INTERNATIONAL JOURNAL FOR NUMERICAL AND ANALYTICAL METHODS IN GEOMECHANICS
Int. J. Numer. Anal. Meth. Geomech. 2008; **32**:1087–1106
Published online 14 September 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/nag.660

Estimation of effective stress parameter of unsaturated soils by using artificial neural networks

C. Kayadelen*, †, ‡

Department of Civil Engineering, Nigde University, 51100 Nigde, Turkey

SUMMARY

Great efforts are required for determination of the effective stress parameter χ, applying the unsaturated testing procedure, since unsaturated soils that have the three-phase system exhibit complex mechanical behavior. Therefore, it seems more reasonable to use the empirical methods for estimation of χ. The objective of this study is to investigate the practicability of using artificial neural networks (ANNs) to model the complex relationship between basic soil parameters, matric suction and the parameter χ. Five ANN models with different input parameters were developed. Feed-forward back propagation was applied in the analyses as a learning algorithm. The data collected from the available literature were used for training and testing the ANN models. Furthermore, unsaturated triaxial tests were carried out under drained condition on compacted specimens. ANN models were validated by a part of data sets collected from the literature and data obtained from the current study, which were not included in the training phase. The analyses showed that the results obtained from ANN models are in satisfactory agreement with the experimental results and ANNs can be used as reliable tool for prediction of χ. Copyright © 2007 John Wiley & Sons, Ltd.

Received 18 April 2007; Revised 1 August 2007; Accepted 7 August 2007

KEY WORDS: effective stress; unsaturated soil; matric suction; artificial neural networks; feed-forward back propagation

INTRODUCTION

The mechanical behavior of soils such as volume change and shear strength is governed by the change in inter-particle forces defined as effective stress. Terzaghi [1] defined the effective stress concept as 'all the measurable effects of a change in stress, such as compression, distortion and a change in shearing resistance, are exclusively due to changes in the effective stress'. Experimental studies denoted that the mechanical behavior of a saturated soil can be defined as a single-valued

*Correspondence to: C. Kayadelen, Department of Civil Engineering, Nigde University Engineering and Architecture Faculty, 51100 Nigde, Turkey.
†E-mail: ckayadelen@nigde.edu.tr
‡Assistant Professor.

effective stress or one stress state variable (i.e. $(\sigma - u_w)$). Unlike the saturated soils, unsaturated soils are characterized by the presence of air phase, water phase and air–water interface in voids. Their mechanical behavior is more complex than those of saturated soils. It is therefore difficult to define an effective stress concept. It is generally desired that the concept of effective stress for unsaturated soils has a form that can be converted for saturated soil and provides the principle of continuum mechanics. In previous studies, the effective stress equations for unsaturated soils were proposed to ensure a single-valued effective stress or one stress-state variable as proposed for saturated soils.

Bishop [2] proposed well-known effective stress equations for unsaturated soils extending Terzaghi's effective stress principle for saturated soils as below:

$$\sigma' = (\sigma - u_a) + \chi(u_a - u_w) \tag{1}$$

where $(u_a - u_w)$ is the matric suction and the parameter χ, which has a magnitude between 1 and 0, represents the variation of condition from fully saturated to total dryness. The main availability of this approach is that the variation of shear strength due to the change in total stress, pore-water and pore-air pressure is governed by a single-stress variable.

Gray and Schrefler [3] reported that Bishop's equation is not thermodynamically consistent, since it is derived under assumption of incompressible grain. According to them a thermodynamic equation is formulated at the microscale and then averaged to the macroscale so that it is consistent thermodynamically. This provides consistency between the models and parameters at the two scales. Therefore, they made use of a functional dependence of the macroscale thermodynamic energy of the phase and interfaces that is consistent with the microscale situation to recover Bishop's equation. Also they emphasized that the Bishop parameter is shown to be a fraction of the solid phase surface area in contact with the wetting phase. Aitchison [4] noted that when the soils are loaded with the same stress path, the parameter χ may equal the specific value and also proposed that the effects of σ and $(u_a - u_w)$ on the mechanical behavior of unsaturated soils should be separately evaluated. Burland [5] reported that determination of the parameter χ is fairly difficult and it takes different values when it is separately valuated for shear strength or volume change. Fredlund *et al.* [6] examined the behavior of unsaturated soils by using multiphase continuum mechanics and pointed out that unsaturated soils can be characterized by using two independent stress variables, namely $(\sigma - u_a)$ and $(u_a - u_w)$. According to these researches, the effective stress equation should not depend on soil properties. However, Khalili *et al.* [7] stated that the suggestion of Fredlund *et al.* [6] may be valid for the materials with single phase and effective stress should be considered to be a function of material properties for the multiphase structures like soils.

In the literature, there exist numerous experimental evidences which support Bishop's equation [7–9]. The main difficulty in using this equation is to determine the parameter χ. Experimental measurement of χ involves a sophisticated, costly unsaturated testing procedure and especially requires long time for fine-grained soils. Hence, many researchers have developed several empirical relationship between χ and soil properties. Oberg and Sallfors [10] suggested a hypothesis for the determination of shear strength of unsaturated soils, roughly assuming S_r instead of the parameter χ in Bishop's effective stress equation. However, Coleman [11] pointed out that soil properties have considerable effects on parameter χ and a relationship between χ and S_r should not be established directly. Khalili and Khabbaz [12] demonstrated a unique relationship between parameter χ and the ratio of matric suction, $(u_a - u_w)$, over the air entry value (AEV) $(u_a - u_w)_b$, for the determination of shear strength and proposed $\chi = [(u_a - u_w)/(u_a - u_w)_b]^{-0.55}$. The other approach relating to the determination of χ was introduced by Xu [8]. The author defined parameter χ by using

DOI: 10.1002/nag

surface fractal dimension of soil pores, (D_s), as $\chi=[(u_a-u_w)/(u_a-u_w)_b]^{(3-D_s)}$. The author found reasonable agreement on comparison with the test results. In conclusion, when the previous studies are evaluated, it appears that there is a relationship between parameter χ and soil parameters. However, this relationship is fairly complex and should be described by more than one soil property rather than only a single parameter. From this point of view, the use of artificial neural network (ANN) which can allow developing complex relations appears to be reasonable for describing the parameter χ using soil properties. In recent years, ANN, which can solve the problem by using heuristic knowledge or pattern matching technique, has been widely used in the geotechnical engineering. It has been successfully applied to many geotechnical problems such as stress–strain relationship, slope stability, foundation design, soil classification, design of earth-retaining structures, pile capacity, settlement of foundation and soil improvement [13–19].

In the current study, it was attempted to use the ANN approach in determining parameter χ. For that purpose, a series of unsaturated triaxial tests were performed under consolidated drained condition. In the tests in which compacted soil specimens from Turkey were utilized, the matric suction was controlled by using axis translation technique. Furthermore, 41 shear strength data sets were gathered from the literature. Thirty-six of them were used to train the several compiled ANN models. Remainder of data sets and test results obtained from the present study were used to test the ANN models. In order to deduce effects of basic soil properties on the parameter χ, six different soil properties such as matric suction, angle of shearing resistance, AEV, sand fraction, silt+clay fraction and plasticity index were used as input parameters in the various ANN models.

