Neural Network Forecast Model in Deep Excavation

J. C. Jan[1]; Shih-Lin Hung, M.ASCE[2]; S. Y. Chi[3]; and J. C. Chern[4]

Abstract: Diaphragm wall deflection is an important field measurement in deep excavation. The monitoring data are applied to evaluate the construction performance to avoid a supporting system failure or damages incurred to adjacent structures. Despite the numerous case histories of construction projects and several forecasting methods, no method accurately forecasts the performance of construction due to the complicated geotechnical and construction factors affecting the behavior of the diaphragm wall. This work predicts the diaphragm wall deflection by using the adaptive limited memory–Broyden-Fletcher-Goldfarb-Shanno supervised neural network. Eighteen case histories of deep excavations with four to seven excavation stages are selected for training and verification. In addition, the knowledge representation adopts measured wall deflections of previous excavation stages as inputs to the network. Doing so substantially reduces the importance of soil parameters, which are often extremely fluctuating and difficult to assess. Simulation results indicate that the artificial neural network can reasonably predict the magnitude, as well as the location, of maximum deflection of the diaphragm wall.

DOI: 10.1061/(ASCE)0887-3801(2002)16:1(59)

CE Database keywords: Neural networks; Excavation; Geotechnical engineering; Sensitivity analysis; Algorithm; Diaphragm wall.

Introduction

In urban areas, braced diaphragm walls are generally applied to deep excavations in soft soil. A satisfactorily braced diaphragm wall not only provides for a safe excavation but also minimizes deformations in the surrounding ground, which are of utmost importance in avoiding costly damages to adjacent buildings. Observational methods are frequently employed in deep excavation projects to ensure safe construction. Peck (1969b) first compiled the observational data in deep excavations and tunneling in soft grounds. By adopting various construction methods, he summarized the feasibility of excavations. Peck (1969a) also recommended the observational method not be adopted unless the designer has a plan of action for every unfavorable situation that may arise.

The estimation of lateral wall deflections and ground settlements has received substantial attention from practicing engineers and researchers. Finite element analyses are extensively applied to estimate wall deflections in deep excavations. Clough and Hansen (1981) demonstrated how anisotropy clay affects braced wall systems by performing finite element analyses. Powrie and Li (1991) employed finite element analyses to investigate the effects of soil/wall/prop stiffness and the preexcavation earth pressure coefficient. Hashash and Whittle (1996) conducted a series of numerical experiments, that applied nonlinear finite element analyses to investigate how the embedment length, support conditions, and stress history profile affect the undrained deformations around a braced diaphragm wall in a deep excavation.

The accuracy of ground movement prediction through finite element analyses heavily depends on the constitutive behavior of the soil. The soil parameters applied in constitutive models are generally obtained from laboratory tests. However, the test results are often not representative of the in situ soil behavior due to factors such as sample disturbance, change of in situ environment and effects of construction. To minimize the effects of soil parameters and construction factors, Gioda and Sakurai (1987) proposed back analysis procedures to obtain modified soil parameters through the fitting of computed wall deformations and field measurements. For example, Whittle et al. (1993) implemented a finite element analysis with an MIT-E3 soil model to simulate the top-down construction of a seven-story, underground parking garage on Post Office Square in Boston, Mass. According to their results, soil deformation measurements provide valuable information on how to evaluate the constitutive model for describing soil behavior. In addition, several mathematical optimization approaches have been utilized to determine modified soil parameters. Ou and Tang (1994) employed optimization techniques to determine soil parameters for finite element analysis in deep excavation. Chi et al. (1999) developed an information construction approach for deep excavation. Their investigation applied an optimization process to calculate back analyzed soil parameters, which were then used to predict the wall deflection of subsequent excavation stages. Although back analysis combined with finite element analysis provides an effective numerical approach for engineers to estimate wall deflections and surface settlements in excavations, the back analysis technique is limited to the number of soil parameters and the soil model applied.

Artificial neural networks (ANNs) form a class of systems that are derived from biological neural networks. Learning is an important feature of artificial neural networks. Several supervised

[1]PhD, Dept. of Civil Engineering, National Chiao Tung Univ., 1001 Ta Hsueh Rd., Hsinchu Taiwan 300, R.O.C.

[2]Associate Professor, Dept. of Civil Engineering, National Chiao Tung Univ., 1001 Ta Hsueh Rd., Hsinchu Taiwan 300, R.O.C. (corresponding author). E-mail: slhung@cc.nctu.edu.tw

[3]Researcher, Geotechnical Researcher Center, Sinotech Engineering Consultants, Inc., Basement No. 7, Ln. 26, Yat-Sen Rd., Taipei Taiwan 105, R.O.C.

[4]Manager, Geotechnical Researcher Center, Sinotech Engineering Consultants, Inc., Basement No. 7, Ln. 26, Yat-Sen Rd., Taipei Taiwan 105, R.O.C.

