Estimating Model Parameters of Conditioned Soils by using artificial network

Zichang Shangguan [a,b]
([a] School of Civil and Hydraulic Engineering, Dalian University of Technology, Dalian 16023,China)
([b] Institute of Civil Engineering, Dalian Fishery University, Dalian 116023, P R China)
shangguan@dlfu.edu.cn

Shouju Li [c,1], Wei Sun[d] and Maotian Luan[a]
([c] State Key Lab. of Struct. Anal. for Ind. Equip., Dalian University of Technology, Dalian, 116024, China)
([d] School of Mechanical Engineering, Dalian University of Technology, Dalian 116024, China)
lishouju@dlut.edu.cn; sunei@dlut.edu.cn; mtluan@dlut.edu.cn

***Abstract*—The parameter identification of nonlinear constitutive model of soil mass is based on an inverse analysis procedure, which consists of minimizing the objective function representing the difference between the experimental data and the calculated data of the mechanical model. The ill-poseness of inverse problem is discussed. The classical gradient-based optimization algorithm for parameter identification is also investigated. Neural network models are developed for estimating model parameters of conditioned soils in EBP shield. The weights of neural network are trained by using the Levenberg-Marquardt approximation which has a fast convergent ability. The parameter identification results illustrate that the proposed neural network has not only higher computing efficiency but also better identification accuracy. The results from the model are compared with simulated observations. The models are found to have good predictive ability and are expected to be very useful for estimating model parameters of conditioned soils in EBP shield.**

***Index Terms*—parameter estimation, neural network, inverse problem, shield machine**

I. Introduction

The problem of parameter identification in distributed parameter system has been studied extensively during the last three decades. The term “distributed parameter system” implies that the response of the system is governed by a partial differential equation and parameters imbedded in the equation are spatially dependent. The inverse problem of parameter identification concerns the optimal determination of the parameters by observing the dependent variable collected in the spatial and time domains [1]. With inverse models used to determine parameter values that optimize the fit of the model results to the field observations for a given model configuration, parameter values and other aspects of the model are adjusted until the dependent variables match field observations [2]. In order to estimate petrophysical parameters from Resistivity Imaging data, the theory of the forward and inverse problem that relates the electrical resistivity with PP was developed by Shevnin . Each field survey should include a VES survey, groundwater resistivity measurements in order to determine the groundwater salinity, and collecting some representative soil samples in the study site for resistivity measurements as function of pore water salinity in laboratory, creating a soil petrophysical model of the site. This technology can be used for the characterization of uncontaminated and oil contaminated sites [3]. Tan presented a fast and robust technique for the experimental identification of soil. This technique, based on the Newton–Raphson method, estimates the unknown parameters of the soil mechanics equations by minimizing the error between measured failure forces and estimated failure forces. The measured failure forces can be obtained by measuring the forces acting on the bucket during the excavation operation, while the estimated failure forces were obtained by an analytical soil model [4]. Mertens developed a coupled simulation global optimization technology and evaluated its suitability for multi-objective inverse modeling. In particular, the trade-off between goodness of fit against leachate volume and soil moisture content in unirrigated and irrigated lysimeters was evaluated [5].

Shield tunneling has become a well-established tunnel construction method in various ground conditions. It is characterized by relatively complex interactions between the soil, the tunnel boring machine (TBM), the hydraulic jacks, the tunnel lining and the tail void grout. It is sometimes difficult to clearly identify correlations between key parameters from measurement data due to the varying conditions of the measurements and the resulting large scatter. If realistic numerical models are used, simulations represent a useful tool to identify and quantify such correlations. To choose a constitutive

Manuscript received January 1, 2009; revised June 1, 2009; accepted July 1, 2009.
Corresponding Author: Shouju li

