

Discussion Session 5:

Modelling (rheological, mathematical, mechanical models)

Modélisation (modèles rhéologiques, mathématiques, mécaniques)

Neural network based prediction of silty sands behavior

S.S. Yasrobi & M. Banimahd

Tarbiat Modarres University, Tehran, Iran

ABSTRACT: The behavior of silty sands under triaxial compression test was modelled using Multi-Layer Perceptrons (MLPs), a well known type of Artificial Neural Networks. This model represents the effects of fines shape, fines content, grain size distribution, confining pressure, stress–strain path and relative density on deviator stress and pore water pressure variation of silty sands. The results of relatively large number of monotonic undrained triaxial compression tests on silty sands with different fine contents and fine shapes were used in this study. To choose the best structure of model, single and dual output networks were developed and effect of number of hidden processing units on the performance of networks was studied. Furthermore, an early stopping technique was utilized in training phase to improve generalization of the model. The good agreement between measured data and modelling results demonstrates the desired capability of MLP in representing complex stress–strain behavior of silty sands.

1 INTRODUCTION

The stress–strain relationships of soils have been recognized to be important factors in the analysis of load–deformation problems. During last four decades, many mathematical models have been introduced to approximate experimentally observed behavior of different soils. To develop a mathematical model after making some assumptions based on adopted theory (e.g. Elasticity, Plasticity) and employing yielding and failure criteria, mathematical expressions are formulated. Then the model parameters are determined by an inverse analysis from the results of laboratory and in-situ tests before using in a numerical code. The strength of such models depends on their complexities (Faruque 1987). In other word, more complex a model is, more accuracy is expected, but at the same time it requires more parameters to be used (Duncan & Chang 1970). In spite of considerable complexities of these constitutive models, because of inadequate understanding of the subject and all factor involved, it is not possible to capture the material response along all complex stress path, densities and confining pressures. Furthermore the complexities of constitutive models, in many cases, inhibit their incorporation in general purpose of numerical codes, thus restricting their usefulness in engineering practice (Shin & Pande 2000).

A fundamentally different approach, Artificial Neural Network (ANN) has been used recently in geo-material behavior modelling. ANN is over simplified

simulation of human brain. It is able to learn and generalize from experimental data even if they are noisy and imperfect. This ability allows this computational system to learn complex constitutive relationships of materials directly from the results of experiments. Therefore in opposite of conventional models it needs no prior knowledge, constant and assumption about deformation of geomaterials.

Up to now, to the authors' best knowledge, ANN has been applied to constitutive modelling of rocks (Millar & Calderbank 1995), clays (Penumadu et al. 1994), clean sands (Ellis et al. 1995, Penumadu & Zhao 1999, Ghaboussi & Sidatra 1998, Najjar & Zhang 2000), gravels (Penumadu & Zhao 1999) and residual soils (Zhu et al. 1998). In present study, Multi-Layer Perceptron was used to predict undrained behavior of silty sands. A silty sand is considered as a delicate composite matrix containing a sand-grain-matrix and a silt-fine-matrix (Thevanayagam 1998). It has been indicated that fines have an important effect on shear strength of silty sands. In general as the fines content increases, initially the steady state strength at the same void ratio decreases, followed by an increase in shear strength with further increase in fines content beyond about 30% (Pitman et al. 1994, Zlatovic & Ishihara 1995). The similar trend has also been observed at a given relative density (Singh 1995, Thevanayagam 1998). Moreover Yasrobi (1996) investigated the effect of fines shape on undrained behavior of silty sands. He stated that angularity of fine particles decreases

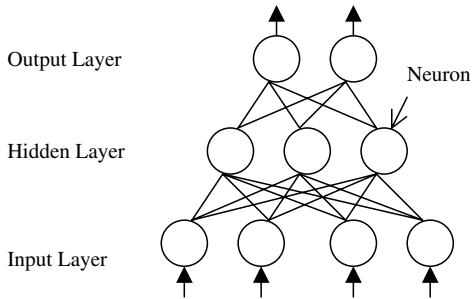


Figure 1. Typical MLP structure with one hidden layer.

the potential of limited strain softening behavior whereas it increases the potential of strain softening behavior. Regarding to the role of fines, in the proposed model percentage of fines and shape of fines are taken into account as well as natural properties of sands, effective confining pressure and relative density of soils.