Note. Discussion open until June 1, 2002. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on May 22, 2000; approved on January 3, 2001. This paper is part of the ***Journal of Computing in Civil Engineering***, Vol. 16, No. 1, January 1, 2002. ©ASCE, ISSN 0887-3801/2002/1-59–65/$8.00+$.50 per page.

and unsupervised neural network learning algorithms have been developed and explored in a number of various domains (Adeli and Hung 1995; Haykin 1994). ANN learning models can effectively deal with qualitative, uncertain, and incomplete information. Therefore, ANN is highly promising for modeling complicated problems in which the governing equations are difficult to define. Flood and Kartam (1994a,b) provided a discourse on the understanding, usage, and potential for application of artificial neural networks in civil engineering. According to their study, artificial neural networks can be implemented in mapping, transitory, and optimization problems, as well as model dynamic processes. Based on supervised learning algorithms, several other researchers have applied neural network learning models in civil engineering (Hajela and Berke 1991; Ghaboussi et al. 1991; Kang and Yoon 1994; Stephen and Vanluchene 1994; Elkordy et al. 1994). In geotechnical engineering, Ni et al. (1996) applied ANN to evaluate failure potential of slopes, and Goh (1994) introduced the application of neural networks to evaluate liquefaction potential. Juang et al. (1999) presented a technique of training ANNs with the aid of fuzzy sets theory. The technique involved modules for preprocessing input parameters and postprocessing network output. They indicated that the fuzzy set-based ANN models could be trained with greater efficiency.

This work attempts to predict the diaphragm wall deflection in deep excavation using a supervised limited memory–Broydon-Fletcher-Goldfarb-Shanno (L-BFGS) ANN learning model (Hung and Lin 1994). The training data are collected from the construction projects in the Taipei basin. For comparison, the conventional finite element analysis, which involves optimization back analysis to calculate soil parameters, is also applied to evaluate these diaphragm wall deflections in deep excavation.

Artificial Neural Networks (ANNs)

The ANNs form a class of systems that are derived from biological neural networks. The topology of an ANN model consists of a number of simple processing elements, called nodes, that are interconnected to each other. Interconnection weights that represent the information stored in the system are used to quantify the strength of the interconnections; these weights hold the key to the functioning of an ANN. ANNs have been used in a broad range of applications, including classification, pattern recognition, function approximation, optimization, prediction, and automatic control. Among the many different types of ANN, the feedforward, multilayered, supervised neural network with the error back-propagation (BP) algorithm, the so-called BP network (Rumelhart et al. 1986), is by far the most commonly applied neural network learning model owing to its simplicity. The architecture of BP networks, displayed in Fig. 1, consists of an input layer, one or more hidden layers, and an output layer.

Before an ANN can be used in the application, it needs to learn or be trained from an existing training set that consists of pairs of input-output elements. The training of a supervised neural network using BP learning algorithm usually involves two stages. The first stage is the data feed forward. The output of each node is defined as

$$net_j = \sum_{i=1}^{n} W_{ij} o_i + \theta_j \tag{1}$$

$$o_j = f(net_j) \tag{2}$$

where W_{ij}=weight associated with the ith node in the preceding layer to the jth node in the current layer; o_i=output of ith node in the preceding layer; θ_j=threshold value of node j in the current layer; o_j=output of node j in the current layer; and function f=activation function, which has to be differentiable. Herein, the sigmoid function is used as the activation function and is defined as

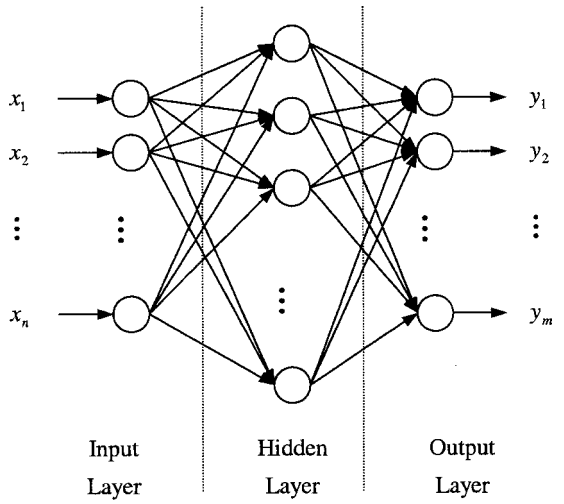

Fig. 1. Feedforward network with one hidden layer

$$f(x) = \frac{1}{1+e^{-x}} \tag{3}$$

The second stage is error back-propagation and adjustment of the weights through the network. In the training process, system error function is used to monitor the performance of the network. This system error function is defined as

$$E = \frac{1}{2} \sum_{p=1}^{P} \sum_{k=1}^{K} (d_{pk} - o_{pk})^2 \tag{4}$$

where P=number of instances in the training set; and d_{pk} as well as o_{pk}=desired and calculated output of the kth output node for the pth instance, respectively. The standard BP algorithm uses a gradient descent approach with a constant step length (learning ratio) to train the network

$$W_{ij}^{(k+1)} = W_{ij}^{(k)} + \Delta W_{ij} \tag{5}$$

$$\Delta W_{ij} = -\eta \frac{\partial E}{\partial W_{ij}} \tag{6}$$

where η=learning ratio, which is a constant in the range of [0,1]. The superscript index k denotes the kth learning iteration.

BP supervised neural network learning models, however, always take a long time to learn. Moreover, the convergence of a BP neural (BPN) network is highly dependent upon the use of a learning rate, η. Thus, several different approaches developed to enhance the learning performance of the BP learning algorithm have been applied. One approach is to develop more effective learning algorithms with the objective of reducing the learning time. Moller (1993) developed a scaled conjugate gradient algorithm for fasting the supervised learning. Adeli and Hung (1994) developed an adaptive conjugate gradient neural network (AdCGN) learning algorithm and applied it to structural engineering. Sanossian and Evans (1995) used a gradient range-based heuristic method for accelerating neural network. Another approach is to develop a parallel algorithm on multiprocesser computers with the objective of reducing the overall computing time. For in-

J. Comput. Civ. Eng. 2002.16:59-65.

stance, Adeli and Hung (1993) presented a concurrent Ad-CGN learning algorithm to a large-scale pattern recognition problem. Significant improvement for the BPN algorithm in computing time was reported in their work. The third approach is the development of hybrid neural network learning algorithms. Hung and Adeli (1994) presented a parallel hybrid genetic/neural network learning algorithm. They reported a superior convergence property of the parallel hybrid neural network learning algorithm as compared with a BPN learning algorithm. Besides, the performance of neural networks in engineering applications can be significantly improved by selecting a suitable representational framework to present the training input/output pattern pairs. Gunaratnam and Gero (1994) discussed the effect of representation of input/output pairs for training instances on the learning performance of the BPN learning algorithm in the problems of structural design. The dimensionless representation is reported to result in a simpler mapping function and makes it possible to train network on a small training set and still have the capability for reasonable accurate predictions.