doi:10.4304/jsw.5.3.296-303

model, one must verify its ability to represent the real behaviour of the soil under the considered loading. Excavating a tunnel leads to a complex loading which will generate in particular a rotation of principal stresses. Two soil constitutive models are used in this study: the elastic-perfectly-plastic Mohr-Coulomb model and an elastic-plastic model developed in the Ecole Centrale de Lyon. The Duncan model for describing deformation behaviors of soils uses isotropic hyperbolic stress-strain relationship. The elastic parameters from this relationship vary according to the stress state. The Young's modulus increases with confining pressure and decreases with increasing shear stress. In a finite element simulation of a geotechnical problem, calibrations of the models used to reproduce soil behavior often pose significant challenges. Real soil is a highly nonlinear material, with both strength and stiffness depending on stress and strain levels. Numerous constitutive models have been developed that can capture many of the important features of soil behavior. An effective and more objective way to calibrate a soil model employs inverse analysis techniques to minimize the difference between experimental data (laboratory or field tests) and numerically computed results. Kasper (2006) modeled the advancement of the step-by-step tunnel construction process by using a three-dimensional finite element model, which takes into account all relevant components of shield tunnelling. The material behaviour of the saturated soil and the tail void grout is modeled by a two-field finite element formulation in conjunction with an elasto-plastic Cam-Clay model for the soil and a hydration-dependent constitutive model for the grout. The analyses provide valuable information with regard to the significance of the investigated parameters and demonstrate the complexity of the various interactions in shield tunneling [6]. Bernat (1998) implemented a numerical simulation using a finite element method in the aim of developing a procedure to predict the movement induced by shield tunnelling in soft soil [7]. Plasticity models for frictional materials such as concrete, soil and rock are most conveniently represented in principal stress space. The general characteristics of the yield surface are described by its crosssectional shape on the deviatoric plane and its trace on meridian planes. In frictional materials the form of the failure envelope is significantly affected by the value of the confining pressure. On a meridian plane the shape of the failure surface can curve like a parabola, while on the deviatoric plane the shape can vary from a curvilinear triangle at low confining pressures to nearly circular at high confining pressures. A classical plasticity model that takes into account the shape of the failure surface on the deviatoric plane is the Mohr–Coulomb model [8]. The Mohr-Coulomb criterion has been extensively used to model the static and dynamic behaviour of retaining walls in both analytical and numerical studies, the main advantage being its simplicity. In the field of computational geomechanics, more sophisticated soil models have been proposed with take account of the inter-mediate principal stress, giving rise to an increase in the equivalent friction angle when analyzing plane strain problems [9]. Artificial neural network is a powerful tool of information processing. Due to its strong ability of modeling linear and nonlinear relationship, it has been widely used in optimization, calibration, parameter identification, modeling and pattern recognition [12-22]. The aim of the paper is to propose a new inversion algorithm to estimate model parameters, and to demonstrate the usefulness of the neural network for parameter estimate of nonlinear constitutive model of conditioned soils.

II. Basic Mechanical Properties Of Conditioned Soils

In order to provide a simple framework encompassing the most important characteristics of conditioned soil stress-strain behaviour using material parameters available from conventional laboratory tests, Duncan and Chang [10] developed a hyperbolic model based on the hyperbolic equation proposed by Kondner [11]. Because of its simplicity and wide applicability, Duncan and Chang's model is adopted herein to define the non-linear stress-strain behaviour of the conditioned soil in chamber of shield machine. The stress-strain behavior of conditioned soil depends on a number of different factors including density, water content, structure, drainage conditions, strain conditions, duration of loading, stress history, confining pressure, and shear stress. The hyperbolic equation proposed by Kondner is expressed as follows

$$(\sigma_1-\sigma_3)=\frac{\varepsilon}{a+b\varepsilon}. \tag{1}$$

Where σ_1 and σ_3 are the major and minor principal stresses; ε is the axial strain; a and b are constants whose values may be determined experimentally. By expressing the parameters a and b in terms of the initial tangent modulus value and the compressive strength, (1) can be rewritten as follows

$$(\sigma_1-\sigma_3)=\frac{\varepsilon}{[\frac{1}{E_i}+\frac{\varepsilon R_f}{(\sigma_1-\sigma_3)_f}]}. \tag{2}$$

Where R_f is the failure ratio, which always has a value less than unity, the value of R_f has been found to be between 0.75 and 1.00. $(\sigma_1-\sigma_3)_f$ is the compressive strength. Experimental studies by Janbu have shown that the relationship between initial tangent modulus and confining pressure may be expressed as

$$E_i=Kp_a(\frac{\sigma_3}{p_a})^n. \tag{3}$$

Where E_i is the initial tangent modulus; p_a is the atmospheric pressure expressed in the same pressure units as E_i; K is a modulus number; n is the exponent determining the rate of variation of E_i with σ_3. If it is assumed that failure will occur with no change in the value of σ_3, the relationship between compressive strength and confining pressure may be expressed

conveniently in terms of the Mohr-Coulomb failure criterion as

$$(\sigma_1-\sigma_3)_f=\frac{2c\cos\varphi+2\sigma_3\sin\varphi}{1-\sin\varphi}. \quad (4)$$