DETERMINATION OF χ FROM THE UNSATURATED SHEAR STRENGTH TEST

The parameter χ is experimentally determined from the unsaturated shear strength tests. As known, the shear strength equation for saturated soils is as given below:

$$\tau_0=c'+(\sigma-u_w)\tan\phi' \tag{2}$$

In this equation $(\sigma-u_w)$ is the effective normal stress. For the unsaturated soils, Bishop's effective stress equation (Equation (1)) is put into the Equation (2), unsaturated shear strength equation is found as shown below:

$$\tau=c'+[(\sigma-u_a)+\chi(u_a-u_w)]\tan\phi' \tag{3}$$

Note that Equation (3) suggested by Bishop and Blight [20] consists of two parts. The first part of Equation (3) is related to saturated shear strength when the pore-air pressure, u_a, is equal to the pore-water pressure, u_w. This part is a function of normal stress, since the shear strength parameters c' and ϕ' are constants for a saturated soil. The second part of this equation is the contribution of matric suction which is called suction strength (τ_{us}). In other words, the difference between Equations (3) and (2), namely $(\tau-\tau_0)$ is equal to τ_{us} as given below:

$$\tau_{us}=\chi(u_a-u_w)\tan\phi' \tag{4}$$

The parameter χ is derived from Equation (4)

$$\chi=\frac{\tau_{us}}{(u_a-u_w)\tan\phi'} \tag{5}$$

DOI: 10.1002/nag

If the effective cohesion (c') and angle of shearing resistance (ϕ') are independent of suction, χ could be determined using the method described above. Experimental evidences confirming this assumption have been reported by many researchers [7, 21–24]. On the other hand, Vanapalli *et al.* [25] and Fredlund and Rahardjo [26] stated that the angle of shearing resistance (ϕ') can be adopted to be constant although it may slightly increase as the matric suction increases.

EXPERIMENTAL PROCEDURE

A total of 45 unsaturated triaxial tests were conducted on compacted soil specimens from Turkey in an attempt to measure the contribution of matric suction to shear strength. Tests were performed under consolidated and drained condition by controlling the constant matric suction. For that purpose the conventional triaxial test apparatus for saturated soils was modified. The system essentially includes a modified triaxial cell, pressure–volume controllers, pore-water change transducers, diffused air volume indicator (DAVI), plumbing arrangements and data acquisition system. The system enables computer-controlled stress or strain rate testing and can give real-time graphical outputs. The apparatus allows the controlling of matric suction using axis translation technique suggested by Hilf [27]. In this technique described in detail by Fredlund and Rahardjo [26], the pore-air pressure is artificially raised above atmospheric pressure to increase the pore-water pressure by the same amount to positive values so that the cavitation risk of water in the measuring system is prevented. In order to keep the air pressure applied on specimen at a constant value and to drain pore water, a saturated ceramic disk with AEV of 1500 kPa was fixed onto a pedestal connected to the measuring system water compartment. The ceramic disc was sealed on to the pedestal using epoxy resign along its periphery. The grooves inside the water compartment work as water channels to flush the air bubbles accumulated as a result of diffusion. The diffused air volume changes were measured by means of DAVI proposed by Fredlund and Rahardjo [26]. A volume change transducer with a sensitivity of 0.01 cm^3 was connected to the water compartment below the high air entry ceramic disc to measure the volume change of pore water–water content variation.

Following setting up in the modified triaxial cell, each sample was saturated prior to application of matric suction so that initial matric suction of the specimens is reduced to zero. The saturated soil samples were consolidated under a preselected confining pressure of σ_3, while the pore-air and pore-water pressures were controlled at pressures of u_a and u_w, respectively. The volume of water flowing out of the specimen was recorded by the volume change transducer. When there is no longer a tendency for the overall volume change and the flow of water from the specimen, it was assumed that the consolidation reached an equilibrium condition. Upon obtaining a stress and matric suction equalization under applied pressure (i.e. σ_3, u_a and u_w), the soil specimens were sheared by compression at a strain rate of 0.004%/min. Each test at least took 25 days.

In the tests, the soils taken from three different locations of Adana city in Turkey were used. The basic soil properties of specimens are summarized in Table I. The tests were performed on compacted specimens prepared by standard proctor compaction procedure at the optimum water content. For the compaction, the air-dried soils were pulverized and they were passed through a 2-mm sieve. Thereafter they were mixed and wetted to optimum water content and they were cured for at least three days in sealed bags for the moisture equilibrium. The compacted specimens were prepared at dimensions of 50 mm in diameter and 100 mm in height (i.e. length to diameter ratio L/D of 2) to minimize the effects due to end platens of the apparatus and to reduce the

DOI: 10.1002/nag

Table I. Soil properties.

Soil	Specimen types	γ_s (kN/m^3)	(ϕ') (°)	AEV (kPa)	Sand (%)	Silt+clay %	I_p
Adana Havuzlubahçe	Compacted	26.6	28	45	8	92	21
Adana Mıdık	Compacted	25.1	19	100	—	100	32
Adana Şakirpaşa	Compacted	24.2	22	80	6	94	37

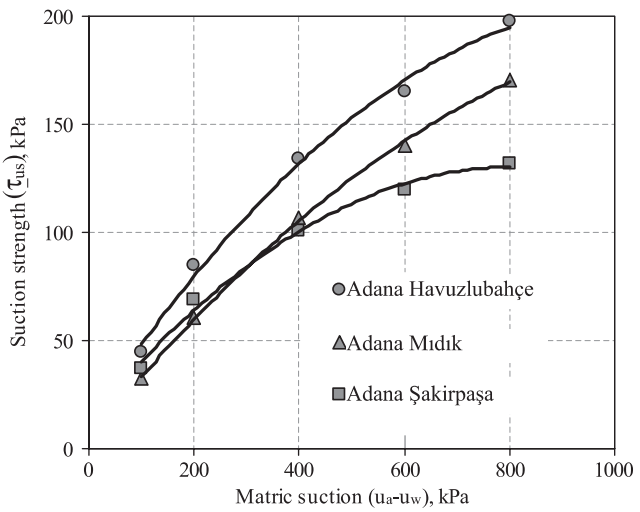

Figure 1. Suction strength *vs* matric suction.

likelihood of buckling during testing. The AEV of the compacted specimens were obtained from the soil–water characteristic curves which were determined from the pressure plate tests.

The test results in terms of suction strength against matric suction are illustrated in Figure 1. It can be seen that the suction strength varies nonlinearly with the matric suction. The variation of the effective stress parameter χ with the matric suction is shown in Figure 2. The parameter χ decreases as the matric suction increases. It was observed that the unsaturated shear tests are time consuming. In order to bring the soil specimens to matric suction equilibrium, long time was needed.