Hung and Lin (1994) developed a more effective adaptive L-BFGS learning algorithm based on the approach of a L-BFGS quasi Newton second-order method (Nocedal 1980) with an inexact line search algorithm. In the conventional BFGS method, the approximation $\mathbf{H}_{k+1}$ to the inverse Hessian matrix of function $E(\mathbf{W})$ is updated by

$$\mathbf{H}_{k+1}=(\mathbf{I}-\rho_k\mathbf{s}_k\mathbf{y}_k^T)\mathbf{H}_k(\mathbf{I}-\rho_k\mathbf{y}_k\mathbf{s}_k^T)+\rho_k\mathbf{s}_k\mathbf{s}_k^T\equiv\mathbf{V}_k^T\mathbf{H}_k\mathbf{V}_k+\rho_k\mathbf{s}_k\mathbf{s}_k^T \tag{7}$$

where $\rho_k=1/\mathbf{y}_k^T\mathbf{s}_k$; $\mathbf{V}_k=\mathbf{I}-\rho_k\mathbf{y}_k\mathbf{s}_k^T$; $\mathbf{s}_k=\mathbf{W}_{k+1}-\mathbf{W}_k$; $\mathbf{y}_k=\mathbf{g}_{k+1}-\mathbf{g}_k$; and $\mathbf{g}_k=\partial E/\partial\mathbf{W}$. Instead of forming the matrix $\mathbf{H}_k$ in BFGS method, we save the vectors $\mathbf{s}_k$ and $\mathbf{y}_k$. These vectors first define and then implicitly and dynamically update the Hessian approximation using information from the last few iterations, say *m* in the work. Therefore, the final stage of the adjusting weights in a BP-based ANN is modified as

$$\mathbf{W}^{(k+1)}=\mathbf{W}^{(k)}+\alpha_k\mathbf{d}_k \tag{8}$$

The search direction is given by

$$\mathbf{d}_k=-\mathbf{H}_k\mathbf{g}_k+\beta_k\mathbf{d}_{k-1} \tag{9}$$

where $\beta_k=[\mathbf{y}_{(k-1)}^T\mathbf{H}_{(k-1)}\mathbf{g}_{(k-1)}/[\mathbf{y}_{(k-1)}^T\mathbf{d}_{(k-1)}]$

The step length α_k is adapted during the learning process through a mathematical approach—the inexact line search algorithm. This is used in the L-BFGS learning algorithm instead of a constant learning ratio (Hung and Lin 1994). The inexact line search algorithm is based on three sequential approaches—bracketing, sectioning, and interpolation. The bracketing approach brackets the potential step length, α, between two points, through a series of function evaluations. The sectioning approach then uses the two points of the bracket as the initial points, reducing the step size piecemeal, and locating the minimum between points, e.g., α_1 and α_2, to a desired degree of accuracy. Finally, the quadratic interpolation approach uses the three points, α_1, α_2, and $(\alpha_1+\alpha_2)/2$, to fit a parabola to determine the step length, α_k. Consequently, the step length α_k is required to satisfy the following conditions in each iteration (Hung and Lin 1994):

$$E(\mathbf{W}_k+\alpha_k\mathbf{d}_k)\leq E(\mathbf{W}_k)+\beta\alpha_k[\nabla E(\mathbf{W}_k)^T\mathbf{d}_k];$$
$$\beta\in(0,1)\quad\text{and}\quad\alpha_k>0 \tag{10}$$

$$\nabla E(\mathbf{W}_k+\alpha_k\mathbf{d}_k)^T\mathbf{d}_k\geq\theta[\nabla E(\mathbf{W}_k)^T\mathbf{d}_k];$$
$$\theta\in(\beta,1)\quad\text{and}\quad\alpha_k>0 \tag{11}$$

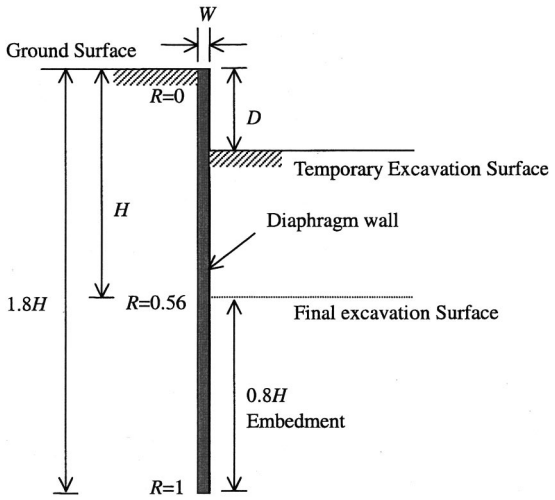

Fig. 2. Illustration of diaphragm wall structure

$$\nabla E(\mathbf{W}_k+\alpha_k\mathbf{d}_k)^T\mathbf{d}_{(k+1)}<0 \tag{12}$$

Hence, the problem of trial and error selection of a learning ratio in the BP algorithm was circumvented in the adaptive L-BFGS learning algorithm.

