{ki} is the real output of neuron k under the output layer, ε is the vector with ε^k as an element.

The formula of error function can be obtained by using the Jacobian matrix J . The following formula is used to calculate the sum of squares of errors between iteration P and iteration $(P+1)$.

$$E = \frac{1}{2} \|\varepsilon(X^P) + J(X^{P+1} - X^P)\|^2 + \mu \|X^{P+1} - X^P\|^2 \quad (4)$$

This paper will compare L- M algorithm with conjugate gradient method and the Newton method. And the result is : among the three algorithms of BP neural network, the L - M algorithm shows a highest rate of convergence and a highest precision.

3. Evaluation of the Model

3.1 Model Training

The neural network model becomes relatively complex if the number of hidden layer is more than 2. Therefore, this model uses a single-layer model. The l is number of nodes in every layer of the model. According to Eq.2, setting l as 4.54 by calculating, which means the number of node is 4 approximately. Make the void ratios under different stress as input items and pre-consolidation stresses under different stress as output items. Also, make void ratios under different stress in experiments as input items and corresponding pre-consolidation stresses under different stress as output items. The process of sample training is shown below.

- (1) Input the data of 5 group training samples and a set of forecasting and apply the normalized processing;
- (2) Initialize the thresholds and the weights of the neural network, preliminary select the transfer functions and the number of hidden layer nodes
- (3) Set expectations of error and the times of calculation
- (4) Make simulation calculations and compare their errors. If the expectations are not achieving, then go back to step 2 and amend the parameters, passing functions and add or reduce the number of nodes in order to reach the expectations.
- (5) Apply different algorithm types to make comparison for the trained neural networks, and observe the convergence speeds and error performances meanwhile.
- (6) Choose the optimized algorithm and use it to forecast for the target data.
- (7) End the training process.

The experimental data is shown in Table 1, and the curves of compression tests are shown in the Fig.4. The group A、B、C、D and F are study groups and E is forecasting group. The pre-consolidation stresses are solved precisely by previous researchers, and the corresponding void ratios are calculated via linear interpolation.

Table.1 Data of compression tests of soil

Test number	A	B	C	D	E	F
Stress (p /kPa)	Void ratio e under different stresses					
25	1.023	0.912	0.835	0.703	0.643	0.600
50	1.019	0.903	0.810	0.700	0.642	0.598
100	1.005	0.881	0.777	0.692	0.636	0.593
200	0.933	0.838	0.728	0.682	0.620	0.587
400	0.853	0.775	0.674	0.663	0.577	0.574
800	0.768	0.707	0.604	0.604	0.528	0.529
1600	0.680	0.616	0.529	0.529	0.480	0.479
3200	0.589	0.525	0.449	0.449	0.431	0.430
Predicting e	1.016	0.839	0.725	0.683	0.618	0.571

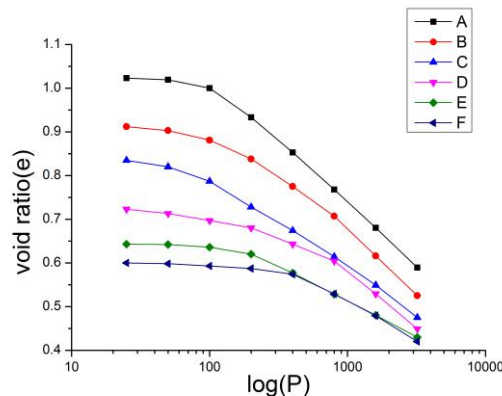


Fig.4 e -log p curves compression tests of soil

3.2 Evaluation of different algorithms

3.2.1 Conjugate gradient method:

The training is ended after 10 times calculation and errors cannot reach the expectation. The result is shown in the Fig.5 and Fig.6.

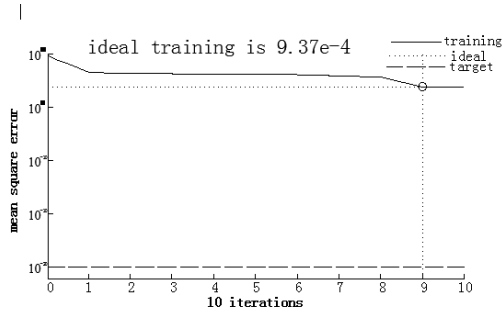


Fig.5 The algorithm error analysis of Conjugate gradient method

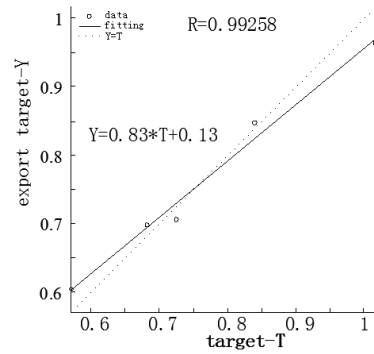


Fig.6 Algorithm regression analysis

3.2.2 Quasi-Newton method

The target error is obtained after 85-times training. The result is shown in the Fig.7 and Fig.8.