VIEW OF THE ARTIFICIAL NEURAL NETWORK

In the past decade, due to difficulties in solutions of the complex engineering systems, researches have started to study ANN inspired by the behavior of the human brain and the nervous system. Each ANN model can be differently organized according to the same basic structure. There are three main layers in the ANN structure; a set of input nodes, one or more layers of hidden nodes and a set of output nodes. Each layer basically contains a number of neurons working as an independent processing elements and densely interconnected with each other. The neurons using the parallel computation algorithms are simply compiled with adjustable connection weights,

DOI: 10.1002/nag

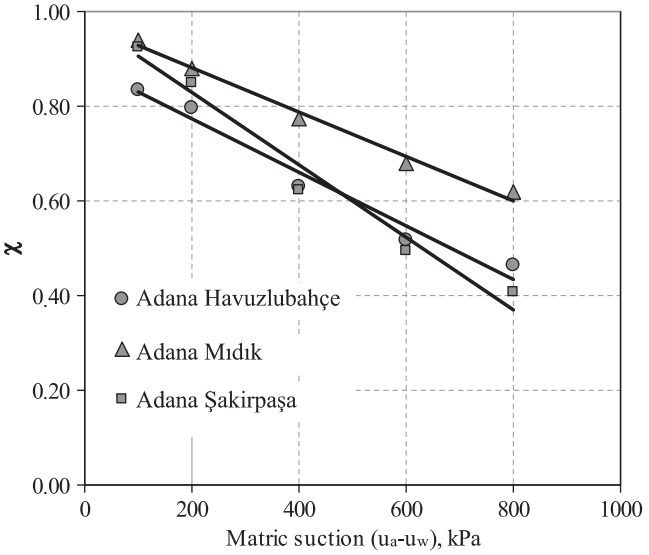

Figure 2. Effective stress parameter *vs* matric suction.

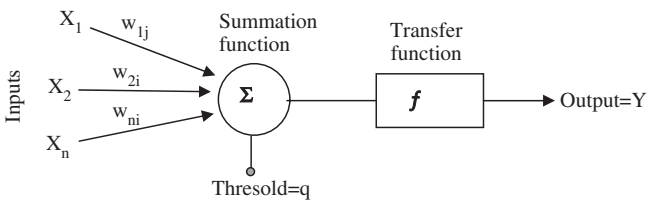

Figure 3. A simple ANN architecture with one neuron.

summation function and transfer function. A simple ANN architecture with one neuron is illustrated in Figure 3. The methodology of ANNs is based on the learning procedure from the data set that presented it from the input layer and testing with other data set for validation. A network is trained by using a special learning function and learning rule. In ANNs analyses, some functions called learning functions are used for initialization, training, adaptation and performance function. During the training process, a network is continuously updated by a training function which repeatedly applies the input variables to a network till a desired error criterion is obtained. Adapt functions are employed for the simulation of a network, while the network is updated for each time step of the input vector before continuing the simulation to the next input. Performance functions are used to grade the network results [28]. In this study, gradient descent with momentum and adaptive learning rate (traingdx), gradient descent with momentum weight and bias learning function (learngdm) and mean-squared error (MSE) were used for training function, adapt function and performance function, respectively. In the learning stage, network initially starts by randomly assigning the adjustable weights and threshold values for each connection between the neurons in accordance

DOI: 10.1002/nag

with the selected ANNs model. After the weighted inputs are summed and the threshold values are added, they are passed through a differentiable nonlinear function defined as a transfer function. This process is continued until a particular input captures their output (i.e. target) or as far as the lowest possible error can be obtained by using an error criterion. In other words, the network training is the determination of weights and biases [29, 30]. An ANN model can be differently composed in terms of architecture, learning rule and self-organization. The most widely used ANNs are the feed-forward, multilayer perceptrons trained by back-propagation algorithms based on gradient descent method (FFBP). This algorithm can provide approximating to any continuous function from one finite-dimensional space to another for any desired degree of accuracy. The superiority of FFBP is that it sensitively assigns the initial weights values and therefore it may yield closer results than the others. Also this algorithm has easier application and shorter training duration.

THE BACK-PROPAGATION PARADIGM

The back propagation is a learning algorithm most widely used in ANNs. The main characteristic of this paradigm is that it works sending inputs forward and then propagating errors calculated using a certain error criteria backwards. In this algorithm, the learning procedure based on supervised rule is continued by adjusting the weights until the minimal error is obtained. A feed-forward network configuration is plotted in Figure 4.

The back-propagation process of error is performed by two steps; the first step is a feed-forward phase in which the output from any node is calculated by propagating the input value given from the input nodes. The second step is a backward phase in which connection weight values are corrected by using error criteria. An output value for any neuron in the hidden layer is computed as follows:

$$Y_j = f\left[\sum_{i=1}^{n} X_i w_{ij} + q_j\right], \quad i = 1, \ldots, n, \quad j = 1, \ldots, m \tag{6}$$

where Y_j is the output value of any neuron in the hidden layer, $X_i = \{X_1, X_2, \ldots, X_n\}$ is the input vector, w_{ij} is the weights from the ith neuron to jth neuron, q_j is a threshold value (bias value), f is the transfer function or activation function and n and m are the numbers of neurons in the input and hidden layers, respectively. The output value of the any output neuron is computed

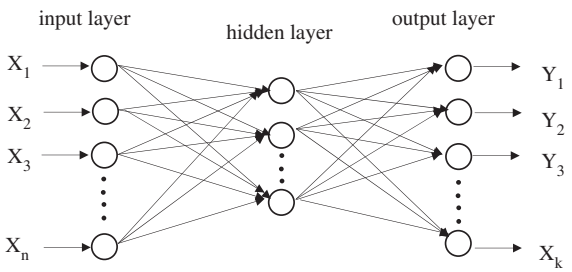

Figure 4. A multilayer ANN configurations.

DOI: 10.1002/nag

as follows:

$$Y_t = f\left[\sum_{j=1}^{m} Y_{jt} w_{ij} + q_t\right], \quad t = 1, \ldots, k \tag{7}$$

where k is the number of neurons in the output layer.

ANNs employ various transfer functions to establish a relationship between the input and output variables at each neuron layer. Log sigmoid and tangent sigmoid given below are the commonly utilized transfer functions. The networks would be able to map nonlinear input–output relationship by means of nonlinearity of these transfer functions:

$$f(x, w) = \frac{1}{(1+\mathrm{e}^{\sum Xw+q})} \quad \text{log-sigmoid} \tag{8}$$

$$f(x, w) = \frac{(1-\mathrm{e}^{\sum Xw+q})}{(1+\mathrm{e}^{\sum Xw+q})} \quad \text{tan-sigmoid} \tag{9}$$

A network propagates the inputs or outputs mapping by using connection weights as far as the lowest possible error can be obtained by using an error criterion. The error is the difference between the calculated output and the target value. MSE for all input patterns is computed from the following equation:

$$E = \frac{1}{N} \sum_{s=1}^{N} \sum_{t=1}^{k} (T_{st} - Y_{st})^2 \tag{10}$$

where N is the number of data set pattern, T_{st} is actual or target value for the sth pattern and Y_{st} is the neural network output value for the sth pattern,

The back-propagation algorithm aims at iteratively minimizing error. According to the gradient descent method, the back-propagation paradigm adjusts the weights by calculating the gradient (δ_t) for each neuron on the output layer with the following equations:

$$\delta_t = Y_t(1-Y_t)(T_t - Y_t) \tag{11}$$

The error gradient δ_j is then recursively determined for the hidden layers by computing the weighted sum of the errors at the previous layer

$$\delta_j = Y_j(1-Y_j) \sum_{t=1}^{k} \delta_t w_{jt} \tag{12}$$

The change in connection weights is updated by using error gradients (Equation (12)) as shown below:

$$\Delta w_{ij}(r) = -\eta \frac{\partial E}{\partial w_{ij}} = \eta \delta_j x_i \tag{13}$$

$$w_{ji}(r+1) = w_{ji}(r) + \Delta w_{ji}(r) \tag{14}$$

The weight change after the tth data presentation is

$$\Delta w_{ji}(r) = \eta \delta_j x_i + \alpha \Delta w_{ji}(r-1) \tag{15}$$

where η is the learning rate, α is the momentum rate and r is the iteration number.