Knowledge Representation of Wall Deflection Problem

Deep excavations are widely conducted in the construction of underground structures and the foundations of high-rise buildings. Therefore, a large amount of monitoring data has been accumulated. However, due to the complexity of factors that affect the behavior of deep excavation, the information cannot be applied effectively to solve new problems. Herein, supervised ANN with adaptive L-BFGS learning models are adopted to accurately predict the diaphragm wall deflections employing accumulated monitoring data. The measurement data and wall deflections from deep excavations, are collected for training purposes. The underlying notion of applying ANN model to predict staged construction problems is that during an excavation, an accurate prediction of the succeeding stage can be derived from the information of two or more previous stages as input to the network. In doing so, the causes and effects of factors that determine the behavior of the modeled problems do not need to be fully understood.

Fig. 2 depicts a wall structure system, where the wall length is assumed to be 1.8 times the final excavation depth, *H*; *R*, an index of the observation point, is the normalized depth of the measuring point; and *W* and *D* denote the thickness of the diaphragm wall and the depth of the current excavation, respectively. Combining the aforementioned terms with monitored data, each instance consists of seven inputs and one output.

Inputs

- Wall thickness: W
- Depth of excavation surface: D
- Equivalent SPT-N value between the depth of $D-0.25H$ and $D+0.25H$: $\overline{N}$
- Index of observation point: R
- Wall deflection of observation point in $(i-3)$th excavation stage: Δ_R^{i-3}

J. Comput. Civ. Eng. 2002.16:59-65.

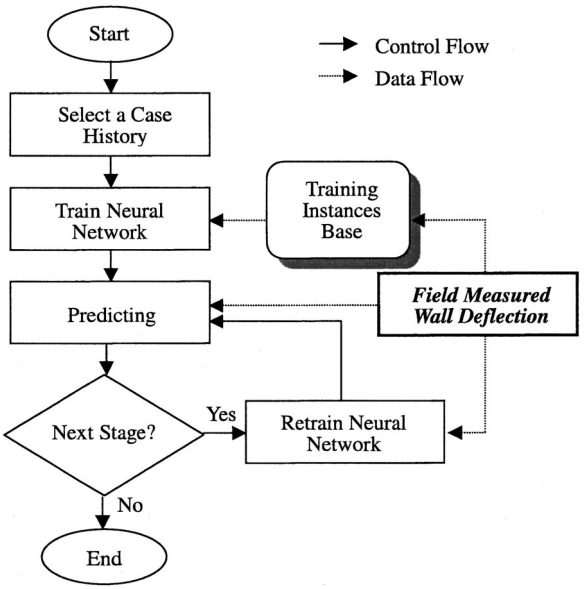

Fig. 3. Flow chart of verification

- Wall deflection of observation point in $(i-2)$th excavation stage: Δ_R^{i-2}
- Wall deflection of observation point in $(i-1)$th excavation stage: Δ_R^{i-1}

Output

Wall deflection of observation point in ith excavation stage Δ_R^i. Notably, in these instances, if i equals 1, Δ_R^{i-1}, Δ_R^{i-2} and Δ_R^{i-3} are zero; if i equals 2, Δ_R^{i-2} and Δ_R^{i-3} are zero; and if i equals 3, Δ_R^{i-3} is also zero.

In this study, 18 case histories of deep excavation with 4–7 excavation stages each, resulting in a total of 93 sets of wall deflection, are used to establish the instance base. Each diaphragm wall is discretized into 18 uniform subintervals with 19 nodal points. Therefore, a total of 1,767 (93×19) instances are generated and employed to train and verify the learning performance of the network.

Computational Results

A hidden layer feed-forward network with 7 input nodes, 15 hidden nodes, and 1 output node was used to solve the diaphragm wall deflection problem. Fig. 3 displays the flow chart of the verification process, as summarized in the following:

1. Select a case history for verification;
2. Train the neural network with other case histories;
3. Predict the wall deflection of subsequent excavation stage;
4. After the excavation stage is completed, append the measured wall deflection of the excavation stage to the training instances;
5. Retrain the neural network; and
6. Repeat Steps 3–5 until the prediction is complete.

The 18 case histories of deep excavation in the Taipei Basin are sequentially used to verify the prediction accuracy of a neural network. Herein, only the third to seventh excavation stages (57 sets of wall deflection) are of concern, because engineering failures seldom occur in the first and second excavation stages. Hence, a total number of 1,083 (57×19) instances are adopted for verification. Fig. 4 presents the forecasted maximum deflections of the 57 sets of wall deflection. The correlation coefficient between predicted and measured maximum wall deflections is 0.9081. The number of cases with relative error of predicted maximum wall deflection in the range of [0,10%], [10,20%], and greater than 20% are 28, 16, and 13, respectively. If we define that the prediction is failed as the relative percentage error exceeds 20, then 77% of cases are acceptable predictions.