2 MULTI-LAYER PERCEPTRON (MLP)

MLP is composed of simple processing units referred to as neurons, which are arranged in layers: input, output and one or more hidden layers. A typical MLP consisting of three layers of neurons is shown in Figure 1. Each neuron in a layer is connected to all neurons of next layer via weighted connections. A neuron receives the weighted inputs from neurons of previous layer. These weighted inputs are summed and passed through a transfer function to produce neuron output.

MLP is capable to represent mapping from one multivariate space of information to another after training by a set of data representing that mapping (Garett 1994). During training process, the inputs are presented to MLP; they are weighted and processed as discussed before to produce network outputs. The outputs are compared with desired values, error is calculated and according to this error, the weights are adjusted. The training process is repeated until error reduces to an acceptable criterion. Once training is finished the weights are frozen and the calibrated MLP is ready to compute outputs corresponding to inputs of trained domain.

3 DATABASE

A total number of 129 monotonic strain-controlled triaxial tests under undrained conditions were used in the database. The tests were performed by Yasrobi (1996) on loose to medium dense samples of Toyoura sand, its mixture with different percentage of angular and round fines and Astaneh silty sand to study the effect of shape

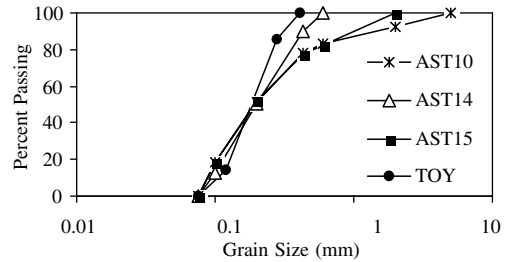


Figure 2. Grain size distributions of sands used in database.

and percentage of fines on steady state strength of silty sands. Toyoura, Japanese standard sand, is a clean uniform sand with subround to subangular grains. To make angular silt, Toyoura sand was crushed and passed through #200 sieve (0.074 mm). Also, as round fines, glass spheres of smaller than 0.074 mm were used. These non plastic fines were mixed with Toyoura sand to make a series of mixture with 10%, 16.7% and 30% of angular fines (TS1, TS2 and TS3) and the same percentages of round fines (TG1, TG2 and TG3). Another group of silty sands was taken from shallow depth (up to 2 m) of three boreholes (AST10, AST14 and AST15) in Astaneh, a city in the north of Iran, which had experienced dramatic damage due to liquefaction during earthquake 1990. These soils have subround to subangular sand grains and fines content ranging from 18% to 21% by weight. Different sand grain size distributions used in this database are shown in Figure 2.

Wet tamping method was used to prepare specimens of 50 mm in diameter and 100 mm in height. The saturated samples were consolidated by gradual increase of cell pressure to desired value of confining pressure between 0.02 to 0.32 Mpa. Then the samples with relative densities ranging from 12% to 60% were axially strained at 1 mm per min and 0.4 mm per min for clean and silty sands respectively. Axial strain (ϵ), deviator stress (q), pore water pressure (u) and confining pressure (σ_3) were recorded every second by a data acquisition system. The tests were terminated once they reached to steady state. Physical and index properties of different test groups such as maximum and minimum void ratio (e_{max} , e_{min}) and G_s , which were measured according to Japanese standard (JSSMFE), are summarized in Table 1.

In present study, generally deviator stress–strain and pore water pressure–strain curves were processed incrementally up to the strain corresponding to steady state. However, for the samples, which had experienced steady state after the strain level of 20%, the curves were considered until the strain of 20%. Using constant strain increment, the stress states corresponding to increments should be interpolated from experimental data. So, to prepare patterns representing the curves

Table 1. Physical and index properties of test groups.

Test groups	G_s	e_{max}	e_{min}	Fines content (%)	Fines shape
TOY	2.646	0.965	0.615	0	–
TS1	2.646	0.918	0.557	10	–
TS2	2.646	0.887	0.517	16.7	A*
TS3	2.646	0.905	0.451	30	–
TG1	2.628	0.781	0.571	10	–
TG2	2.618	0.823	0.523	16.7	R**
TG3	2.596	0.715	0.499	30	–
AST10	2.766	1.097	0.669	18	SA***
AST15	2.778	1.158	0.703	21	SR****
AST14	2.747	1.081	0.633	23	SA

*Angular; **round; ***subangular; ****subround.

directly from original data, a varying strain increment between 0.25% and 0.30% was used. A total number of 7475 patterns were created for training, validating and testing of the model.