DOI: 10.1002/nag

Back propagation can converge to local minimum sometimes and therefore global minimum may not be reached. To overcome this problem, the momentum rate is used [28]. The other problem in the training stage is overfitting happening during the learning. Although the error of the training may be very small value, when the new data are evaluated by using the same network, the error may be relatively high. The network memorizes the training data, but it does not generalize the new case. To improve the generalization, a method called Bayesian regularization is used. This method uses a performance function modified by adding a term that consists of the mean of the sum of squares of the network weights and biases as follows:

$$\mathrm{MSE}_{\mathrm{reg}} = \gamma \mathrm{MSE} + (1-\gamma)\mathrm{MSw} \tag{16}$$

where MSE is the mean sum of squares of the network errors, MSw is the sum of squares of the networks weights and γ is the performance ratio [28, 31]. One of the effective ways to avoid overfitting is to have as much data as possible. For this purpose one of the criteria is that the number of the training case should be at least 30 times of the weights in the networks. In addition, model selection and early stopping can be used to improve the generalization.

DATABASE

To train and test the ANN models developed in this paper, the shear strength test results for the determination of χ under different matric suctions were collected from various sources published in the unsaturated soil literature. This paper contains data compiled from 41 different unsaturated shear strength tests. Thirty-six of them shown in Table II were used for training the ANN models. Table II shows the index properties, angle of shearing resistance and AEV of soils used for the analysis. Remainder of data set given in Table III and data obtained from tests conducted in this study were used for testing. Data sets used for testing were not included in the training stage.

Development of ANNs models for prediction of parameter χ

The potential of using ANNs for the estimation of parameter χ is investigated by developing different ANN models. It is previously pointed out that the main factor affecting parameter χ is the basic soil properties. This paper has, therefore, taken into consideration five basic soil properties such as angle of shearing resistance (ϕ'), AEV, sand fraction (S), silt+clay fraction ($M+C$) and plasticity index (I_{p}) as input parameters for the ANN models. The experimental data from the published researches were divided into two groups. The first group consisting of 145 data points from 36 different tests in Table II was used to train the network and to develop different ANN models. The second group consisting of 30 data points from five different tests in Table III and current study (this data set was not included in the training stage) was used to test the accuracy of the developed models.

The ANN toolbox of MATLAB computer-aided Software [28] was used to perform the necessary computation. In order to develop the most appropriate ANN architecture, the number of neurons in the hidden layer and a number of multilayer networks with different transfer functions were tried to predict parameter χ. They were, therefore, varied until the convergence was achieved in the MSE. The developed network models for the training are given in Table IV. The feed-forward neural networks that consist of multilayer perceptrons trained back-propagation algorithms were

DOI: 10.1002/nag

Table II. Details of the soil properties and soil type by several published researches.

References	ϕ' (°)	AEV (kPa)	Sand (%)	Silt+clay (%)	I_p
Bishop and Blight [20], Selset Boulder clay	27.3	45*	40	60	17
Cui and Delage [32], Jossigny silt	22	182	4	96	18
De-campos and Carrillo [33], Rio de Janeiro residual soil	28.7	38	69.1	30.9	18.4
De-campos and Carrillo [33], yellow collovium soil	26.4	54	50.3	49.7	22.7
Escario and Saez [34], Madrid grey clay	22.5	200	1	99	36
Escario and Saez [34], red clay of Guadalix de la Sierra	32	100	13.5	86.5	13.6
Escario and Saez [34], Madrid clayey sand	39	95	83	17	17
Gan and Fredlund [35] (saprolitic soil) decomposed fine ash tuff	30	40	40	60	16
Gan *et al.* [36], GT-16-N5 Glacial till	25.5	115	28	72	18.7
Gan *et al.* [36], Glacial till	25.5	153	28	72	18.7
Gan *et al.* [36], Glacial till	25.5	35	28	72	18.7
Gulhati and Satija [37], Dhanauri clay (high density)	28.5	60	5	95	23.5
Huat *et al.* [38], Malaysia residual soil	26	75*	45	55	49
Khalili and Khabbaz [12], compacted kaolin	22	395	—	100	33
Khalili *et al.* [7], SJ10a, Hume Dam in Southeastern Australia	29	95	33	67	21
Khalili *et al.* [7], SJ10b, Hume Dam in Southeastern Australia	29	125	25	75	12
Khalili *et al.* [7], SJ11, Hume Dam in Southeastern Australia	30	200	65	35	6
Krahn *et al.* [39], Notch Hill silt	35	90	5	95	25
Lee *et al.* [19], Okchun soil (wet compaction)	36.8	100*	94.79	4.84	—
Lee *et al.* [19], Okchun soil (dry compaction)	34.6	40	94.79	4.84	—
Lee *et al.* [19], Yungi soil (wet compaction)	41.21	35*	98.44	1.56	—
Lee *et al.* [19], Seochang soil (wet compaction)	36.04	50*	78.71	13.89	—
Lee *et al.* [19], Chochiwon soil (dry compaction)	32.05	85*	93.57	6.43	—
Maatouk *et al.* [40], Trois Rivieres soil	40	85*	16	84	7
Miao *et al.* [22], Nangyang expansive soil	21.3	25	6.7	93.3	31.8
Oloo and Fredlund [41], Indian Head compacted till	22	38	28	72	19
Oloo and Fredlund [41], Botkin compacted silt	28	40	52.5	47.5	6
Pereira and Fredlund [42], Geneiss compacted soil	14.6	5	52	48	12
Rassam and Williams [43], Kidston taillings	40.7	7	75	25	—
Rassam and Williams [44], Kidston taillings	40.7	35.9	75	25	—
Satija [45], Dhanauri clay (low density)	29	25	5	95	23.5
Vanapalli *et al.* [25], Indian Head Glacial till	23	75	28	72	18.7
Vanapalli *et al.* [25], Indian Head Glacial till	23	32	28	72	18.7
Wang *et al.* [46], Botkin silt	31	40	41.8	58.2	15.6
Wheeler and Sıvakumar [47], Speswhite kaolin	25	86	2	98	30
Xu [8], Ningxia expansive soil	24	40	25	75	21.7

*AEV values were obtained from the suction strength *vs* matric suction curves.

DOI: 10.1002/nag

Table III. Database used for testing of the ANN models.

References	ϕ'	AEV	Sand	Silt+clay	I_p
Escario and Juca [48], Madrid gray clay	22.5	200	1	99	36
Lee *et al.* [19], Chochiwon soil (wet compaction)	34.35	60	93.57	6.43	—
Khalili and Khabbaz [12], sand–clay mixture	33.5	115	75	25	—
Rassam and Williams [44], Kidston taillings	40.7	20	75	25	—
Vanapalli *et al.* [25], Indian Head Glacial till	23	110	28	72	18.7

Table IV. The neural network models used in the determination of χ.