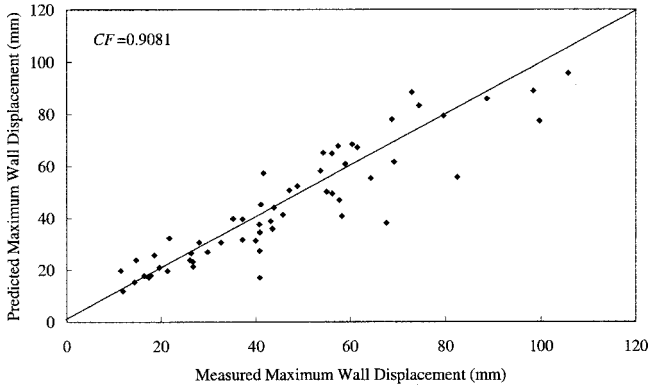

Fig. 4. Computational results of predicted maximum wall deflection

Furthermore, the average error of wall deflection for all verification instances is 5.3 mm. The average error of the seven largest deflection points of each excavation stage of case histories is 6.4 mm. The average error of wall deflection of embedded points is 4.8 mm. As a result, the largest prediction error occurs near the maximum wall deflection nodal point. Additionally, for maximum wall deflection, 38.6% of computed case histories occur at the same nodal point, as compared with measured case histories; 49.1% of maximum wall deflection of computed case histories are only inaccurate by one nodal point. These results imply that more than 87% of the predicted maximum deflection point is less than one nodal point from the observed location. Furthermore, the wall shape of predicted wall deflection resembles that of the measured wall deflection.

For comparison, the finite element analysis involving optimization back analysis calculating soil parameters, proposed by Chi et al. (1999), is employed as a reference. The hyperbolic stress-strain relationship and Mohr-Coulomb plasticity are adopted for the soil model. Four most important soil parameters are obtained by optimization back analysis calculation. The four soil parameters are the ratio of Young's modulus of clay over undrained shear strength, the ratio of Young's modulus of sand over SPT-N, the variation of undrained shear strength, and the undrained shear strength at ground level.

Three excavation projects are adopted to compare the computational performances of ANN prediction model and the reference approach. Both Cases A and B applied a 0.6-m thick diaphragm wall, however, the depths of the wall were 23 and 21.5 m, respectively. Furthermore, the walls were constructed with the bottom-up construction method. The third case, Taipei National Enterprise Center (TNEC) Case, is the construction project in Taiwan. Ou et al. (1998) studied this case, and Chi et al. (1999) investigated optimization of back analysis by incorporating it into a finite element analysis. Case TNEC used a 0.9-m thick and 35-m depth diaphragm wall as the earth-retaining structure. Differing from the other cases, however, a top-down construction method was applied.

J. Comput. Civ. Eng. 2002.16:59-65.

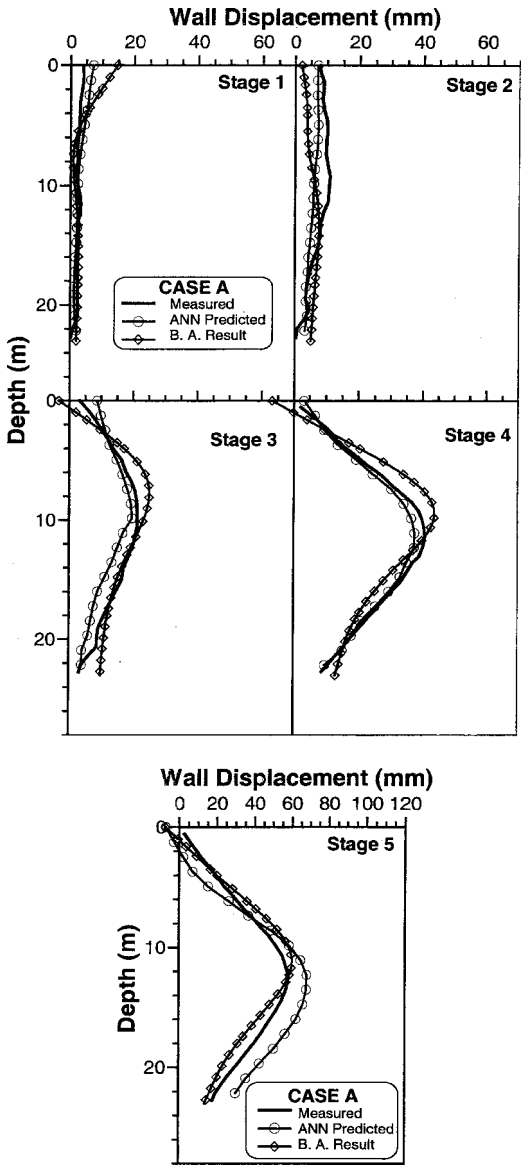

Fig. 5. (a) Back analysis results and artificial neural network (ANN) prediction of case A (excavation stages 1–4) and (b) Back analysis results and ANN prediction of case A (excavation stage 5)

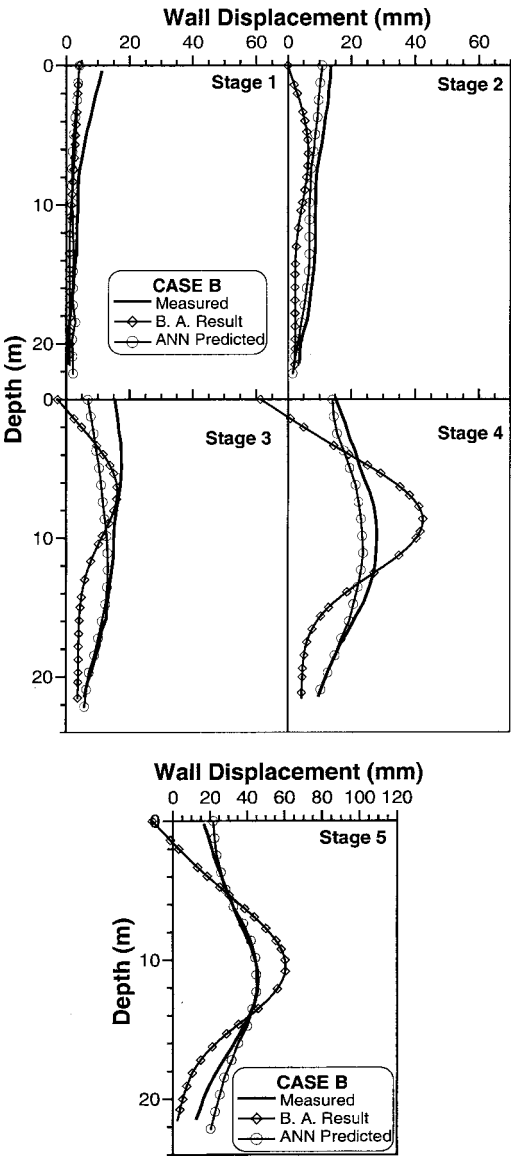

Fig. 6. (a) Back analysis results and artificial neural network (ANN) prediction of case B (excavation stages 1–4) and (b) Back analysis results and ANN prediction of case B (excavation stage 5)