Where c and φ are the Mohr-Coulomb strength parameters. (3), (4) combination with (2), provide a means of relating stress to strain and confining pressure by means of the five parameters K, n, c , φ and R_f. The stress-strain relationship may be employed very conveniently in incremental stress analyses because it is possible to determine the value of the tangent modulus corresponding to any point on the stress-strain curve. If the value of the minor principal stress is constant, the tangent modulus may be expressed as

$$E_t=\frac{\partial(\sigma_1-\sigma_3)}{\partial\varepsilon}. \quad (5)$$

Where E_t is the tangent modulus of soil. The values of tangent modulus and Poisson ratio for any stress condition may be expressed as

$$E_t=Kp_a(\frac{\sigma_3}{p_a})^n[1-\frac{R_f(1-\sin\varphi)(\sigma_1-\sigma_3)}{2C\cos\varphi+2\sigma_3\sin\varphi}]^2. \quad (6)$$

$$\nu_i=\frac{G-F\lg(\frac{\sigma_3}{p_a})}{1-A^2}. \quad (7)$$

$$A=\frac{D(\sigma_1-\sigma_3)}{Kp_a[\frac{\sigma_3}{p_a}][1-\frac{R_f(\sigma_1-\sigma_3)(1-\sin\varphi)}{2c\cos\varphi+2c\sigma_3\sin\varphi}]}. \quad (8)$$

This expression for tangent modulus may be employed very conveniently in incremental stress analyses, and constitutes the essential portion of the stress-strain relationship.

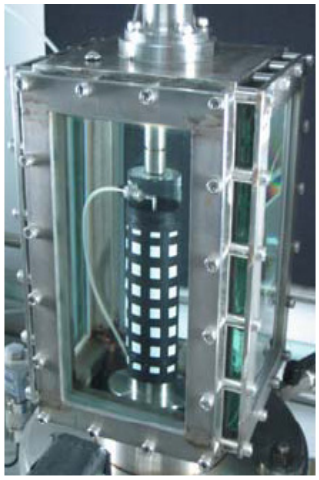

Figure 1. Sketch map of laboratory test of conditioned soil for tree-dimensional compression.

Laboratory tests were conducted in a tree-dimensional compression machine to study the characteristics of the conditioned soil. Fig. 1 shows the sketch map of laboratory test of conditioned soil for tree-dimensional compression. Fig. 2 demonstrates relationship between stress and axial strain. Fig. 3 depicts the variation of radial strain versus axial strain.

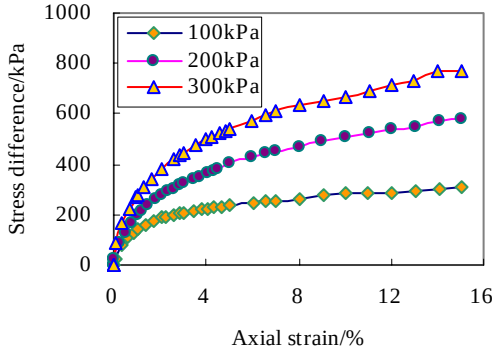

Figure 2. Relationship between stress and axial strain.

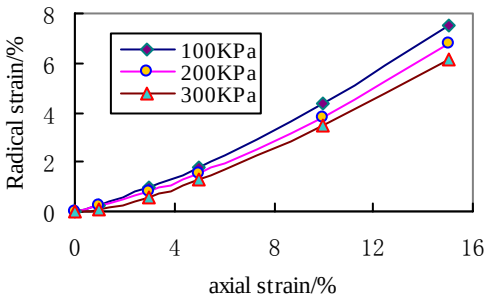

Figure 3. Variation of radial strain versus axial strain.

III. Basic Characteristics Of Inverse Problem Of Parameter Estimation

Parameterization of the system denotes the discovery of a minimal set of model parameters whose values completely characterize the system. Forward modeling denotes the discovery of the physical laws allowing us, for given values of the model parameters, to make predictions on the results of measurements on some observable parameters. Inverse modeling denotes use of the actual results of some measurements of the observable parameters to infer the actual values of the model parameters [18]. A particular choice of model parameters is a parameterization of the system. Independently of any particular parameterization, it is possible to introduce an abstract space of points, a manifold, each point of which represents a conceivable model of the system. This manifold is named the model space. To obtain information on model parameters, we have to perform some observations during a physical experiment, i.e., we have to perform a measurement of

some observable parameters. We can thus arrive at the abstract idea of a data space, which can be defined as the space of all conceivable instrumental responses.

The inverse problem is often ill-posed [1]. The ill-posedness is generally characterized by the nonuniqueness and instability of the identified parameters. The instability of the inverse solution stems from the fact that small errors in measurement dada will cause serious errors in the identified parameters. The uniqueness problem has a great practical importance, because in the case of nonuniqueness, the identified parameters will differ according to the initial estimate of the parameters, and there will be no reason for the estimated parameters to be close to the true parameters. As a consequence, the responses of the model and system may differ for inputs different from those that have been used for identification. The uniqueness problem in parameter identification is intimately related to identification. The notation of identifiability addresses the question of whether it is all possible to obtain unique solution of the inverse problem for unknown parameters of interest in a mathematical model, from data collected in the spatial and time domains. Identifiability is usually not achievable in the case of point measurements where data is only available at a limited number of locations in the spatial domain.