Model	Inputs	Structure	Transfer function
I	(u_a-u_w), ϕ', AEV, S, $M+C$, I_p	6–3–1	logsig–logsig
II	(u_a-u_w), AEV, S, $M+C$, I_p	5–4–1	tansig–tansig
III	(u_a-u_w), ϕ', S, $M+C$, I_p	5–4–1	logsig–logsig
IV	(u_a-u_w), ϕ', AEV, I_p	4–4–1	tansig–tansig
V	(u_a-u_w), ϕ', AEV, S, $M+C$	5–3–1	logsig–logsig

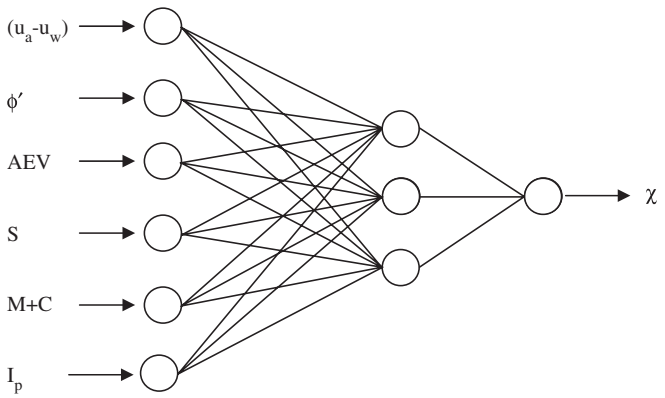

Figure 5. The architecture of feed-forward artificial neural networks with six inputs.

employed for this study. In Figure 5, the architecture of the Model I formed by using feed-forward ANNs with six inputs is shown.

In the ANN analyses, the data set is generally normalized to obtain better convergence. Prior to the training stage, a certain range in which the inputs and target values fall is determined. Thus, the data set used in this study was scaled between 0 and 1 using the following equation [49, 50]:

$$U_{\text{normalized}} = \frac{U_{\text{actual}} - U_{\min}}{U_{\max} - U_{\min}} \tag{17}$$

where $U_{\text{normalized}}$ is the normalized value of the observed variable, U_{actual} is the actual value of the observed variable, $U_{\max}$ is the maximum observation value of the data set and $U_{\min}$ is the minimum

DOI: 10.1002/nag

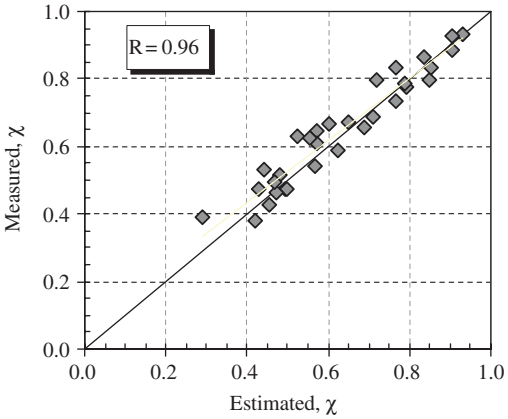

Figure 6. Comparison between the estimated by Model I and measured values of χ.

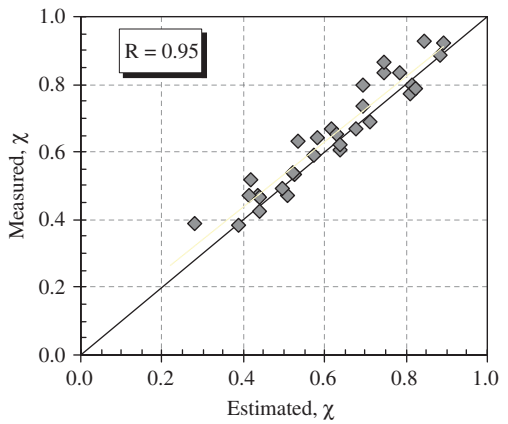

Figure 7. Comparison between the estimated by Model II and measured values of χ.

observation value of the data set. The normalized data set was then used to train neural networks to obtain the final weights. In the end of the analyses, the network outputs were postprocessed to convert the data back into unnormalized units.

RESULTS AND DISCUSSION

The focus of this section of the paper was to examine and discuss the results obtained from the ANN models in Table IV. In Model I, six basic soil properties were presented in the network as input parameters to determine the relationship between soil properties and parameter χ. In order to develop the different ANN models, the input parameters were also individually excluded from the input parameters. As previously mentioned, developed ANN models were tested by data sets

DOI: 10.1002/nag

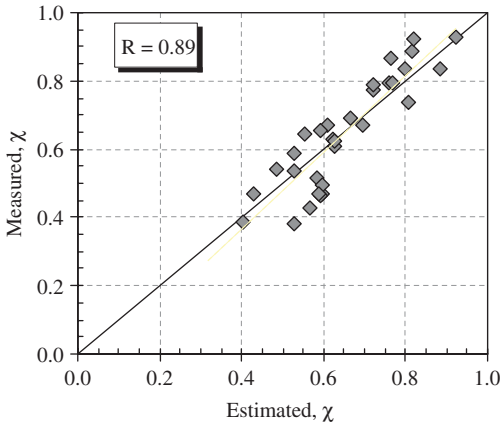

Figure 8. Comparison between the estimated by Model III and measured values of χ.

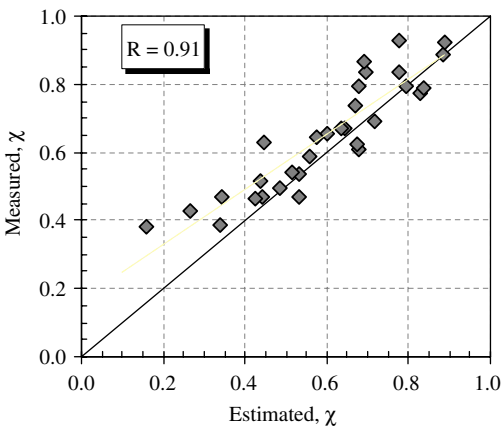

Figure 9. Comparison between the estimated by Model IV and measured values of χ.

from the literature and current study, which were not employed in the training stage. To evaluate how accurate the results of the developed ANN models are, the coefficient of correlation (R), the MSE, mean (μ) and standard deviation (σ) of the ratio and estimated χ_{ANN} to the measured $\chi_{measured}$ ($\chi_{ANN}/\chi_{measured}$) were used as statistical verification tools.

Estimated χ values were graphically compared with the measured χ values in Figures 6–10. As can be seen, the ANN models were found to be able to learn the relationship between the parameter χ and basic soil properties. Table V gives the statistical performance of all the models. It appears that there is a relatively good agreement between the ANN predictions and the actual data. This can be interpreted from the R values ranging from 0.89 to 0.96 in Table V. R values of the models reflect the overall error performances of the models. According to Smith [51], if a proposed model gives $R>0.8$, there is a strong correlation between measured and estimated values over all the data available in the database. The average value and standard deviation are in the

DOI: 10.1002/nag

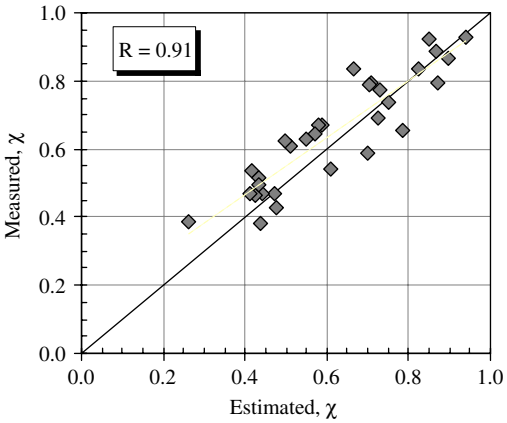

Figure 10. Comparison between the estimated by Model V and measured values of χ.