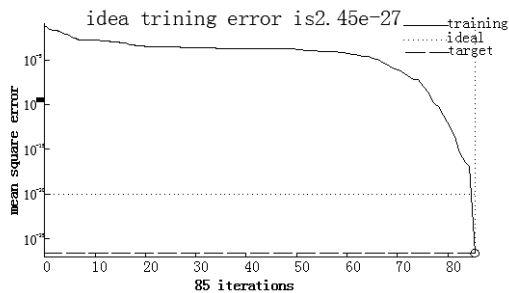


Fig.7 The algorithm error analysis of Quasi-Newton method

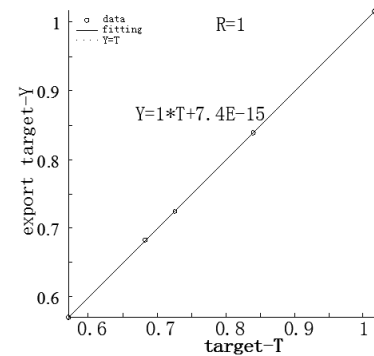


Fig.8 Algorithm regression analysis

3.2.3 L-M method

The target error is obtained after 49-times training. The result is shown in the Fig.9 and Fig.10.

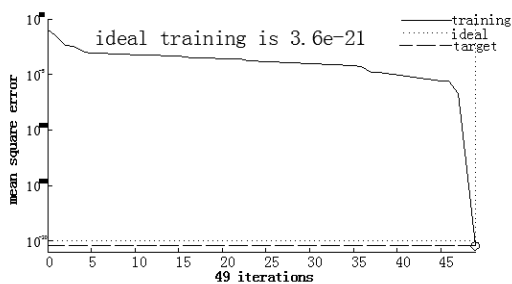


Fig.9 The algorithm error analysis of L-M method

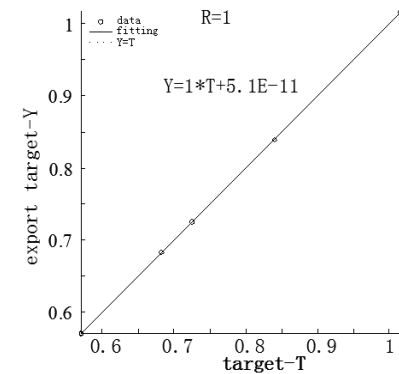


Fig.10 Algorithm regression analysis

Compared the results of three methods, it is found that L-M algorithm has a faster rate of convergence and lower output errors compared to conjugate gradient method and quasi-Newton method. Therefore, L-M method is applied in the following analysis of pre-consolidation stresses. The major network parameters are:

```
net=newff(minmax(P),[4,1],{'softmax','purelin'},'trainlm');
inputWeights=net.IW{2,1};
inputbias=net.b{2};
layerWeights=net.LW{2,1};
layerbias=net.b{2};
net.trainParam.show = 10;
net.trainParam.lr = 0.08;
```

```
net.trainParam.mc = 0.9;
net.trainParam.epochs = 1000;
net.trainParam.goal = 1e-20;
net.trainParam.min_grad=1e-15;
```

The network training function using 'trainlm' function, and in this model the 'trainlm' uses the L-M algorithm. When modifying the approximated second-order training rate, the designed L-M algorithm can make the learning time shorter. For complex problems, this method in the running process needs to consume larger amounts of resources, but the precision is higher. The network performance function uses the mean square error (mse) function. Transfer function of output layer uses the Logsig function

3.3 Model validation

Calculating Results: set group E as the forecasting group and others as study groups. The model has taken 10-times calculations and each calculation has 50 to 100 steps, which means the rate of convergence is quite fast. After ten-times calculations, the forecasting value of void ratio e in the forecasting group is 0.6185. The pre-consolidation stress is calculated via using the linear interpolation after obtaining the value of e . The value of pre-consolidation stress is forecast as 206.79kPa. And the relative error between the accurate value (209kPa) and forecasting value (206.79kPa) is 1.06% (Fig.11). To improve the accuracy, it is recommended to takes more times of calculations.

Similarly, rotating the another group of data to predict the pre-consolidation stress in turn and leaving other groups as training groups to judge the calculation stability of the model. Then repeat the operations of the neural network model to obtain a predicting value with a relatively small error. Meanwhile, the Casagrande plotting method and the numerical plate method are used to estimate the pre-consolidation stress of the compression test data. Compared their accuracy and stability with the neural network model by analyzing the values of errors in this paper, the advantage of the neural network model is easy to perceive. The result is shown in the Fig.12. It can be seen that the errors of neural network model are significantly smaller than other methods. It is shown in the Table.3 that the mean value and the sample variance of neural network model are both the smallest, which reflects its applicability to solve pre-consolidation stresses is much better than traditional methods.

From what has been discussed above, using BP neural network model to predict pre-consolidation stress has a great advantage in the aspect of using data. If the original data can be classified according to different soil properties in the case of data source is relatively large, the model can improve its accuracy in predicting. The pre-consolidation stresses with different levels are predicted using the corresponding set of data groups. Making full use of the previous test data as learning group data sources can predict pre-consolidation stress more conveniently and accurately.