Inverse problems correspond to the particular case where the data space D and model space M have fundamentally different physical meaning and where we are interested in translating information from the data space into the model space. It is assumed that a reasonable approximation for representing the physical theory relating the model parameters m to the observable parameters d can be written under the form of a probability density for d given and possible m

$$\Theta(d,m)=\theta(d\,|m)\mu_M(m). \tag{9}$$

Where $\Theta(d, m)$ denotes probability density, $\mu_M(m)$ is marginal probability density, which denotes the homogeneous probability density over the model space manifold M; $\theta(d|m)$ is conditional probability density. The prior information in the joint manifold $D\times M$ takes the special form

$$\rho(d,m)=\rho_D(d)\rho_M(m). \tag{10}$$

Where $\rho_D(d)$ is the probability density in the data space that represents the information found by the measurement. Above equation means that information in the space of observable parameters (data) has been obtained (from measurements) independently of the prior information in the model space. In particular, the homogenous limit of this last equation is

$$\mu(d,m)=\mu_D(d)\mu_M(m). \tag{11}$$

Where $\mu(d, m)$ represents the homogenous state of information; $\mu_M(m)$; $\rho_M(m)$ denotes the probability density representing prior information. For the posterior information in the model space, this gives

$$\sigma_M(m)=k\rho_M(m)\int_D d\,\frac{d\rho_D(d)\theta(d\,|m)}{\mu_D(d)}. \tag{12}$$

Where $\rho_M(m)$ is the prior probability density, which represents both information obtained on the observable parameters d and a prior information on the model parameters m. sometimes, we shall write above equation as

$$\sigma_M(m)=k\rho_M(m)L(m). \tag{13}$$

Where k is a constant and $L(m)$ is the likelihood function

$$L(m)=\int_D d\,\frac{d\rho_D(d)\theta(d\,|m)}{\mu_D(d)}. \tag{14}$$

The existence of the solution simply means that $\sigma_M(m)$ is not identically null. If this were case, it would indicate the incompatibility of the experimental results, the prior hypothesis on model parameters, and the theoretical information, thus showing that some uncertainty bars have been underestimated.

The uniqueness of the solution is evident when by solution we mean the probability density $\sigma_M(m)$ itself and is simply a consequence of the uniqueness of the conjunction of sates of information. Of curse, $\sigma_M(m)$ may be very pathological, but that would simply mean that such is the information we possess on the model parameters. The information itself is uniquely defined. The a prior information that the (unknown) model m is a sample of a known Gaussian probability density whose mean is m_{prior} and whose covariance matrix is C_M. The a prior probability density over the model space M is

$$\rho_M(m)=C\exp[-\frac{1}{2}\frac{(m-m_{\text{Prior}})^T(m-m_{\text{Prior}})}{C_M}]. \tag{15}$$

This probability density is assumed to be a prior in the sense that it is independent of the result of the measurements on the observable parameters d. To solve forward problem means to predict the error-free values of the observable parameters d that would correspond to a given model m.

$$d=g(m). \tag{16}$$

Where g is called the forward operator. It expresses mathematical model of the physical system under study. Some measurements on the observable parameters d whose results can be represented by a Gaussian probability density centered at d_{obs} and with covariance C_D

$$\rho_D(d)=C\exp[-\frac{1}{2}\frac{(d-d_{obs})^T(d-d_{obs})}{C_D}]. \tag{17}$$

The a posterior probability density in the model space is expressed as follows

$$\sigma_M(m)=C\exp[-S(m)]. \tag{18}$$

The misfit function $S(m)$ is defined as

$$S(m)=\frac{1}{2}[\|g(m)-d_{obs}\|_D^2+\|m-m_{prior}\|_M^2]. \tag{19}$$

If the equation $d=g(m)$ solving the forward problem is linear, this gives

$$d=Gm. \tag{20}$$

The misfit function is expressed as follows

$$S(m)=\frac{1}{2}[\|Gm-d_{obs}\|_D^2+\|m-m_{prior}\|_M^2]. \quad (21)$$

As the misfit function is quadratic in *m*, the posterior probability density $\sigma_M(m)$ is a Gaussian probability density, so there must be a point $\tilde{m}$ and a covariance matrix $\tilde{C}_M$ such that the posterior probability density can be written

$$\sigma_M(m)=C\exp[-\frac{1}{2}\frac{(m-\tilde{m})^T(m-\tilde{m})}{\tilde{C}_M}]. \quad (22)$$

The basic problem, in this linear case, is then the evaluation of the center *m* and the covariance $\tilde{C}_M$ of the posterior covariance probability density $\sigma_M(m)$. The center of poster Gaussian is given by

$$\tilde{m}=m_{prior}+\frac{C_M G^T}{(GC_M G^T+C_D)}(d_{obs}-Gm_{prior}). \quad (23)$$