Figs. 5(a and b) depict the back analysis results and ANN prediction model of Case A, where the back analysis results are curve fittings rather than predictions of measured wall deflections. According to these figures, the optimization back analysis has acceptable convergence in fitting measured wall deflections. Also, the ANN model performs acceptable predictions. Similarly, Figs. 6(a and b) illustrate the back analysis results and ANN prediction model of Case B. The optimization back analysis in this case has a poor convergence in fitting measured wall deflections. For example, the magnitudes of maximum wall deflections of excavation stages 4 and 5 are visibly larger than the measured results, and the deflection shapes of excavation stages 3, 4, and 5 are markedly dissimilar to the measured wall deflections. However, the predicted wall deflections from the ANN model are acceptable.

In Case TNEC, the back analyzed soil parameters from previous excavation stages are applied for finite-element analysis to predict wall deflections of the subsequent excavation stage. In doing so, the prediction begins at the second excavation stage. For comparison, soil parameters based on local empirical formulas are also employed herein. Figs. 7(a and b) display the computed wall deflections from ANN prediction model and the finite-element analysis. The ANN model gives acceptable prediction of wall deflections, except for excavation stages 2 and 3. Analysis results obtained from finite-element analysis, based on local empirical formulas, markedly differ from those of measured wall deflection in all excavation stages. In addition, analysis results obtained from the finite-element analysis, using back analyzed soil parameters, satisfactorily predict the maximum wall deflection except for excavation stages 2 and 3. However, the predicted wall deflections of the embedded points are inaccurate except for excavation stages 2 and 7.

Sensitivity Analysis

After an ANN model is successfully trained, the relative strength of effect for input element on output data can be derived based on

J. Comput. Civ. Eng. 2002.16:59-65.

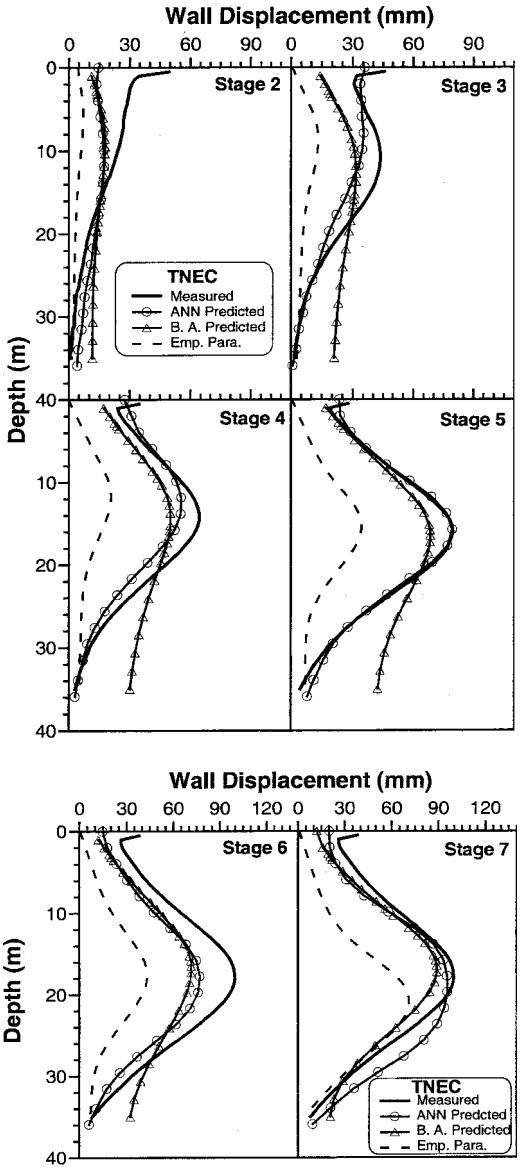

Fig. 7. (a) Comparison of back-analysis and artificial neural network (ANN) predictions of the case of the Taipei National Enterprise Center (TNEC) (excavation stages 2–5) and (b) Comparison of back analysis and ANN predictions of case TNEC (excavation stages 6 and 7)

the weights stored in the network (Yang and Zhang 1997). This work adopted an importance index, I_{ik}, to express the degree of sensitivity for each input parameter x_i on one of data in output o_k. Hereinafter, the process of sensitivity analysis is summarized as follows:

1. ABP-based ANN model has been successfully trained. The computed output for any node then can be yielded through the network and expressed as

$$o_k = f(net_k), \quad net_k = \sum_{jn} o_{jn} W_{jn,k} + \theta_k \tag{13}$$

$$o_{jn} = f(net_{jn}), \quad net_{jn} = \sum_{jn-1} o_{jn-1} W_{jn-1,jn} + \theta_{jn} \tag{14}$$

$$\vdots \tag{15}$$

$$o_{j1} = f(net_{j1}), \quad net_{j1} = \sum_{i} x_i W_{i,j1} + \theta_{j1}$$

where x_i is ith input parameter; and o_k, o_{jn}, and o_{j1} denote the computed output for output node j, hidden node j_n, and input node j_1, respectively.