Table V. Statistical performance of the ANN models.

Model	R	MSE	μ	σ
I	0.96	0.004	0.980	0.091
II	0.95	0.007	0.944	0.105
III	0.89	0.010	1.021	0.147
IV	0.91	0.019	0.891	0.186
IV	0.91	0.011	0.956	0.130

Table VI. Connection weights and biases for Model I.

Weights input layer and hidden layer					
13.7570	−0.8844	−2.6310	−0.2588	−0.3780	−0.8122
−4.7263	−0.7602	10.4190	4.1644	9.2179	−2.8012
5.7394	−12.1160	−4.3082	8.5455	−0.9156	0.6158
Weights hidden layer and output layer					
−7.0362	5.9498	5.9123			
Biases in hidden layer					
1.5289	−3.8330	10.6950			
Biases in output layer					
−5.5192					

range of 0.891–1.021 and 0.091–0.186, respectively. As shown in Table V, Model I with six inputs and Model II excluding ϕ' have fairly high value of R. Model I exhibits a better performance from the standard deviation (σ) and mean (μ) point of view. Model III excluding AEV from the input shows the poor performance in comparison with the other ANN models. The behavior of models was also evaluated in terms of underestimating or overestimating. This evaluation was presented in

DOI: 10.1002/nag

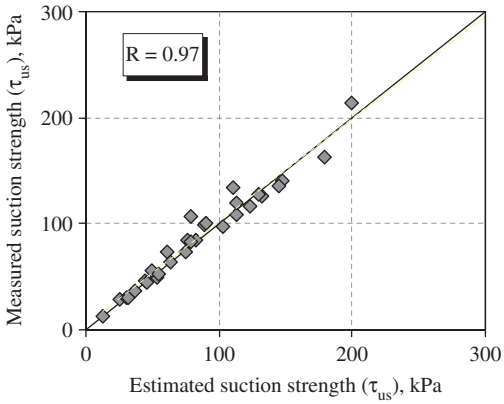

Figure 11. Comparison between estimated from Model I and measured suction strength.

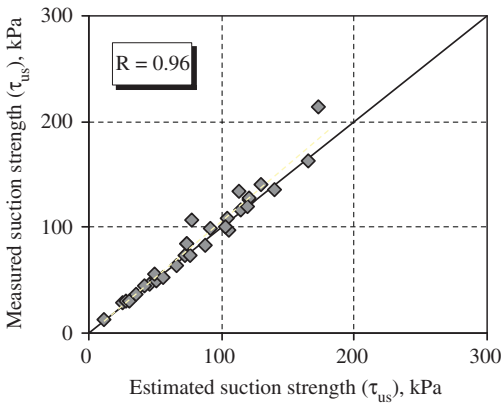

Figure 12. Comparison between estimated from Model II and measured suction strength.

terms of finding out the mean (μ) and the standard deviation (σ) of the ratio $\chi_{ANN}/\chi_{measured}$. The μ value greater than 1.0 means overestimating and less than 1.0 means underestimating. According to Table V Model I appears as the best model as μ values of all models is evaluated. The connection weights and biases for Model I were given in Table VI.

The ANN models developed herein were evaluated with respect to suction strength. Substituting parameter χ estimated from the ANN models into Equation (4), suction strength values were calculated. Figures 11–15 show the comparison of the estimated suction strength values with the measured values. As far as these figures are concerned, Model I which has the highest R appears to be the best model. On the other hand, one can clearly see that other models give good correlation between the estimated and measured suction strength values.

Consequently, when the results in all figures and tables are evaluated, it can be concluded that ANN models can be used for the prediction of parameter χ and suction strength of unsaturated soils. Considering the unsaturated soil tests to be fairly difficult, time consuming and expensive,

DOI: 10.1002/nag

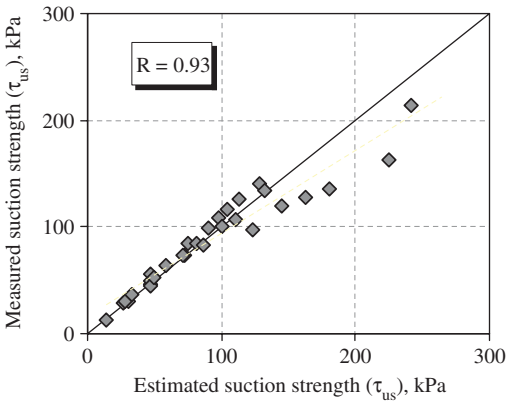

Figure 13. Comparison between estimated from Model III and measured suction strength.

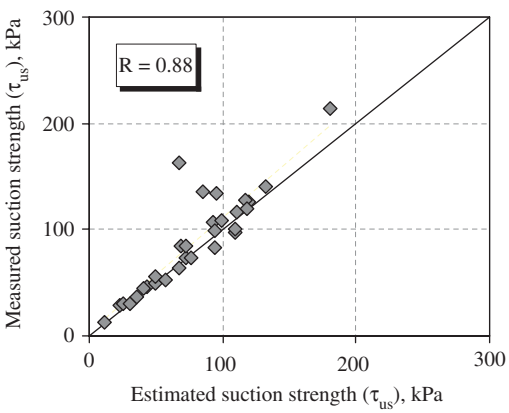

Figure 14. Comparison between estimated from Model IV and measured suction strength.

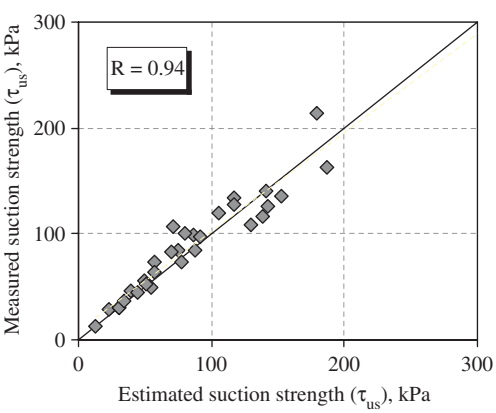

Figure 15. Comparison between estimated from Model V and measured suction strength.