Point m has been defined as the center of the poster Gaussian. It could have been defined as the point realizing the minimum of the least-squares misfit function. From this perspective, m is the best point in the sense that it is close to the prior point m_{prior}, and the predicted data Gm are close to the observed data d_{obs}. If the equation $d=g(m)$ solving the forward problem is actually nonlinear, there is no simplification in equation giving the posterior probability density in the model space

$$\sigma_M(m)=C\exp[-S(m)]. \quad (24)$$

If the results of the observations can be described by a Gaussian probability with mean vector, *d*, and covariance matrix C_D, the a priori information can be described by a Gaussian probability with mean vector, mp, and covariance Matrix, C_M, and if the equation solving the forward problem is quasi linear in the region of the model space with significant a posteriori probability, then the a posteriori probability in the model space is approximately Gaussian. The mean *m* of the approximated Gaussian probability minimizes the misfit function

$$S(m)=\frac{1}{2}[\|g(m)-d_{obs}\|_D^2+\|m-m_{prior}\|_M^2]. \quad (25)$$

And can be obtained using an iterative process.

IV. Classical Gradient-based Algorithms For Parameter Identification And Its Development

For modeling purposes, the objective is to determine model parameters, *m*, from a limited number of observations of strain, ε, scattered the field so that a certain criterion is optimized. If the classic least square error is used to represent the output error, the objective function to be minimized is expressed as follows

$$\min J(m)=[\varepsilon^c(m)-\varepsilon^o]^T[\varepsilon^c(m)-\varepsilon^o]. \quad (26)$$

Where *J* is the objective function of parameter estimation; ε^c is the vector of calculated strains at observation points, based upon some estimated values of parameters, m, and ε^o is the vector of observed strains. The Gauss-Newton algorithm has proven to be an effective algorithm to perform minimization. The popularity of the algorithm stems from the fact that it does not require the calculation of the Hessian matrix as is required by the Newton method and the rate of convergence is superior when compared to the classical gradient searching procedures. The algorithm is basically developed for unconstrained minimization. However, constrains such as upper and lower bounds are easily incorporated in the algorithm with minor modifications. The algorithm starts with a set of initial estimates of parameters and converges to a local optimum. If the objective function is convex, the local optimum would be the global optimum. Due to the presence of noise in the observations, the inverse problem is usually nonconvex, and hence only a local optimum can be assured in the minimization [1]. The algorithm generates the following parameter sequence for an unconstrained minimization problem

$$m^{k+1}=m^k-\rho^k d^k. \quad (27)$$

with

$$A^k d^k=g^k. \quad (28)$$

$$A^k=[J_J(m^k)]^T[J_J(m^k)]. \quad (29)$$

$$g^k=[J_J(m^k)]^T e^k. \quad (30)$$

$$e^k=[\varepsilon^c(m^k)-\varepsilon^o]. \quad (31)$$

Where *e* is error vector; J_J is Jacobian matrix of strain with respect to model parameters *m*; ρ^k is step size; d^k is Gauss-Newton direction vector, *k* denotes number of iteration. The step size ρ^k can be determined by a quadratic interpolation scheme such that $J(m^{k+1})<J(m^k)$, or simply by a trial-and error procedure. The covariance matrix of the estimated parameters is defined by [22]

$$\text{Cov}(\text{m}^e)=E[(m^r-m^e)(m^r-m^e)^T]. \quad (32)$$

Where m^e is estimated parameters; m^r is true parameters; *E* is mathematical expectation. An approximation of the covariance matrix of the estimated parameters in nonlinear regression can be represented by the following form

$$\text{Cov}(\text{m}^e)=\frac{1}{M-L}\frac{J(m^e)}{A}. \quad (33)$$

Occasionally, the direction matrix, *A*, may become ill-conditioned. Some modification method are made in order for the algorithm to continue, the typical examples are to apply damping least squares method developed by Levenberg and Marquardt, or regularizing least squares method proposed by Tibonov. The Levenberg-Marquardt algorithm is an iterative technique that locates the minimum of a multivariate function that is expressed as the sum of squares of non-linear real-valued functions. The strategy of altering the diagonal elements of *A* is called damping and μI is referred to as the damping term.

$$[A^k+\mu^k I]d^k=g^k. \quad (34)$$

Where *I* is a unity matrix. If the updated parameter vector $m+\Delta m$ with Δm computed from (34) leads to a reduction in the error *e*, the update is accepted and the

process repeats with a decreased damping term. Otherwise, the damping term is increased, the augmented normal equations are solved again and the process iterates until a value of Δm that decreases error is found. The process of repeatedly solving (34) for different values of the damping term until an acceptable update to the parameter vector is found corresponds to one iteration of the Levenberg-Marquardt algorithm. The Levenberg-Marquardt algorithm approximates the normal gradient descent method, while if it is small, the expression transforms into the Gauss-Newton method. After each successful step the constant μ is decreased, forcing the adjusted weight matrix to transform as quickly as possible to the Gauss-Newton solution. When after a step the errors increase the constant μ is increased subsequently.