4 MODEL DEVELOPING

In order to model stress–strain curves up to strain of 20%, three layer perceptrons with one and two outputs were developed. A three-layer perceptron with differentiable transfer function and sufficient number of neurons in hidden layer can approximate any nonlinear relationship (Hornik 1989). Details of MLPs (A1, A2 and B1) which were used in this study are presented in Table 2. As transfer functions, for A1, a tan-sigmoid function and a log-sigmoid function were used in hidden and output layers respectively. For A2 and B1, tan-sigmoid functions were used in both layers. Inputs are relative density of samples after initial consolidation (D_r), effective confining pressure (σ'_3), coefficient of curvature (C_c) and coefficient of uniformity (C_u) of sands, fine shape index (I_s), fine percentage by weight (P_f) and current state of stress–strain: deviator stress (q_i), pore water pressure (u_i), strain (ϵ_i) and strain increment ($\Delta\epsilon_i$). Outputs are the next state of stress (q_{i+1} , u_{i+1}). Fine shape index (I_s) is equal to: 0 for angular fines and clean sands, 0.25 for subangular, 0.5 for subangular to subround, 0.75 for subround and 1 for round fines. Since shape and mean size of sand grains in database were relatively the same, they were not considered as inputs of the model. Boundaries of inputs and outputs are given in Table 3.

Matlab 6.12, a popular numeric computation and visualization software, and its neural network toolbox were used to train and test MLPs. For training, the Levenberg-Marquardt algorithm was used. This method, which is an approximation of Newton's method, has been shown to be one of the fastest algorithms for

Table 2. Inputs and outputs of developed networks.

MLPs	Inputs	Output(s)
A1	$\Delta\epsilon_i, \epsilon_i, q_i, C_c, C_u, D_r, \sigma'_3, P_f, I_s$	q_{i+1}
A2	$\Delta\epsilon_i, \epsilon_i, u_i, C_c, C_u, D_r, \sigma'_3, P_f, I_s$	u_{i+1}
B1	$\Delta\epsilon_i, \epsilon_i, q_i, u_i, C_c, C_u, D_r, \sigma'_3, P_f, I_s$	q_{i+1}, u_{i+1}

Table 3. Boundaries of inputs and outputs of the model.

Limits	Inputs				Output(s)
	ϵ_i (%)	ϵ_i (%)	C_c	C_u	q_{i+1} (MPa)
Max	0.35	19.98	1.05	3.11	0.577
Min	0.25	0	0.73	1.74	0

Limits	Inputs				Output(s)
	D_r (%)	P_f (%)	I_s	σ'_3 (MPa)	u_{i+1} (MPa)
Max	62	30	1	0.314	0.212
Min	12	0	0	0.031	-0.114

training moderate size MLPs (Hagan & Menhaj 1994). To improve generalization of MLPs, training should be stopped when overfitting is started. Overfitting makes MLPs to memorize training patterns in such a way that it can not generalize well to new data. In present study, cross validation technique was used as the stopping criterion. In this technique database is divided to three sets: training, validation and testing. Training set is used to update networks weights. During this process, the error on validation set is monitored. When the error on validation set begins to increase, the training should be stopped because it is considered as the best point of generalization. Finally, testing data is fed into the networks to evaluate its performance. In this study, the results of 107, 10 and 12 triaxial tests were used for training, validation and testing of networks respectively. It should be noted that the trained MLPs in this study build stress–strain curves incrementally. It is simply done by starting at a stress-strain free state ($\epsilon = 0$, $q = 0$ and $u = 0$) and using MLPs to estimate next stress state. Then this stress state is fed back to network to generate another stress states which are themselves inputs for approximating new stress states. This loop is continued up to desired strain level less than 20%.

In order to choose optimum number of hidden neurons, the performance of MLPs with hidden neurons varying from 3 to 25 was studied regarding to error measures: correlation coefficient (R) and Mean Absolute Error (MAE). For initial assessment, these measures on validation set, a part of unseen data which controls training process, were considered to represent both simulation ability (against seen data) and prediction ability (against unseen data) of networks.