Table2. Results and errors of 10-times computations

Computation times	1	2	3	4	5
Predicting void ratios	0.6198	0.6173	0.6213	0.6116	0.6168
Predicting pre-consolidation stress (kPa)	200.9013	212.4324	191.7411	238.8581	214.6968
Error (%)	-3.87498	1.642294	-8.25785	14.28619	2.725718
Computation times	6	7	8	9	10
Predicting void ratios	0.6203	0.6192	0.6167	0.6121	0.6185
Predicting pre-consolidation stress (kPa)	198.1083	203.8405	215.4016	236.5542	207.0721
Error (%)	-5.21136	-2.46864	3.062952	13.18384	-0.92243

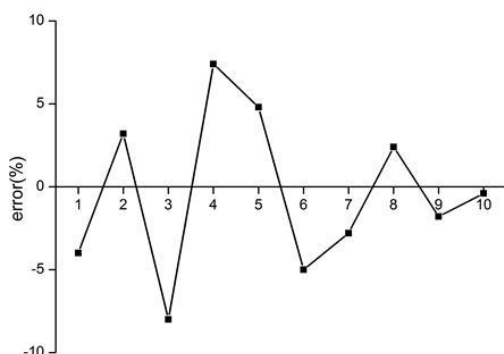


Fig.11 The variation trend of forecasting errors

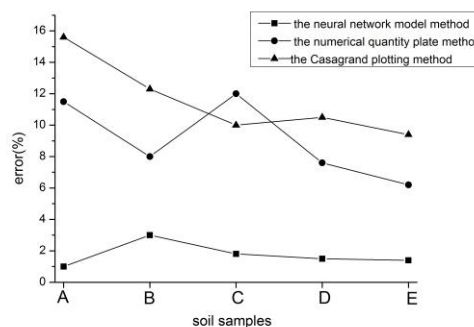


Fig.12 The forecasting errors of three methods

Table3. The error analysis of three different methods

Methods	Mean value(%)	sample variance
Neural network model	1.974	0.287
Casagrande plotting method	12.321	0.917
Numerical plate method	9.462	0.840

4. Conclusion

This paper establishes an artificial intelligent model to solve pre-consolidation stress based on BP artificial neural network by using the MATLAB platform. The BP neural network model successfully predicts the pre-consolidation stress in practical engineering via ‘training’ the system. After comparing the results with previous methods, it is proved that the BP neural network model can accurately predict pre-consolidation stress. The success to solve pre-consolidation stresses by the new method means the artificial neural network model has a good applicability in calculations of rock and soil mechanics.

Acknowledgements

This work was financially supported by the mechanical properties of decay characteristics and mesoscopic structure evolution mechanism of Granite residual soil foundation pit engineering, the national natural science fund project of China (No.11672320).

References

- [1] Yang Xiujuan, Jia Yonggang, Liu Hongjun, and Shan Hongxian: Characteristics and Causes of the Preconsolidation Stress of Soils in the Yellow River Delta[J]. Journal of Ocean University of China, 2009, (03): 215-221.
- [2] Dawidowski, J. B., and Koolen, A. J: Computerized termination of the precompression stress in compaction of field core samples [J]. Soil Tillage Res, 1994, (31): 277-282.
- [3] Casagrande A. The Determination of the Pre-consolidation Load and Its Practical Significance[J]. Proc. of First ICMFE, 1936(3):60—64.
- [4] Van Zelst, T.W., An investigation of the Factors Affecting Laboratory Consolidation of Clay, Proc. of 2nd ICSMFE, Vol.7, 1945, pp.52~61.
- [5] Schmertmann J H: The undisturbed consolidation behavior of clay[J]. Transactions of ASCE, 1955, 120(2): 1201—1226.

- [6] AS Gregory, WR Whalley, CW Watts, NRA Bird, PD Hallett, AP Whitmore: Calculation of the compression index and precompression stress from soil compression test data [J] . Soil & Tillage Research, 2006, 89: 45 ~ 57.
- [7] Liu Yong-hai, Zhu Xiang-rong, Chang Lin-yue: Determining pre-consolidation pressure by mathematic analysis based on Casagrande method [J]. Rock and Soil Mechanics 2009,(01):211-214+220
- [8] Dong, X., Wang, S., Sun, R. et al. Design of artificial neural networks using a genetic algorithm to predict saturates of vacuum gas oil [J]. Petroleum Science, 2010,(01):118-122.
- [9] Zhiguo Zhang: Study on artificial neural networks and their applications in Geoscience [D]. Jiling University, 2006
- [10] Real estate appraisal system based on GIS and BP neural network[J]. Transactions of Nonferrous Metals Society of China,2011,(S3):626-630
- [11] A new neural network model for the feedback stabilization of nonlinear systems [J]. Journal of Zhejiang University (Science A: An International Applied Physics & Engineering Journal),2008,(08):1015-1023.