In Levenberg-Marquardt algorithm, the damping term is adjusted at each iteration to assure a reduction in the error e. If the damping is set to a large value, matrix F in (34) is nearly diagonal and the Levenberg-Marquardt algorithm update step Δm is near the steepest descent direction. Moreover, the magnitude of Δm is reduced in this case. Damping also handles situations where the Jacobian is rank deficient and A is therefore singular. In this way, LM can defensively navigate a region of the parameter space in which the model is highly nonlinear. If the damping is small, the LM step approximates the exact quadratic step appropriate for a fully linear problem. LM is adaptive because it controls its own damping: it raises the damping if a step fails to reduce e, otherwise it reduces the damping. In this way LM is capable to alternate between a slow descent approach when being far from the minimum and a fast convergence when being at the minimum's neighborhood. The LM algorithm terminates when at least one of the following conditions is met: 1) The magnitude of the gradient of $e^{\mathrm{T}}e$, i.e.g in the right hand side of (33), drops below a threshold δg. 2) The relative change in the magnitude of & Δm drops below a threshold δm. 3) The error $e^{\mathrm{T}}e$ drops below a threshold δe. 4) A maximum number of iterations $k_{\max}$ is completed. To overcome instability, Tihonov proposed regularizing the least squares method by adding a term to the least squares criterion so that it becomes

$$\min J(m) = [\varepsilon^c(m) - \varepsilon^o]^T [\varepsilon^c(m) - \varepsilon^o] + \alpha m^T m. \tag{35}$$

Where α is a nonnegative scalar. Tihonov proved that if the solution of forward problem is linear in the parameters, then the solution of inverse problem exists, is unique, and depends continuously on the observations [22].

V. Artificial Neural Network Model For Parameter Estimating In Conditioned Soil In EPB Shield

Least squares are popular for solving inverse problems because they lead to the easiest computations. Their only drawback is the lack of robustness, i.e., their strong sensitivity to a small number of large errors in a data set. An artificial neural network model is a system with inputs and outputs based on biological nerves. The system can be composed of many computational elements that operate in parallel and are arranged in patterns similar to biological neural nets. A neural network is typically characterized by its computational elements, its network topology and the learning algorithm used. The neural networks have such strength as processing large amounts of sensory information, collective and parallel processing capability and adaptation by learning.

The architecture of BP networks, depicted in Fig. 4, includes an input layer, one or more hidden layers, and an output layer. The nodes in each layer are connected to each node in the adjacent layer. Notably, Hecht-Nielsen proved that one hidden layer of neurons suffices to model any solution surface of practical interest.

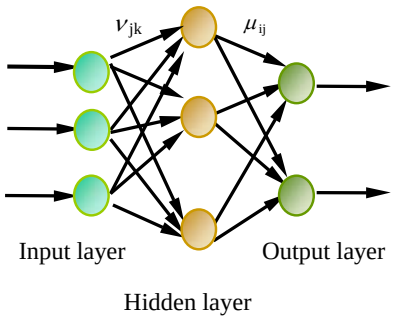

Figure 4. A schematic of neural network model.

Hence, a network with only one hidden layer is considered in this study. Before an ANN can be used, it must be trained from an existing training set of pairs of input-output elements. The training of a supervised neural network using a BP learning algorithm normally involves three stages. The first stage is the data feed forward. The computed output of the i-th node in output layer is defined as follows

$$y_i = f\left(\sum_{j=1}^{N_h} \left(\mu_{ij} f\left(\sum_{k=1}^{N_i} \nu_{jk} x_k + \theta_j\right) + \lambda_i\right)\right). \tag{36}$$

Where μ_{ij} is the connective weight between nodes in the hidden layer and those in the output layer; ν_{jk} is the connective weight between nodes in the input layer and those in the hidden layer; θ_j or λ_i is bias term that represents the threshold of the transfer function f, and x_k is the input of the kth node in the input layer. Term N_i, N_h and N_o are the number of nodes in input, hidden and output layers, respectively. The transfer function f is selected as Sigmoid function:

$$f(\cdot) = 1/[1 + \exp(-\cdot)]. \tag{37}$$

The second stage is error back-propagation through the network. During training, a system error function is used to monitor the performance of the network. This function is often defined as follows