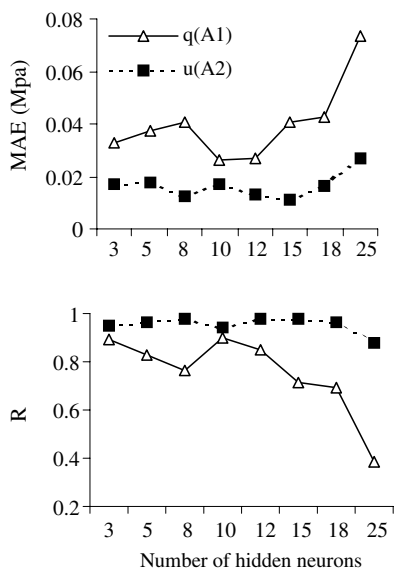


Figure 3. Effect of number of hidden neurons on performance of networks A1 and A2 over validation set.

Therefore, in the first stage the most promising networks were filtered out using the error measures on validation set. Figure 3 presents the error measures on validation set for A1 and A2. It can be seen that A1 with 3 and 10 hidden neurons gave good result while satisfying performance was observed for all A2 networks except the one with 25 hidden neurons. In second stage, performance of the most promising networks was monitored on testing and training set. The absolute optimum network with the best overall performance on seen and unseen data was selected. Following this procedure, 10 and 15 hidden neurons were chosen for A1, A2 respectively. For B1, the network with 15 hidden neurons showed the best overall performance.

In Table 4, error measures of single networks (A1 and A2) and dual output network (B1) are compared. From the table, it can be seen that generally single output networks yielded better MAE and R on validation and testing sets while the error measures of single and dual output networks on training set were relatively the same. So as final structure of model, single output networks (A1 and A2) with equal simulation ability (against training set) and better prediction ability (against validation and testing set) were selected. The minimum correlation coefficients of A1 and A2 on all three sets are 0.90 and 0.94. According to Smith (1986) explanation for $R > 0.8$, the proposed model represents strong correlation between measured and estimated values over all data available in the database.

Table 4. Comparison of single and dual output networks.

Error measures	MLPs	Training set		Validation set		Testing set	
		q	u	q	u	q	u
R ($\times 10^{-2}$)	A1	92	—	90	—	90	—
	A2	—	94	—	98	—	96
	B1	92	95	84	96	81	96
MAE* ($\times 10^{-3}$)	A1	25	—	26	—	38	—
	A2	—	14	—	11	—	14
	B1	25	13	26	23	37	17

* In Mega Pascal (MPa).

5 SIMULATIONS AND PREDICTIONS

Networks A1 and A2 were used in simulating 107 test results of training set. A typical comparison of model simulations with laboratory results for 4 tests is shown in Figure 4. Another part of laboratory results, which form validation and testing set, was used to examine predictability of model. Figure 5 and Figure 6 represent model predictions associated with experimental results for two tests of each set. As it can be seen from the figures, model results on different strain levels are in good agreement with measured values. Furthermore all three different observed behavior of silty sands, strain softening, limited strain softening and strain hardening are captured well by the model. This model also represents the effect of dilative and contractive behavior of silty sands on the variation of pore water pressure reasonably well. It is worthy to mention that the observed deviation between predicted and measured stress–strain curves is not only due to model deficiency but partly due to uncertainties, errors and inconsistency of data used for training and testing model. Unfortunately, as stated by Najjar & Zhang (2000), it is impossible to differentiate between deficiencies relating to modelling methodology and those relating to experimental inconsistencies and inaccuracies.