2. The variance of output with the change of each input parameter can be derived. The variance is represented by the following differential equation:

$$\frac{\partial o_k}{\partial x_i} = \frac{\partial o_k}{\partial net_k} \frac{\partial net_k}{\partial o_j} \cdots \frac{\partial net_i}{x_i}$$

$$= \sum_{j_n} \sum_{j_{n-1}} \cdots \sum_{j_1} [W_{j_n k} f'(net_k) W_{j_{n-1} j_n}$$

$$\times f'(net_{j_n}) \ldots W_{i j_1} f'(net_{j_1})] \tag{16}$$

where j_n, j_{n-1}, . . . , and j_1 denote hidden nodes in the nth, $(n-1)$th, . . . , and first hidden layer, respectively; $W_{j_n k}$ denotes weight between the kth output node and hidden node j_n; $W_{j_{n-1} j_n}$ denotes weight between the hidden nodes j_{n-1} and j_n; $W_{i j_1}$ denotes weight between the ith input node and the hidden node j_1; net_k, net_{jn} and net_{j1} denote weighted sums of kth output node, the hidden node j_n and j_1, respectively; and f' denotes differential function of the activation function f.

3. The total variance of training instances as a temporary variable T_{ik} of x_i and o_k

$$T_{ik} = \sum_{p} \left| \left(\frac{\partial o_k}{\partial x_i} \right)_p \right| \tag{17}$$

where p denotes pth training instances.

4. The importance index, I_{ik}, of input parameter x_i to output data o_k, therefore, can be calculated as

$$I_{ik} = \frac{T_{ik}}{T_{\max}} \tag{18}$$

where $T_{\max}$ is maximum sum of variance.

If the network converges, the terms I_{ik} for each instance also converge to constants. Obviously, a large value of I_{ik} indicates more effect on the output data. A positive value of I_{ik} implies the positive relation; i.e., the change of output data is proportional to the change of input variable. On the other hand, a negative value of I_{ik} specifies negative action. The output data have no relation with the input variable when their I_{ik} equals zero. Herein, the importance indexes for each input, W, D, N, R, Δ_R^{i-3}, Δ_R^{i-2} and Δ_R^{i-1}, to the output, Δ_R^i, are derived via the aforementioned sensitivity analysis process. They are 0.3204, 0.121, 0.1313, 0.173, 0.271, 0.3759, and 1, respectively. Revealed from these results, the input parameters D, N, and R have little influence in the prediction of the diaphragm wall deflection. Alternately, the wall deflection in the last excavation stage is the most important factor for the prediction. Restated, incorporating the previous excavation stages into the knowledge representation, which reflects both the soil and construction factors for the case to be predicted, has an important effect on the ANN prediction model.

Conclusions

In this work, the adaptive L-BFGS supervised neural network was applied to predict the diaphragm wall deflection of deep excavations. Training data were collected from the construction projects in the Taipei Basin. To compare, the conventional finite element analysis involving optimization back analysis calculating soil parameters is also applied to evaluate these diaphragm wall deflections in deep excavation. The conclusions are summarized as follows:

J. Comput. Civ. Eng. 2002.16:59-65.

1. This study presents a neural network prediction model for deep excavation in geotechnical engineering to assess the safety of retaining systems during construction. Due to the historical information that is gathered during excavations, the limitation of understanding cause and effect, which in turn determines the behavior of the obstacles being modeled, is avoided. The advantage of the ANN prediction model is that it does not require a rigorous understanding of cause and effect. Moreover, the soil models are not significant to the predictions of wall deflections in deep excavations as compared with other factors.
2. In deep excavations, the artificial neural network can reasonably predict the magnitude as well as the location of maximum deflection of diaphragm wall. In addition, reliable predictions on wall deflections in the embedded position are also achieved. These results imply that the shape of the predicted wall matches that of the measured wall. As a result, neural network prediction allows engineers to estimate the wall structure system prior to the next excavation stage.
3. Input wall deflections of the three previous excavation stages provides a more accurate prediction of the ensuing excavation stage. Also, as compared with other conventional approaches, the importance of soil parameters is markedly reduced. The simulation results indicate that the monitoring data involving various information reflects both the uncertainty and the construction factors of field situations.

Acknowledgment

The writers would like to thank the Sinotech Engineering Consultants, Inc. of the Republic of China for financially supporting this research under Grant No. SEC/R-GT-99-01.

References

Adeli, H., and Hung, S. L. (1993). ''A concurrent adaptive conjugate gradient learning algorithm on MIMD shared memory machine.'' *J. Supercomp. Appl.*, 7(2), 155–166.

Adeli, H., and Hung, S. L. (1994). ''An adaptive conjugate gradient learning algorithm for effective training of multilayer neural networks.'' *Appl. Math. Comput.*, 62(1), 81–102.

Adeli, H., and Hung, S. L. (1995). *Machine learning—Neural networks, genetic algorithms, and fuzzy systems*, Wiley, New York.

Chi, S. Y., Chern, J. C., and Wang, C. C. (1999). ''Information construction approach for deep excavation.'' *Proc., Int. Symposium Geotechnical Aspects of Underground Construction in Soft Ground*, Balkema, Rotterdam, The Netherlands, 471–476.

Clough, G. W., and Hansen, L. A. (1981). ''Clay anisotropy and braced wall behavior.'' *J. Geotech. Eng.*, 107(7), 893–913.