$$E(w)=\frac{1}{2}\sum_{p=1}^{P}(\sum_{i=1}^{N_o}(y_i^p-o_i^p)^2. \quad (38)$$

Where y_i^p and o_i^p denote the practical and desired value of output node i for training pattern p, p is the number of sample. Training methods based on back-propagation offer a means of solving this nonlinear optimization problem based on adjusting the network parameters by a constant amount in the direction of steepest descent, with some variations depending on the flavor of BP being used. The optimization algorithm used to train network makes use of the Levenberg-Marquardt approximation. This algorithm is more powerful than the common used gradient descent methods, because the Levenberg-Marquardt approximation makes training more accurate and faster near minima on the error surface. The weight adjustment using Levenberg-Marquardt algorithm is expressed as follows:

$$w(k+1)=w(k)-\frac{J^T(k)e(k)}{[J^T(k)J(k)+\mu I]}. \quad (39)$$

Where $w(k)$ is the vector of network parameters(net weights and element biases) for iteration k, J is the Jacobian matrix; μ is a constant, I is a unity matrix; e is an error vector.

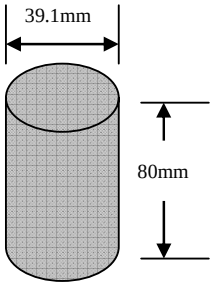

Figure 5. Dimension of test sample.

The Levenberg-Marquardt algorithm approximates the normal gradient descent method, while if it is small, the expression transforms into the Gauss-Newton method. After each successful step the constant μ is decreased, forcing the adjusted weight matrix to transform as quickly as possible to the Gauss-Newton solution. When after a step the errors increase the constant μ is increased subsequently. The training sample data are regularized before they are trained. The topology structure of neural network is composed of 31*3=93 input nodes, 100 hidden nodes and 8 output nodes. The node number of input layer is determined by number of observed deformation data. The node number of hidden layer is determined by test, which approaches to double of node number of input layer. The node number of output layer agrees with number of identified parameters. Fig. 5 shows the dimension of test sample. Fig. 6 shows objective function versus training epochs. The estimated parameters for the conditioned soil are listed in Table .

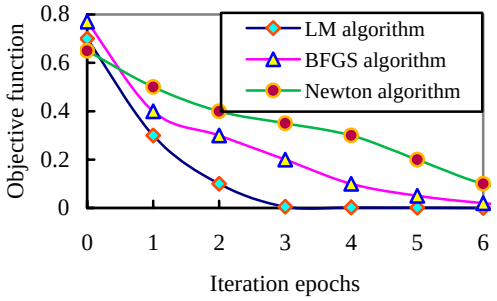

Figure 6. Objective function versus training epochs for different training algorithm.

Table I.

HYPERBOLIC MATERIAL CONSTANTS OF CONDITIONED SOIL

Parameter	C /kpa	φ	R_f	K	*n*	D	F	G
value	0.6	35	0.71	80.7	0.65	0.03	0.04	0.33

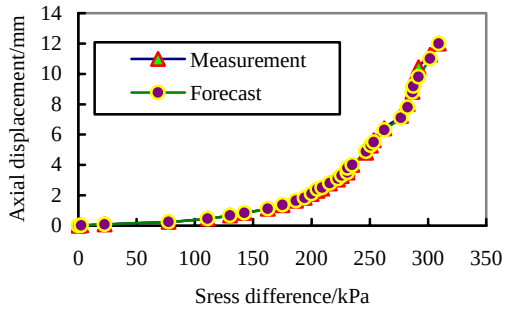

Figure 7. Comparison of observed axial displacements of test

VI. CONCLUSION

This work presents a novel procedure for estimating the model parameters of conditioned soils by using neural network approach. The proposed method of estimating the model parameters of conditioned soils has been verified by excellent agreement between the observed deformation of test sample and forecasted ones based on parameter identification.

ACKNOWLEDRMENT

The research is funded by the National Basic

Research Program (973 Program) (No. 2007CB714006) and the National Natural Science Foundation of China (No. 90815023).

REFERENCES

[1] W. W-G. YEH, "Review of parameter identification procedures in groundwater hydrology: the inverse problem," *Water Resour. Res.*, vol. 22, pp. 95–108, February 1986.

[2] E. P. Poeter, and M. C. Hill, "Inverse models: A necessary next step in ground-water modeling," *Ground Water*, vol. 35, pp. 250–260, March 1997.

[3] V. Shevnin, O. D. Rodríguez, and A. Mousatov, "Estimation of soil petrophysical parameters from resistivity data: Application to oil-contaminated site characterization," *Geofísica Internacional*, Vol. 45, pp. 179–193, 2006.

[4] C. P. Tan, Y. H. Zweiri, and K. Althoefer, "Online Soil Parameter Estimation Scheme Based on Newton–Raphson Method for Autonomous Excavation," *IEEE/ASME Transactions on Mechatronics*, vol. 10, pp. 221–230, April 2005.

[5] J. Mertens, R. Stenger, and G. F. Barkle, "Multiobjective inverse modeling for soil parameter estimation and model verification," *Vadose zone Journal*, vol. 5, pp. 917–933, June 2006.