6 CONCLUSION

Herein, a model based on two single output MLPs was introduced to estimate the variation of deviator stress and pore water pressure with axial strain in undrained triaxial test condition for silty sands. The model generates deviator stress–strain and pore water pressure–strain curves incrementally up to strain level of 20%, starting from a free stress–strain state. To develop the model, an attempt was made to choose the best structure of networks regarding to number of hidden and output neurons. Moreover training process was controlled using an early stopping technique to

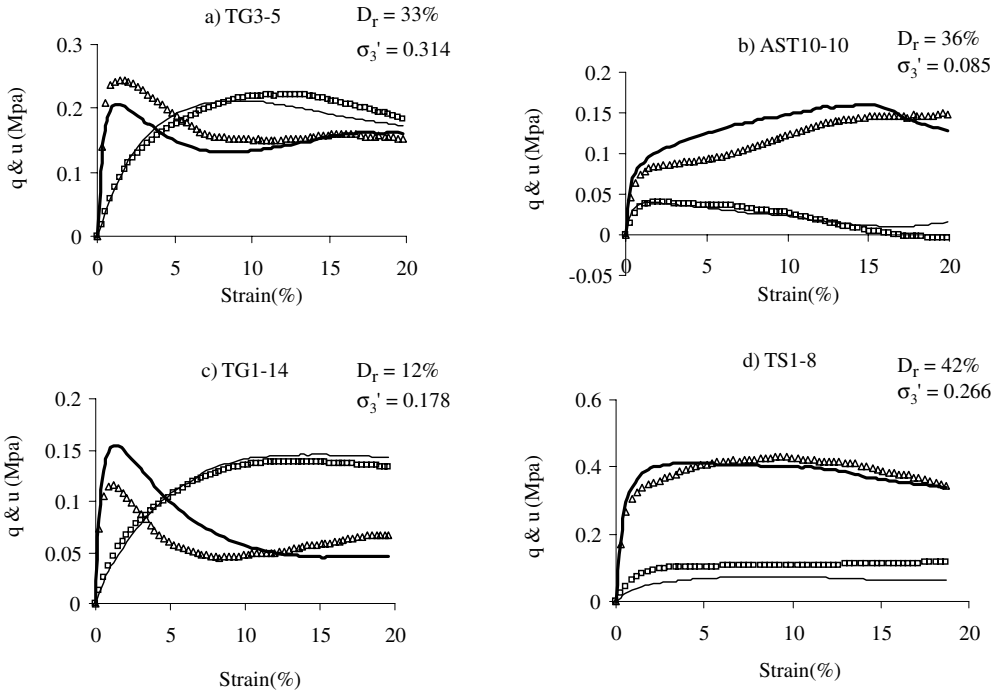


Figure 4. Simulation ability of the model against training data (— q (Lab.); Δ q (A1); — u (Lab.); \square u (A2)).

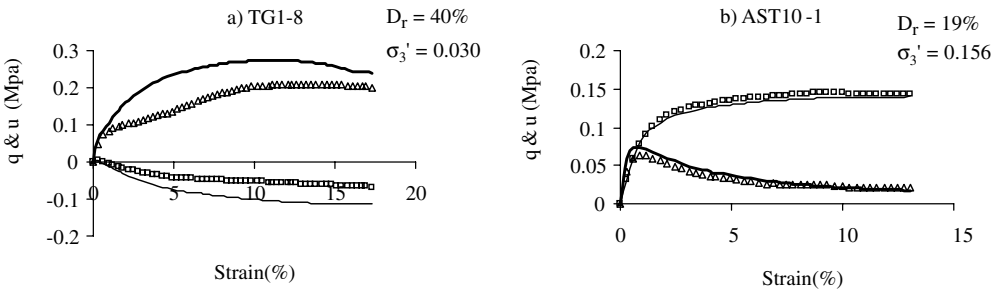


Figure 5. Prediction ability of the model against validation data (— q (Lab.); Δ q (A1); — u (Lab.); \square u (A2)).

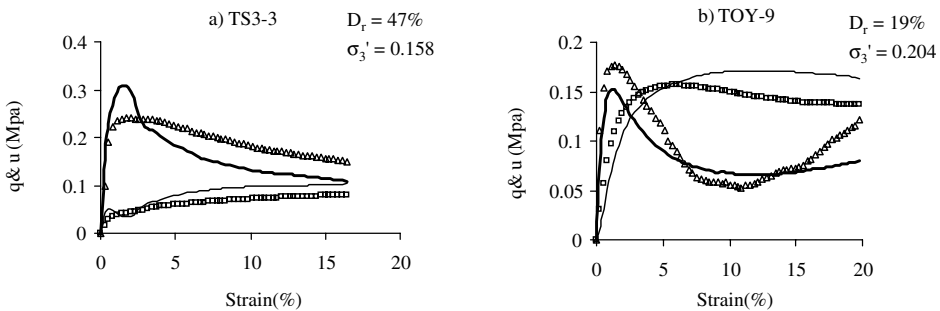


Figure 6. Prediction ability of the model against testing data (— q (Lab.); Δ q (A1); — u (Lab.); \square u (A2)).

avoid overtraining. Unlike conventional constitutive models, no assumption was made and no constant should be determined. The model characterizes behavior of silty sands simply using natural properties, the density and stress level of them. Overall, the proposed modelling methodology in this study is efficient in predicting and simulating the complex stress-strain behavior of such soils as judged by error measures and practical observation.