DOI: 10.1002/nag

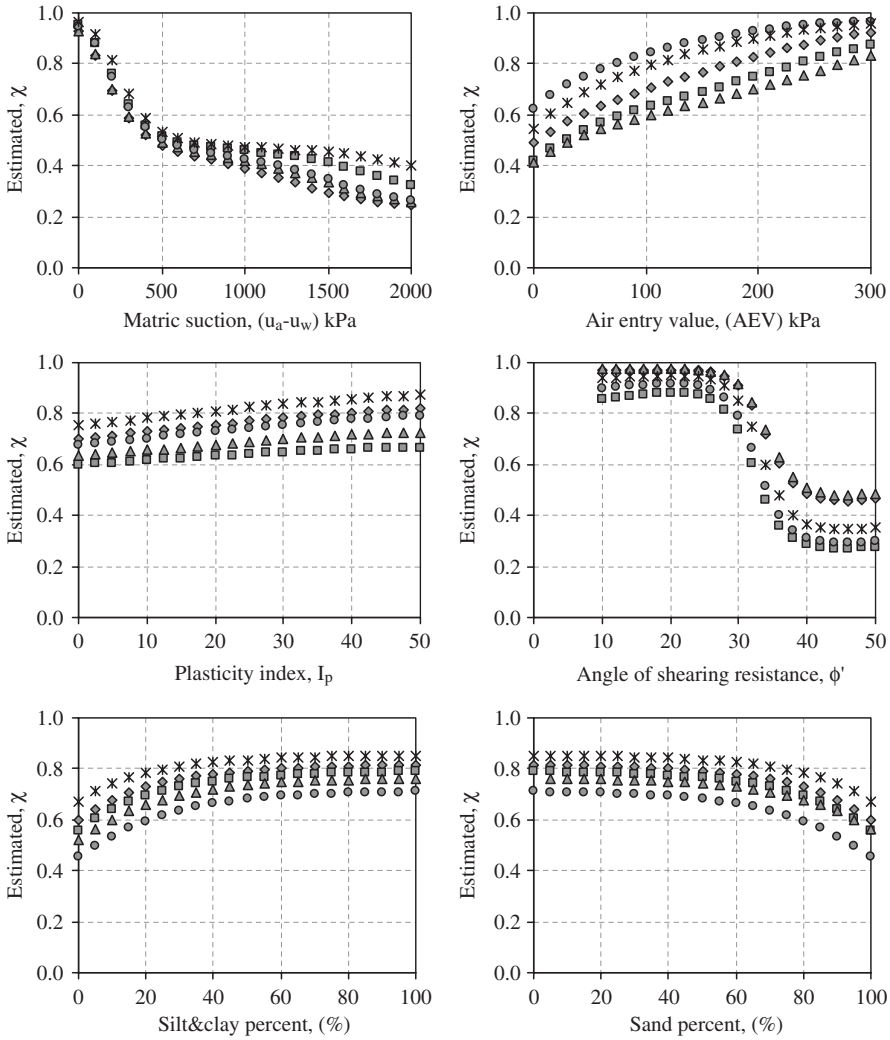

Figure 16. Behavior of Model I with respect to input parameters.

it can be emphasized that the use of ANN in the determination of unsaturated soils parameters is important and reasonable.

Parametric study for the determination of trend of parameter χ

In order to extract how the input parameters individually influence the behavior of Model I, a series of parametric study described by Lee and Lee [18] were performed. Firstly, five random value sets of input parameters were selected. While one parameter was changed in the possible range (i.e. $0<(u_a-u_w)<2000$, $0<\text{AEV}<300$, $0<I_p<50$, $0<\phi'<50$, $0<\text{sand fraction}<100$, $0<\text{silt and clay fraction}<100$), the others were kept constant. Data sets composed in this way

DOI: 10.1002/nag

were presented to Model I and trend of parameter χ was observed. Figure 16 shows the results of the parametric study. Parameter χ decreases when the matric suction $(u_a - u_w)$, angle of shearing resistance (ϕ') and sand fraction increase. On the other hand, it increases when the plasticity index (I_p), silt and clay fraction and AEV increase.

SUMMARY AND CONCLUSION

The main objective of this study was to investigate the applicability of the ANN technique in the determination of effective stress parameter (χ) of unsaturated soils. Furthermore, the suction strength behavior and χ values of compacted soil specimens taken from Adana city in Turkey were determined experimentally for different matric suctions performing unsaturated triaxial tests. The test results showed that the matric suction contributes to the shear strength of unsaturated soils. This contribution called suction strength exhibits nonlinear behavior with respect to the matric suction. It was also observed that parameter χ calculated from the measured suction strengths decreases as the matric suction increases.

Several ANN architectures were compiled to establish an interrelationship between the basic soil properties and parameter χ. The back-propagation algorithm was used in the analyses. The five ANN models which have different input parameters were trained by data collected from 36 different unsaturated tests with 145 data points from the literature and then they were tested by the data collected from five different unsaturated tests as well as the data obtained from the present study. Performances of the models were examined in terms of some statistical criteria. The best results and high correlation coefficient were produced by Model I with six input parameters. It was concluded that the results of ANN models are very encouraging for the cases tested and the use of ANN is a more reasonable method for the prediction of parameter χ when the difficulties and limitations in the testing of unsaturated soils are considered.

REFERENCES

1. Terzaghi K. The shearing resistance of saturated soils and the angle between the planes of shear. *Proceedings of the 1st Conference on Soil Mechanics*, Cambridge, MA, vol. 1, 1936; 54–56.
2. Bishop AW. The principle of effective stress. *Tecknish Ukebland* 1959; **106**(39):859–863.
3. Gray WG, Schrefler BA. Analysis of the solid phase stress tensor in multiphase porous media. *International Journal for Numerical and Analytical Methods in Geomechanics* 2007; **31**(4):541–581.
4. Aitchison GD. Separate roles of site investigation, quantification of soil properties and selection of operational environment in the determination of foundation design on expansive soils. *Proceedings of the 3rd Asian Regional Conference on Soil Mechanics and Foundation Engineering*, Haifa, Israel, vol. 3, 1967; 72–77.
5. Burland JB. Some aspects of the mechanical behaviour of partly saturated soils. In *Proceedings of the Conference on Moisture Equilibria and Moisture Changes in Soil Beneath Covered Areas*, Aitchison GD (ed.). Butterworths: London, 1965; 270–278.
6. Fredlund DG, Morgenstern NR, Widger RA. The shear strength of unsaturated soil. *Canadian Geotechnical Journal* 1978; **15**:313–321.
7. Khalili N, Geiser F, Blight GE. Effective stress in unsaturated soils: review with new evidence. *International Journal of Geomechanics* (ASCE) 2004; **4**(2):115–126.
8. Xu FY. Fractal approach to unsaturated shear strength. *Journal of the Geotechnical and Geoenvironmental Engineering* (ASCE) 2004; **130**(3):264–273.
9. Bishop AW, Henkel DJ. *The Measurement of Soil Properties in the Triaxial Test* (2nd edn). Edward Arnold: London, England, 1962; 227.
10. Oberg AL, Sallfors G. Determination of shear strength parameters of unsaturated silt and sands based on the water retention curve. *Geotechnical Testing Journal* 1997; **20**(1):40–48.