[6] T. Kasper, and G.. Meschke, "A numerical study of the effect of soil and grout material properties and cover depth in shield tunneling," *Computers and Geotechnics*, vol. 33, pp. 234–247, June–July 2006.

[7] S. Bernat, and B. Cambou, "Soil-structure Interaction in Shield Tunnelling in Soft Soil," *Computers and Geotechnics*, vol. 22, pp. 221–242, April 1998.

[8] R. I. Borja, K. M. Sama, and P. F. Sanz, "On the numerical integration of three-invariant elastoplastic constitutive models," *Comput. Methods Appl. Mech. Engrg.*, vol. 192, pp. 1227–1258, February 2003.

[9] P. K. Woodward, "Earth pressure coefficients based on the Lade-Duncan failure," *Engineering Structures*, vol. 19, pp. 733–737, September 1997.

[10] J. M. Duncan, and C. Chang, "Nonlinear analysis of stress and strain in soils," *Journal of the Soil Mechanics and Foundations Division, Proceedings of ASCE*, vol. 96, pp. 1629–1653, September–October 1970.

[11] R. L. Kondner, "Hyperbolic stress-strain response: cohesive soils," *Journal of the Soil Mechanics and Foundations Division, Proceedings of ASCE*, vol. 89, pp. 115–143, February 1963.

[12] S. J. Li, and Y. X. Liu, "Parameter identification approaches to fractal model of mass transport for unsaturated soils," *Modern Physics Letters B*, vol. 19, pp. 172–173, December 2005.

[13] S. J. Li, and Y. X. Liu, "Inverse Determination of Model Parameters of Nonlinear Heat Conduction Problem Using Hybrid Genetic Algorithm," *Nonlinear Dynamics and Systems Theory*, vol. 7, pp. 23–34, 2007.

[14] S. J. Li, and Y. X. Liu, "An improved approach to nonlinear dynamical system identification using PID neural networks," *International Journal Of Nonlinear Sciences And Numerical Simulation*, vol. 7, pp. 177–182, 2006.

[15] G. S. Su, X. F. Zhang, and G. Q. Chen, "Identification of structure and parameters of rheological constitutive model for rocks using differential evolution algorithm," *J. Cent. South Univ. Technol.*, vol. 15, PP. 25–28, September 2008.

[16] H. B. Zhao, and S. D. Yin, "Geomechanical parameters identification by particle swarm optimization and support vector machine," *Applied Mathematical Modeling*, vol. 33, pp. 3997–4012, October 2009.

[17] J. E. Andrade, A. M. ASCE, and K. C. Ellison, "Evaluation of a Predictive Constitutive Model for Sands," *Journal of Geotechnical and Geoenvironmental Engineering*, Vol. 134, pp. 1825–1828, December 2008.

[18] A. Tarantola, *Inverse problem theory and methods for model parameter estimation*, Science Press, 2009.

[19] S. Lingireddy, "Aquifer parameter estimation using genetic algorithms and neural networks," *Civil Engineering and Environmental Systems*, vol. 15, pp. 125–144, 1998.

[20] J. Morshed, and J. J. Kaluarachchi, "Parameter Estimation Using Artificial Neural Network and Genetic Algorithm for Free-Product Migration and Recovery," *Water Resour. Res.*, vol. 34, pp. 1101–1113, May 1998.

[21] F. T. Tsai, N. Z. Sun, and W. W-G Yeh, "A combinatorial optimization scheme for parameter structure identification in ground water modeling," *Ground Water*, vol. 41, pp. 156–169, March 2003.

[22] J. Carrera, and S. Neuman, "Estimation of aquifer parameter under transient and steady state conditions: 2. Uniqueness, stability and solution algorithms," *Water Resour. Res.*, vol. 22, pp. 211–227, February 1986.

Zichang Shangguan was born in Fushun, Liaonong Province, China, on May 29, 1959. He received his B.S degree in Civil Engineering from Fuxin Mining Institute in 1982.

Currently, he is a PH.D candidate with Civil Engineering at Dalian University of Technology since 2003. His research interest is Geotechnical Engineering, Sturcture Engineering and Intelligent Computing. He is working at Dalian Fisheris University as a Professor and a Associate Dean of Civil Engineering Institute.

Shouju Li was born in Shenyang, Liaonong Province, China, on October 3, 1960. He received the Ph. D. degree in Engineering Mechanics from the Dalian University of Technology, Dalian, China, in 2004. He was Associate Professor at Department of Engineering Mechanics, Dalian University of Technology form 1994 to 2008.

Now he is a Professor of Dalian University of Technology and teaches and conducts research in the areas of neural network, intelligent optimization, parameter identification applied to soil mechanics and underground engineering fields.