Elkordy, M. F., Cheng, K. C., and Lee, G. C. (1994). ''A structural damage neural network monitoring system.'' *Microcomput. Civ. Eng.*, 9(2), 83–96.

Flood, I., and Kartam, N. (1994a). ''Neural networks in civil engineering. I: Principles and understanding.'' *J. Comput. Civ. Eng.*, 8(2), 131–148.

Flood, I., and Kartam, N. (1994b). ''Neural networks in civil engineering II: principles and understanding.'' *J. Comput. Civ. Eng.* 8(2), 149–162.

Ghaboussi, J., Garrett, J. H., and Wu, X. (1991). ''Knowledge-based modeling of material behavior and neural networks.'' *J. Eng. Mech.* 117(1), 132–153.

Gioda, G., and Sakurai, S. (1987). ''Back-analysis procedures for the interpretation of field measurements in geomechanics.'' *Int. J. Numer. Analyt. Meth. Geomech.*, 11, 555–583.

Goh, A. T. C. (1994). ''Seismic liquefaction potential assessed by neural networks.'' *J. Geotech. Eng.*, 120(9), 1467–1480.

Gunaratnam, D. J., and Gero, J. S. (1994). ''Effect of representation on the performance of neural networks in structural engineering applications.'' *Microcomput. Civ. Eng.*, 9(2), 97–108.

Hajela, P., and Berke, L. (1991). ''Neurobiological computational modes in structural analysis and design.'' *Computers and Struct.*, 41(4), 657–667.

Hashash, Y. M. A., and Whittle, A. J. (1996). ''Ground movement prediction for deep excavations in soft clay.'' *J. Geotech. Eng.*, 122(6), 474–486.

Haykin, S. (1994). *Neural networks—A comprehensive foundation*, Macmillan, New York.

Hung, S. L., and Adeli, H. (1991). ''A hybrid learning algorithm for distributed memory multicomputers.'' *Heuristic-The J. Knowledge Eng.*, 4(4), 56–68.

Hung, S. L., and Adeli, H. (1994). ''A parallel genetic/neural network learning algorithm for MIMD shared memory machines.'' *IEEE Trans. Neural Netw.*, 5(6), 900–909.

Hung, S. L., and Lin, Y. L. (1994). ''Application of an L-BFGS neural network learning algorithm in engineering analysis and design.'' *Proc., 2nd National Conf. on Structural Engineering, Chinese Society of Structural Engineering*, Taiwan, R.O.C. (in Chinese).

Juang, C. H., Ni, S. H., and Lu, P. C. (1999). ''Training artificial neural networks with the aid of fuzzy sets.'' *Microcomput. Civ. Eng.*, 14(6), 407–415.

Kang, H.-T., and Yoon, C. J. (1994). ''Neural network approaches to aid simple truss design problems.'' *Microcomput. Civ. Eng.*, 9(3), 211–218.

Moller, M. F. (1993). ''A scaled conjugate gradient algorithm for fast supervised learning.'' *Neural Networks*, 69, 525–533.

Ni, S. H., Lu, P. C., and Juang, C. H. (1996). ''A fuzzy neural network approach to evaluation of slope failure potential.'' *Microcomput. Civ. Eng.*, 11(2), 59–66.

Nocedal, J. (1980). ''Updating quasi-Newton matrix with limited storage.'' *Math. Comput.*, 35, 20–33.

Ou, C. Y., Liao, J. T., and Lin, H. D. (1998). ''Performance of diaphragm wall constructed using top-down method.'' *J. Geotechnical and Geoenvironmental Eng.*, 124, 798–808.

Ou, C. Y., and Tang, Y. G., (1994). ''Soil parameter determination for deep excavation analysis by optimization.'' *J. Chin. Inst. Eng.*, 17(5), 671–688.

Peck, R. B. (1969a). ''Advantages and limitations of the observational method in applied soil mechanics.'' *Geotechnique*, 19(2), 171–187.

Peck, R. B. (1969b). ''Deep excavations and tunneling in soft ground.'' *Proc., 7th Int. Conf. Soil Mechanics and Foundation Engineering*, Univ. Nacional Autonoma de Mexico Instituto de Ingenira, Mexico City, 225–290.

Powrie, W., and Li, E. S. F. (1991). ''Finite element analysis of an in situ wall porpped at formation level.'' *Geotechnique*, 41(4), 499–514.

Rumelhart, D., Hinton, G., and Williams, R. (1986). ''Learning representations by back-propagation errors.'' *Parallel distributed processing*, Vol. 1, D. Rumelhart et al., eds., MIT Press, Cambridge, Mass., 318–362.

Sanossian, H. Y., and Evans, D. J. (1995). ''Gradient range-based heuristic (GRBH) method for accelerating neural network convergence.'' *Integrated Comp.-Aided Eng.*, 2(2), 147–152.

Stephen, J. E., and Vanluchene, R. D. (1994). ''Integrated assessment of seismic damage in structures.'' *Microcomput. Civ. Eng.*, 9(2), 119–128.

Vanluchene, R. D., and Sun, R. (1990). ''Neural networks in structural engineering.'' *Microcomput. Civ. Eng.*, 5(3), 207-215.

Whittle, A. J., Hashash, Y. M. A., and Whitman, R. V. (1993). ''Analysis of deep excavation in Boston.'' *J. Geotech. Eng.*, 119(1), 69-90.

Yang, Y., and Zhang, Q. (1997). ''A hierarchical analysis for rockengineering using artificial neural networks.'' *Rock Mech. and Rock Eng.*, 30(4), 207–222.

J. Comput. Civ. Eng. 2002.16:59-65.