DOI: 10.1002/nag

11. Coleman JD. Stress–strain relations for partially saturated soils. *Géotechnique* 1962; **12**(4):348–350.
12. Khalili N, Khabbaz MH. A unique relationship for χ for the determination of the shear strength of unsaturated soils. *Géotechnique* 1998; **48**(5):681–687.
13. Hadipriono FC, Diaz CF, Wolfe WE. Toward the development of a knowledge base expert system for determining the causes of foundation failures. *Proceedings of the Conference on Computational Structures Technology*. Civil-Comp Press: Edinburgh, 1991.
14. Kulhawy FH, Trautmann CH. Development of an intelligent computer program for foundation design. In *Computer Methods and Advances in Geomechanics*, Beer G, Booker JR, Carter JP (eds). A.A. Balkema: Rotterdam, 1991; 75–80.
15. Davey-Wilson IEG. Evaluation of artificial-intelligence and hypertext approaches to a geotechnical expert-system. In *Information Technology for Civil and Structural Engineers*, Topping BHV, Khan AI (eds). Civil-Comp Press: Edinburgh, 1993; 109–113.
16. Goh ATC. Backpropagation neural networks for modeling complex-systems. *Artificial Intelligence in Engineering* 1995; **9**(3):143–151.
17. Cai YD. The application of artificial neural-network in determining the blasting classification of rocks. *Proceedings of the 2nd International Conference on Engineering Blasting Technique*, Kunming. Peking University Press: People's Republic of China, Beijing, 1995; 24–27.
18. Lee IM, Lee JH. Prediction of pile bearing capacity using artificial neural networks. *Computers and Geotechnics* 1996; **18**(3):189–200.
19. Lee SJ, Lee SR, Kim YS. An approach to estimate unsaturated shear strength using artificial neural network and hyperbolic formulation. *Computers and Geotechnics* 2003; **30**:489–503.
20. Bishop AW, Blight GE. Some aspects of effective stress in saturated and partly saturated soils. *Géotechnique* 1963; **13**(3):177–197.
21. Tekinsoy MA, Kayadelen C, Keskin S, Söylemez M. An equation for predicting shear strength envelope with respect to matric suction. *Computers and Geotechnics* 2004; **31**(7):589–593.
22. Miao L, Liu S, Lai Y. Research of soil–water characteristics and shear strength features of nanyang expansive soil. *Engineering Geology* 2002; **65**:261–267.
23. Drumright EE, Nelson JD. The shear strength of unsaturated tailings sand. In *The 1st International Conference on Unsaturated Soils*, Paris, Alonso EE, Delage P (eds). A.A. Balkema: Rotterdam, 1995; 45–50.
24. Rohm SA, Vilar OM. *Shear Strength of Unsaturated Sandy Soil*, Alonso EE, Delage P (eds). A.A. Balkema: Rotterdam, 1995; 189–195.
25. Vanapalli SK, Fredlund DG, Pufahl DE, Clifton AW. Model for the prediction of shear strength with respect to soil suction. *Canadian Geotechnical Journal* 1996; **33**:379–392.
26. Fredlund DG, Rahardjo H. *Soil Mechanics for Unsaturated Soils*. Wiley: New York, 1993.
27. Hilf JW. An investigation of pore water pressure in compacted cohesive soils. *Technical Memorandum No. 654*, Bureau of Reclamation. US Department of Interior, 1956.
28. Demuth H, Beale M. *Neural Network Toolbox for Use with MATLAB*. The MathWorks Inc.: Natick, MA, 2001; 840.
29. Basma AA, Kallas N. Modeling soil collapse by artificial neural networks. *Geotechnical and Geological Engineering* 2004; **22**:427–438.
30. Moosavi M, Yazdanpanah MJ, Doostmohammadi R. Modeling the cyclic swelling pressure of mudrock using artificial neural networks. *Engineering Geology* 2006; **87**:178–194.
31. Foresee FD, Hagan MT. Gauss–Newton approximation to Bayesian learning. *Proceedings of the International Joint Conference on Neural Networks*, Houston, TX, U.S.A., 1997; 1930–1935.
32. Cui YJ, Delage P. On the elasto-plastic behaviour of an unsaturated silt. *Geotechnical Special Publication*, vol. 39. ASCE: New York, 1993; 115–126.
33. De-Campos TMP, Carrillo CW. Direct shear testing on an unsaturated0 soil from Rio de Janeiro. *Proceedings of the 1st International Conference on Unsaturated Soils*, Paris, 1995; 31–38.
34. Escario V, Saez J. The strength of partly saturated soils. *Géotechnique* 1986; **36**(3):453–456.
35. Gan JKM, Fredlund DG. Shear strength characteristics of two saprolitic soils. *Canadian Geotechnical Journal* 1996; **33**:595–609.
36. Gan JKM, Rahardjo H, Fredlund DG. Determination of the shear strength parameters of an unsaturated soil using the direct shear test. *Canadian Geotechnical Journal* 1988; **25**:500–510.
37. Gulhati SK, Satija BS. Shear strength of partially saturated soils. *Tenth International Conference on Soil Mechanics and Foundation Engineering*, Stockholm, Sweden, vol. 1.1, 1981; 609–612.

DOI: 10.1002/nag

38. Huat BBK, Ali FH, Abdullah A. Shear strength parameters of unsaturated tropical residual soils of various weathering grades. *Electronic Journal of Geotechnical Engineering* 2005; **10**.
39. Krahn J, Fredlund DG, Klassen MJ. Effect of soil suction on slope stability at Notch Hill. *Canadian Geotechnical Journal* 1989; **26**(2):269–278.
40. Maatouk A, Leroueil S, La Rochelle P. Yielding and critical state of a collapsible unsaturated silty soil. *Geotechnique* 1995; **45**(3):465–477.
41. Oloo SY, Fredlund DG. A method for determination of ϕ^b for statically compacted soils. *Canadian Geotechnical Journal* 1996; **33**:272–280.
42. Pereira JHF, Fredlund DG. Shear strength behavior of a residual soil of gneiss compacted at metastable structured conditions. *Proceedings of the Eleventh Pan American Conference on Soil Mechanics and Geotechnical Engineering*, Iguanzu Falls, Brazil, vol. l, 8–12 August 1999; 369–377.
43. Rassam DW, Williams DJ. Shear strength of unsaturated gold tailings. *Proceedings*, *1st Australia New Zealand Conference on Environmental Geotechnics*, A.A. Balkema: Rotterman, 1997; 469–474.
44. Rassam DW, Williams DJ. Bearing capacity of desiccated tailings. *Journal of Geotechnical and Geoenvironmental Engineering* (ASCE) 1999; **125**(7):600–609.
45. Satija BS. Shear behavior of partly saturated soils. *Ph.D. Thesis*, Indian Institute of Technology, New Delhi, 1978.
46. Wang Q, Pufahl DE, Fredlund DG. A study of critical state on an unsaturated silty soil. *Canadian Geotechnical Journal* 2002; **39**:213–218.
47. Wheeler SJ, Sivakumar V. An elasto-plastic critical state framework for unsaturated soil. *Geotechnique* 1995; **45**(1):35–53.
48. Escario V, Juca J. Strength and deformation of partly saturated soils. *Proceedings*, *12th International Conference on Soil Mechanics and Foundation Engineering*, Rio De Janeiro, vol. 11, 1989; 43–46.
49. Rafiq MY, Bugmann G, Easterbrook DJ. Neural network design for engineering applications. *Computers and Structures* 2001; **79**:1541–1552.
50. Sivrikaya O. Comparison of ANN models with correlative works on undrained shear strength (Su). *Eurasian Soil Science* 2007, accepted.
51. Smith GN. *Probability and Statistics in Civil Engineering*. Collins: London, 1986.

DOI: 10.1002/nag