REFERENCES

- Duncan, J. M. & Chang, C. Y. 1970. Nonlinear analysis of stress and strain in soils. *J. Soil mechanics and foundation division* 96(5M5): 1629–1653.
- Ellis, G. W., Yao, C., Zhao, R. & Penumadu, D. 1995. Stress-strain modelling of sands using artificial neural networks. *J. Geotech. Engrg-ASCE* 43: 195–219.
- Faruque, M. O. 1987. A third invariant dependent cap model for geological materials. *Soils and foundations* 27: 12–20.
- Garett, Jr, J. H. 1994. Where and why artificial neural networks are applicable in civil engineering. *J Comp. Civil Engrg-ASCE* 8(2): 129–130.
- Hagan, M. T. & Menhaj, M. B. 1994. Training feedforward networks with marquardt algorithm. *IEEE transaction on neural networks* 5 (6): 989–992.
- Hornik, K. M., Stinchcombe, M. & White, H. 1989. Multilayer feedforward networks are universal approximator. *Neural networks* 2(5): 359–366.
- Millar, D. L. & Calderbank, P. A. 1995. On the investigation of a multilayer feedforward neural network model of rock deformability behavior. In Fujii (ed.), *Rock mechanics*; Int. Congress, Tokyo: 933–938. Rotterdam: Balkema.
- Najjar, Y. & Zhang, X. C. 2000. Characterizing the 3d stress-strain behavior of sandy soils: a neuro mechanistic approach. In G. M. Filz & D. V. Griffiths (eds.), *Numerical methods in geotechnical engineering*; Geotech. Spec. Pub. 96: 43–57. Virginia: ASCE.
- Penumadu, D., Jin-Nan, L., Chameau, J. L. & Arumugam, S. 1994. Rate dependent behavior of clays using neural networks, *Int. Soc. Soil mechanics and foundation engineering*; 13th Conf., New Delhi: 1445–1448. Oxford & IBH Publ. Co.
- Penumadu, D. & Zhao, R. 1999. Triaxial compression behavior of sands and gravels using artificial neural networks (ANN). *Computers and geotechnics* 24: 207–230.
- Pitman, T. D., Robertson, P. K. & Sego, D. C. 1994. Influence of fines on the collapse of loose sands. *Can. Geotech. J.* 31: 728–739.
- Shin, H. S. & Pande G. N. 2000. On self-learning finite element codes based on monitored response of structures. *Computers and geotechnics* 27: 161–178.
- Singh, S. Re-examination of the effects of fine contents on the liquefaction characteristics of sands. In K. Ishihara (ed.), *Earthquake geotechnical engineering*; Proc. Int. Conf., Tokyo: 829–831. Rotterdam: Balkema.
- Smith, G. N. 1986. *Probability and statistics in civil engineering*. London: Collins.
- Thevanayagam, S. 1998. Effect of fines and confining stress on undrained shear strength of silty sands. *J. Geotech. Geoenviron. Engrg.-ASCE* 124(6): 479–491.
- Yasrobi, S. S. 1996. Effect of amount and shape of fines on steady state behavior of sandy soils. PhD dissertation, Sharif University of Technology, Tehran, Iran.
- Zhu, J. H., Zaman, M. M. & Anderson, S. A. 1998. Modelling of shearing behavior of a residual soil with recurrent neural network. *Int. J. Num. Anal. Meth. Geomech* 22: 671–687.
- Zlatovic, S. & Ishihara, K. 1995. On the influence of non-plastic fines on residual strength. In K. Ishihara (ed.), *Earthquake geotechnical engineering*; Proc. Int. Conf., Tokyo: 239–244. Rotterdam: Balkema.